Pós-Graduação em Ciência da Computação

ALIXANDRE THIAGO FERREIRA DE SANTANA

ENTERPRISE ARCHITECTURE ANALYSIS BASED ON NETWORK PARADIGM: A FRAMEWORK PROPOSAL AND EMPIRICAL EVALUATION

Universidade Federal de Pernambuco
posgraduacao@cin.ufpe.br
www.cin.ufpe.br/~posgraduacao

RECIFE
2017

Alixandre Thiago Ferreira De Santana

Enterprise Architecture Analysis based on Network Paradigm: A Framework Proposal and Empirical Evaluation

A Ph.D. Thesis presented to the Center of Informatics of Federal University of Pernambuco in partial fulfillment of the requirements for the degree of Philosophy Doctor in Computer Science.

ADVISOR: Hermano Perrelli de Moura

RECIFE
2017

Catalogação na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

S232e Santana, Alixandre Thiago Ferreira de
Enterprise architecture analysis based on network paradigm: a framework proposal and empirical evaluation / Alixandre Thiago Ferreira de Santana – 2017.
370 f.: il., fig., tab.

Orientador: Hermano Perrelli de Moura.
Tese (Doutorado) – Universidade Federal de Pernambuco. CIn, Ciência da Computação, Recife, 2017.
Inclui referências e apêndices.

1. Engenharia de software. 2. Gerenciamento de projetos. 3. Análise de redes. I. Moura, Hermano Perrelli de (orientador). II. Título.

005.1 CDD (23. ed.) UFPE- MEI 2017-110

Alixandre Thiago Ferreira de Santana

ENTERPRISE ARCHITECTURE ANALYSIS BASED ON NETWORK PARADIGM: A Framework Proposal and Empirical Evaluation

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutora em Ciência da Computação

Aprovado em: 15/02/2017.

BANCA EXAMINADORA

____

Prof. Dr. Alexandre Marcos Lins de Vasconcelos
Centro de Informática / UFPE

____

Profa. Dra. Simone Cristiane dos Santos Lima
Centro de Informática / UFPE

____

Profa. Dra. Fernanda Araújo Baião Amorim
Departamento de Informática aplicada/UNIRIO

____

Prof. Dr. José Adson Oliveira Guedes da Cunha
Departamento de Ciências Exatas/ UFPB

____

Prof. Dr. Kai Fischbach
University of Bamberg

I dedicate this thesis to my parents, Manoel and Marlene,
and to my sister Andreia.

ACKNOWLEDGEMENTS

It has been a long journey to this point. Excitement and doubts turned into fear, hope, resistance, faith, collaboration, audacity, and personal and professional growth in the last five years. This was probably the period when I had the greatest professional challenges and personal growing experience so far. However, I cannot say that this is only my accomplishment. So many great people crossed my path and helped me in many ways to arrive here. First, I must thank the people who helped me way before this Ph.D. started. My humble, honest and lovely parents, Manoel and Marlene, for their incommensurable devotion in raising their children. They have taught me a lot just by being who they are individually and as a couple. My brother and sister, Adriano and Andreia, who decided to believe in my education. They wanted me to have the opportunities they could not get in the past. I feel really blessed by God in being part of my family.

My dearest teachers who influenced me from kindergarten to post-graduation. On behalf of all of them, I would like to thank my math teacher Isabel, who helped me coming from the Padre Tomaz de Aquino public school, in my little hometown, to make research cooperation with universities over the world. I am aware that I am an exception to the Brazilian public school system. On the other hand, this achievement motivates me to fight for improving it for other people. My Ph.D. advisor, Hermano Perrelli de Moura, for your trust, open mind, curiosity, for giving me science freedom and for the overall support along these years. I also extend my gratitude to the Professor Fischbach, for accept me as a visiting scholar at the University of Bamberg, Germany, and giving me great support to develop my research. During the time in your chair, I was able to meet key people who taught me a lot and allowed to develop this thesis and made me a better professional.

From the German chapter of this thesis, I want to thank some special people. Among them, Daniel Simon. Words are not enough to thank you for the research partnership, for hours of meetings, discussions, emails, and reviews. You played a major role in this thesis. Similarly, Mathias Kreimeyer, for the open doors. I cannot express enough gratitude in having worked with you. Your good will really inspired me for the rest of my life. Still in Germany, in the middle of uncertainties and culture clash, the light came to me as true friends. You were my support base that helped me to keep going. Alexandra Wolf and Mark Berninger, for the laughs, the Ph.D. related talks, the cultural exchange, our international abendessens, and the sensibility when I needed it. Mariam Dopslaf, for sharing your family, for the words which echo in my mind until now. In the darkest hours of winter(rs), I have found a great friend, Ryan Reinke, with such a great heart. Thanks for staying there in the worst and best moments, too. Long live to the memories and others to come. I must say that, without you four, I could not make it. Matthias, Alexander, Nadja, Ewelina, Diana, Lisa, Oliver, Tobias, Cassia, Heloisa, thanks for the time together, I hope to see you all soon. Danke sehr!

Back in Brazil, I was helped by two nice and smart undergraduate students, Amanda Oliveira and Aline Daniele. You do not know how much enjoyed to work with you in the last months, including some weekends. Thanks for saying “yes” to my not-so-normal projects, for the Spotify sessions, the snacks, coffees, and cultural exchange in the working background. You made the final months of this thesis more pleasant in Garanhuns. Glaucia, Chico and Eunice, it was important to have someone to talk about the Ph.D. experience. Thanks for your friendship and special moments shared.

Special thanks to the evaluation committee, professors Alexandre Vasconcelos, Simone Santos and Fernanda Baião for their valuable feedback since the qualifying exam; and to the professors Kai Fischbach and Adson Cunha, for the feedback in the final version of this document.

General thanks for the EA experts, the anonymous reviewers and people in conferences, who provided feedback for this work.

Finally, thanks to the Brazilian government for the financial support through the Ph.D. Sandwich Program. Thanks to professor Assuero Ximenes and my colleagues at the Federal Rural University of Pernambuco - Garanhuns, for the support and comprehension during these last two years.

I have no doubt in saying that alone (perhaps) we go faster, but together we go further. (unknown author).

Recife, February 2017.

Alixandre Santana

ABSTRACT

While the literature on Enterprise Architecture (EA) models, frameworks, and methodologies for EA implementation has many exemplars, the field is still missing mechanisms of EA analysis. EA analysis is the process which uses any technique or method to extract information from EA models about a particular concern, in order to support EA management by the experts or inform stakeholders. In this thesis, we model the EA as a complex network, a concept discussed in network science, to analyze EA structural aspects. During our exploratory study about EA network analysis (EANA), it was clear that the field was still lacking foundational aspects. First, no common language was shared by researchers. Secondly, there was no clarity about what concerns could be analyzed with network analysis initiatives and thirdly, the techniques and methods´ implementation were not clear in the papers. We solve those gaps in order to describe how to perform analysis of EA components and their relationships supported by network measures. The research approach comprehends qualitative methods such as systematic literature review, thematic analysis and design science research method. The research is conducted in three complementary and interrelated phases, aiming at first, to collect and synthesize the available knowledge about the analysis approaches existent in the literature. Next, we aim to trace a comprehensive understanding of the main concepts involved in EANA such as their analysis concerns, modeling decisions, inputs required and steps necessary to perform it. Altogether, this resulted in a set of six proposed artifacts: EANA meta-model, EANA library, EANA process, EANA data derivation strategy. Finally, we investigate the use of those artifacts, evaluating them empirically through their instantiations and/ or with the help of EA experts of three German multinational companies. The evaluation results were positive regarding, among other criteria, the efficacy and utility of the proposed artifacts in their respective contexts. As contributions, we claim the definition of the conceptual foundations of the EANA research field. Complementary, the study is not limited to the theoretical findings since it advances the understanding of empirical network analysis, whereas it offers a library of analysis initiatives, methods to derive EA data and guidelines to help experts through the analysis process (EANA process). Finally, we also add to the EANA knowledge base two new EANA methods which were also empirically evaluated. We expect that results can enhance the awareness researchers and practitioners about the EA network-based analysis´ efficacy and utility, a step necessary to develop more rationally grounded methods and tools to support the EA management considering structural aspects.

Keywords: Enterprise architecture. Network analysis. Structural analysis. Design science. Framework.

Resumo

Enquanto a literatura sobre modelos, frameworks e metodologias de implementação de arquitetura empresarial - AE (do inglês, Enterprise Architecture, EA) é representativa, a pesquisa em AE ainda carece de mecanismos específicos para sua análise. Análise de AE é o processo que usa técnicas ou métodos para extrair informações arquiteturais sobre um aspecto de interesse sobre a AE, a partir de modelos, e com o objetivo de dar suporte aos especialistas no gerenciamento da AE ou ainda pra informar seus stakeholders. Nesta tese, a AE é modelada como uma rede complexa, um conceito originário da teoria de redes, com o objetivo de analisar aspectos estruturais de AE. Durante o estudo exploratório sobre a análise de redes aplicada ao contexto de AE, constatamos a ausência de fundamentos conceituais básicos como, por exemplo, um nivelamento conceitual entre os autores dos trabalhos; desconhecimento sobre a abrangência dos estudos de análise estrutural no contexto de AE e finalmente, carência de informação acerca do processo de análise estrutural realizado nos trabalhos. Nosso objetivo principal na tese é investigar como as métricas e métodos de análise de redes podem sem aplicados no contexto de análise de AE. Métodos qualitativos de pesquisa como revisão sistemática de literatura, análise temática e design science research foram utilizados em três fases complementares e inter-relacionadas. Primeiramente, para coletar e consolidar o conhecimento sobre abordagens de análise de AE existentes na literatura. Numa segunda etapa, o objetivo foi traçar um entendimento abrangente sobre os principais conceitos envolvidos na análise estrutural de AE, mapeando seus métodos e técnicas utilizados, culminando com o design de quatro artefatos propostos: um meta-modelo para análise de redes no contexto de AE; uma biblioteca reunindo as iniciativas de análise extraídas dos artigos; um processo de análise de redes para AE e uma estratégia para derivação de dados. Na terceira e última etapa, investigou-se o uso dos artefatos avaliando-os empiricamente por meio de suas instanciações e da opinião de especialistas em AE de três organizações multinacionais alemãs. Os resultados foram positivos considerando, dentre outros critérios, a eficácia e utilidade dos artefatos propostos nos seus respectivos contextos. Como contribuições, esta pesquisa define os conceitos fundamentais para análise de redes em AE, além de avançar no entendimento acerca da análise empírica de redes naquele contexto, uma vez que apresenta um catálogo de métricas e métodos para derivação de dados, além de um processo para auxiliar os especialistas ao longo da execução da análise. Finalmente, a pesquisa também contribui para a base de conhecimento com dois métodos de análise validados também empiricamente. Com base nos resultados, espera-se corroborar o potencial da análise de AE baseada em redes, sua eficácia e utilidade para pesquisadores e práticos, além de estimular a adoção e desenvolvimento de ferramental para suportar o gerenciamento de AE, considerando seus aspectos estruturais.

Palavras-chave: Arquitetura empresarial. Análise de redes. Análise estrutural. Design science. Framework.

List of Figures

1.1 Structure of the thesis 34

2.1 Evolution of EAM according to Ahlemann et al. (2012). 36
2.2 Evolution of the number of EA studies through the years (SIMON; FISCHBACH, 2013). 37
2.3 Overview of the EA modeling representation (AHLEMANN, 2012) 39
2.4 EA components and relationships considered in the thesis. 40
2.5 Application layer meta-model of Archimate 41
2.6 EA model-based analysis (GAALOUL; GUERREIRO, 2015) 46
2.7 Lankhorst' EA analysis classification. Extracted from LANKHORST (2004) . . . 47
2.8 Consolidated view of EAM according to the findings of KOTUSEV et al. (2015) . 51
2.9 Undirected network generated from the Datatset02, detailed in chapter 3 53
2.10 The biggest node has the biggest centrality value of network, since it is connected with other four nodes 56
2.11 Dyad and triads: (a) - dyad (b) - triad with two relations (c) - triad with three relations 56
2.12 Clusters detected in the network are represented by different colors. 57
2.13 Two different networks with different values for density. 57
2.14 Uni-modal network 58
2.15 bi-modal network 58
2.16 Multi-modal network 58
2.17 Example of a graph. 59
2.18 Example of a heat map. 59
2.19 Example of a ranking list. 60
2.20 Visual representation of the affiliation network and data derivation process. 62
2.21 The Engineering Systems Multiple-Domain Matrix (ES-MDM) (Bartolomei,2007). 64

3.1 Overall SLR process adapted from BOELL; CECEZ-KECMANOVIC (2015) and KITCHENHAM (2007) 72
3.2 Relation between artifact and context (WIERINGA, 2010) 74
3.3 Information systems DSR framework (HEVNER; CHATTERJEE, 2010) 75
3.4 Design science research methodology (DSRM) process model (PEFFERS et al., 2007a) 76

4.1 Distribution of EA analysis research papers per year 98
4.2 Distribution of EA analysis research papers per country 98
4.3 EA analysis research papers per concern category 112
4.4 EA analysis research papers per model category 116

4.5 Levels of models' abstraction and EA contextualization among them (STEEN et al., 2004). 119
4.6 EA analysis research papers per category of approach 130
4.7 Number of models used by each EA analysis approach 136
4.8 Distribution of theoretical, simulated and empirical studies 138
4.9 EA analysis tool support considerations extracted from primary studies 140
4.10 Qualitative and quantitative approaches of EA analysis 141
4.11 Qualitative approaches by concern category 142
4.12 Quantitative approaches of EA analysis by concern category 142
4.13 Distribution of the data collection methods in EA analysis papers 144

5.1 EANA research focus and overlapping areas 149
5.2 Distribution of papers per year for the EANA SLR 154
5.3 Distribution of papers per country for the EANA SLR 154
5.4 Proposed EANA-MM 170
5.5 NAI library building process 175
5.6 The decision mechanism of the GQM-based EANA library 175
5.7 Use case diagram for the EANA library web portal 186
5.8 NAI search's interface 187
5.9 Interface showing the details of an NAI's 187

6.1 The concern and data oriented EANA approaches 193
6.2 The EANA conceptual model 198
6.3 A classification schema of methods for EANA at the component level 199
6.4 A decision schema to select a category of EANA at component level 201
6.5 Activities of the EANA with concern oriented approach 205
6.6 Phases of EANA data-oriented approach 207
6.7 Activities of the EANA with data oriented approach 208
6.8 Phases of EANA data-oriented approach 209

7.1 Graphical representation of a binary DSM 215
7.2 Representation of MDM's concepts – adapted from FURTMEIER; TOMMELEIN (2010) 216
7.3 General concept of data derivation. Adapted from LINDEMANN; MAURER; BRAUN (2008) 221
7.4 Visual representation for the derivation by affiliation. 222
7.5 Visual representation for the derivation by intra-domain relation. 223
7.6 Visual representation for the derivation by transitivity. 223
7.7 Visual representation for the derivation by attribute. 224
7.8 BoxBo derived network and highlighted betweenness outliers. 228

7.9 TecxTec derived network and highlighted eigenvector outliers. 229
7.10 VendorxVendor derived network and eigenvector outliers highlighted 231
7.11 VendorxVendor heat map 231
7.12 Heat map for BuxBu network 233

8.1 Combined EA analysis. Cycle I: EA models are analyzed with expert-based techniques; Cycle II: EA models are converted into EA network data; Cycle III: EA network analysis is performed using network measures; Cycle IV: Combined EA analysis 237
8.2 An application architecture modeled as a network (red nodes represent the ten highest eigenvector centrality values, yellow ones have the ten highest closeness, and green ones have the 10 top betweenness values). 238
8.3 Semantic adopted for modeling the Dataset01 and Dataset02 238
8.4 Cognitive-structural diagnosis analysis: defining the samples to compute recall and precision 241
8.5 Cognitive-structural diagnosis analysis method – pseudo algorithm 242
8.6 Attribute check analysis method – pseudo algorithm 242
8.7 Ishikawa diagram for true positive, false positive, and false negative BOs 246

9.1 Artifacts´ evaluations performed n this thesis 251
9.2 Experience in academy and industry of the respondents 255
9.3 Activities of the EANA with concern oriented approach 259
9.4 Activities of the EANA with data oriented approach 260
9.5 Distribution of the EANA activities of the EANA process from chapter 7 263

10.1 The EANA framework and internal relations among the proposed artifacts 272
10.2 The EANA framework´s instantiation guidelines 273

11.1 Thesis' contributions and the research flow 280

A.1 Elements of a simple sociogram 299
A.2 Network diameter illustration 300
A.3 Difference on network density 300
A.4 Node ERP is the most central system on the network 302
A.5 Bonanich's centrality demonstration 302
A.6 Differences between high local centrality (node B) and high betweenness centrality (node A) 305
A.7 Graph G and its four components 307
A.8 Partition of Graph in two structures according to their internal densities. 308
A.9 Cliques of 3, 4 and 5 nodes 309

A.10 Example of a 3k-core (Scott, 2010) 310

B.1 EANA-MM 312
B.2 GQM-based EANA library – a graphical representation for the “Coupling” NAI . . 313

G.1 A subset of the EANA containing the main fields of all 74 NAIs. 364

H.1 Interface for adding a new NAI 365
H.2 Interface for selecting stakeholder to be associated to a new NAI 366
H.3 Technical layers of the EANA library web portal 367
H.4 Class diagram for the EANA library web portal 368

List of Tables

2.1 Common EA Definitions . . . 38
2.2 EA layers and components terminologies adopted in our work . . . 39
2.2 EA layers and components terminologies adopted in our work . . . 40
2.3 Main EA analysis definitions found in the literature . . . 43
2.3 Main EA analysis definitions found in the literature . . . 44
2.4 EA representation and abstraction levels (adapted from LANKHORST (2004) . . . 45
2.5 Lankhorst' classification in four dimensions. Extracted from LANKHORST (2004) 47
2.6 EA Analysis types according Niemann - Adapted from NIEMANN (2006) . . . 48
2.7 Zachman framework (ZACHMAN, 1987). . . . 52
2.8 Network measures and their contextualization in EA. Based on SCOTT (1992); SIMON; FISCHBACH (2013) . . . 55
2.9 Matrix representation of a network . . . 60

3.1 Philosophical assumptions of the three research perspectives . . . 69
3.2 Distribution of activities of DSRM for each proposed artifact . . . 79
3.2 Distribution of activities of DSRM for each proposed artifact . . . 80
3.2 Distribution of activities of DSRM for each proposed artifact . . . 81
3.3 Main steps and checklist for thematic synthesis proposed by CRUZES; DYBA (2011) 82
3.3 Main steps and checklist for thematic synthesis proposed by CRUZES; DYBA (2011) 83
3.3 Main steps and checklist for thematic synthesis proposed by CRUZES; DYBA (2011) 84
3.4 Datasets description . . . 85
3.5 Classification schema for the evaluation methods of the artifacts . . . 87
3.6 Selected criteria for the artifacts' evaluation methods . . . 89
3.7 Research phases, research questions, methods and outputs . . . 91
3.7 Research phases, research questions, methods and outputs . . . 92

4.1 Research query . . . 94
4.2 Research query string adopted for the EA analysis SLR . . . 94
4.3 Studies returned by each engine. . . . 94
4.4 Inclusion Criteria . . . 95
4.5 Exclusion criteria. . . . 95
4.6 Output of each research phase . . . 96
4.7 Data fields extracted from the EANA primary studies . . . 96
4.7 Data fields extracted from the EANA primary studies . . . 97
4.8 EA analysis concerns´s categorization . . . 99
4.8 EA analysis concerns´s categorization . . . 100
4.8 EA analysis concerns´s categorization . . . 101

4.9 Detailing the EA analysis concerns 101
4.9 Detailing the EA analysis concerns 102
4.9 Detailing the EA analysis concerns 103
4.9 Detailing the EA analysis concerns 104
4.9 Detailing the EA analysis concerns 105
4.9 Detailing the EA analysis concerns 106
4.9 Detailing the EA analysis concerns 107
4.9 Detailing the EA analysis concerns 108
4.9 Detailing the EA analysis concerns 109
4.9 Detailing the EA analysis concerns 110
4.9 Detailing the EA analysis concerns 111
4.10 EA analysis models categories 112
4.10 EA analysis models categories 113
4.10 EA analysis models categories 114
4.11 EA Analysis model categorization (* These are recognized models in EA field and also well defined in theory. Thus we did not include their descriptions.) 114
4.11 EA Analysis model categorization (* These are recognized models in EA field and also well defined in theory. Thus we did not include their descriptions.) 115
4.11 EA Analysis model categorization (* Well defined in theory models are well recognized in the field of EA and are well defined in theory, thus we did not include their descriptions.) 116
4.12 Most used EA models according to our EA analysis literature review 117
4.12 Most used EA models according to our EA analysis literature review 118
4.13 Definition of the EA analysis approaches categories 119
4.13 Definition of the EA analysis approaches categories 120
4.13 Definition of the EA analysis approaches categories 121
4.13 Definition of the EA analysis approaches categories 122
4.14 EA Analysis approaches categorization 122
4.14 EA Analysis approaches categorization 123
4.14 EA Analysis approaches categorization 124
4.14 EA Analysis approaches categorization 125
4.14 EA Analysis approaches categorization 126
4.14 EA Analysis approaches categorization 127
4.14 EA Analysis approaches categorization 128
4.14 EA Analysis approaches categorization 129
4.15 EA Analysis approaches (categories) used for each concern 130
4.15 EA Analysis approaches (categories) used for each concern 131
4.15 EA Analysis approaches (categories) used for each concern 132
4.16 Number of distinct EA analysis approaches in relation to the total 133

4.17 Models (categories) used for each concern 134
4.17 Models (categories) used for each concern 135
4.18 Models (categories) used by each analysis approach category 136
4.18 Models (categories) used by each analysis approach category 137
4.18 Models (categories) used by each analysis approach category 138
4.19 Number of empirical, simulated and theoretical studies per concern 139
4.20 Strengthens and weakness of different data collection methods (YIN, 2003) contextualized in EA (adapted from NÄRMAN et al. (2011). 143

5.1 Research query string adopted for the EANA SLR 150
5.2 Inclusion criteria for papers of the EANA SLR. 151
5.3 Inclusion criteria for papers of the EANA SLR 151
5.4 Screening phases of EANA SLR 152
5.5 Data extracted from the selected primary studies for EANA SLR 152
5.5 Data extracted from the selected primary studies for EANA SLR 153
5.6 Categories of EANA concerns mapped from the EANA SLR 155
5.6 Categories of EANA concerns mapped from the EANA SLR 156
5.6 Categories of EANA concerns mapped from the EANA SLR 157
5.6 Categories of EANA concerns mapped from the EANA SLR 158
5.7 An excerpt of the NAIs identified during the EANA SLR 159
5.7 An excerpt of the NAIs identified during the EANA SLR 160
5.8 Data collection methods used by primary studies of the EANA SLR 162
5.9 Benefits pointed out by authors of primary studies of the EANA SLR 163
5.9 Benefits pointed out by authors of primary studies of the EANA SLR 164
5.9 Benefits pointed out by authors of primary studies of the EANA SLR 165
5.9 Benefits pointed out by authors of primary studies of the EANA SLR 166
5.10 DSR phases and contextualization for the EANA-MM (HEVNER; MARCH; RAM, 2008). 169
5.11 Summary of the EANA conceptual elements 172
5.11 Summary of the EANA conceptual elements 173
5.12 DSR phases and contextualization for the EANA library (HEVNER; MARCH; RAM, 2008). 174
5.13 The GQM-based EANA library 176
5.13 The GQM-based EANA library 177
5.13 The GQM-based EANA library 178
5.13 The GQM-based EANA library 179
5.13 The GQM-based EANA library 180
5.13 The GQM-based EANA library 181
5.13 The GQM-based EANA library 182

5.13 The GQM-based EANA library 183
5.13 The GQM-based EANA library 184
5.13 The GQM-based EANA library 185

6.1 DSR phases and contextualization (HEVNER; MARCH; RAM, 2008). 192
6.2 The concern and data oriented EANA approaches. Adapted from STEEN et al. (2004)194
6.3 Network analysis tools used in this thesis 202
6.3 Network analysis tools used in this thesis 203
6.3 Network analysis tools used in this thesis 204
6.4 Activities, inputs, outputs for EANA concert oriented approach 206
6.4 Activities, inputs, outputs for EANA concert oriented approach 207
6.5 Activities, inputs, outputs for EANA data-oriented approach 208
6.5 Activities, inputs, outputs for EANA data-oriented approach 209

7.1 Viewpoint's requirements defined by IEEE 1741 (IEEE, 2000) 214
7.2 DSR phases and contextualization (HEVNER; MARCH; RAM, 2008) 218
7.2 DSR phases and contextualization (HEVNER; MARCH; RAM, 2008) 219
7.3 EA layers and viewpoint matrix 220
7.4 Used network measures 222
7.5 Operators used to derive EA network data 225
7.6 The EADV-MM partial instantiation 226
7.7 Network analysis at the component level for BuxBu Prod (Phase I) 232
7.8 Network analysis at the component level for BuxBu Prod (Phase I) 234

8.1 Network measures used in this research and their contextualization to the case studies239
8.2 Outputs of the cognitive-structural analysis method 241
8.3 Recall and precision values for availability 244
8.4 Recall and precision values for business criticality, using outdegree and betweenness 244
8.5 Recall and precision values for "core application" using indegree, outdegree, eigenvector and betweenness 245
8.6 Percentage of TOP ranked Business Objects (BOs) and their subMS Completion Criteria supported by PLM 248

9.1 Methods, evaluation criteria for the proposed artifacts 251
9.1 Methods, evaluation criteria for the proposed artifacts 252
9.1 Methods, evaluation criteria for the proposed artifacts 253
9.2 Classification for the evaluation method of the EANA-MM 254
9.3 Classification for the evaluation method of the EANA library 254
9.4 Survey result for the EANA-MM 256
9.4 Survey result for the EANA-MM 257
9.5 Survey result for the EANA library 258

9.6 Classification for the evaluation method for the scenarios one, two and three 260
9.7 Classification for the evaluation method for the fourth scenario 261
9.8 Categorization for the EA network data strategy's evaluation method 265
9.9 Categorization for the EA network data strategy's evaluation method 267
9.10 Participation of the experts during the artifacts´evaluations 268

10.1 Survey result for the EANA-MM 271

11.1 Knowledge contribution classification (adapted from GREGOR; HEVNER (2013) . 281

B.1 Summary of the EANA conceptual elements 312
B.2 EANA meta-model instantiation 314
B.3 Respondent's background form 315
B.4 EANA meta-model - assessment 315
B.4 EANA meta-model - assessment 316
B.5 EANA library evaluation form 316
B.5 EANA library evaluation form 317

C.1 EA Analysis SLR primary studies 318
C.1 EA Analysis SLR primary studies 319
C.1 EA Analysis SLR primary studies 320
C.1 EA Analysis SLR primary studies 321
C.1 EA Analysis SLR primary studies 322
C.1 EA Analysis SLR primary studies 323
C.1 EA Analysis SLR primary studies 324
C.1 EA Analysis SLR primary studies 325
C.1 EA Analysis SLR primary studies 326
C.1 EA Analysis SLR primary studies 327
C.1 EA Analysis SLR primary studies 328
C.1 EA Analysis SLR primary studies 329
C.1 EA Analysis SLR primary studies 330

D.1 EA Analysis SLR Primary Studies classified by concern 331
D.1 EA Analysis SLR Primary Studies classified by concern 332
D.1 EA Analysis SLR Primary Studies classified by concern 333
D.1 EA Analysis SLR Primary Studies classified by concern 334
D.1 EA Analysis SLR Primary Studies classified by concern 335
D.1 EA Analysis SLR Primary Studies classified by concern 336
D.1 EA Analysis SLR Primary Studies classified by concern 337
D.1 EA Analysis SLR Primary Studies classified by concern 338
D.1 EA Analysis SLR Primary Studies classified by concern 339

D.1 EA Analysis SLR Primary Studies classified by concern 340
D.1 EA Analysis SLR Primary Studies classified by concern 341
D.1 EA Analysis SLR Primary Studies classified by concern 342
D.1 EA Analysis SLR Primary Studies classified by concern 343
D.1 EA Analysis SLR Primary Studies classified by concern 344
D.1 EA Analysis SLR Primary Studies classified by concern 345
D.1 EA Analysis SLR Primary Studies classified by concern 346
D.1 EA Analysis SLR Primary Studies classified by concern 347
D.1 EA Analysis SLR Primary Studies classified by concern 348

E.1 EA Analysis SLR primary studies classified by analysis approach 349
E.1 EA Analysis SLR primary studies classified by analysis approach 350
E.1 EA Analysis SLR primary studies classified by analysis approach 351
E.1 EA Analysis SLR primary studies classified by analysis approach 352
E.1 EA Analysis SLR primary studies classified by analysis approach 353
E.1 EA Analysis SLR primary studies classified by analysis approach 354
E.1 EA Analysis SLR primary studies classified by analysis approach 355

F.1 EA Analysis SLR primary studies classified by models 356
F.1 EA Analysis SLR primary studies classified by models 357

LIST OF ACRONYMS

AI	Attribute of Interest
AA	Application Architecture
BIT	Business-IT
BO	Business Object
BU	Business Unit
BP	Business Process
C4ISR	System Command, Control, Computers, Communications, Intelligence, Surveillance, and Reconnaissance
DSR	Design Science Research
DMM	Domain-Mapping Matrix
DoDAF	Department of Defense Architecture Framework
DREPT	Design-Relevant Explanatory/Predictive Theory
DSR	Design Science Research
DSM	Design Structure Matrix
EA	Enterprise Architecture
EANA	Enterprise Architecture Network Analysis
EANA-MM	Enterprise Architecture Network Analysis Meta-Model
EAIL	EA Intelligence Life Cycle
ES-MDM	Engineering Systems Multiple-Domain Matrix
EAM	EA Management
EADV-MM	Enterprise Architecture Derived Viewpoints Meta-Model

GQM	Goal Question Metric
IS	Information Systems
ISDT	Information Systems Directive Design Theory
MIT	Massachusetts Institute of Technology
MDM	Multiple Domain Matrix
NAI	Network Analysis Initiative
NMV	Network Measure Value
SoS	System of Systems Theory
SLR	Systematic Literature Review
TAFIM	Technical Architecture Framework for Information Management
TOGAF	The Open Group Architecture Framework
ZAF	Zachman Architecture Framework

CONTENTS

1 INTRODUCTION 27
1.1 CONTEXT AND MOTIVATION 27
1.2 RESEARCH PROBLEM STATEMENT....31
1.3 RESEARCH QUESTION 32
1.4 RESEARCH SCOPE AND CONTRIBUTIONS 32
1.5 OUTLINE OF THE THESIS 33
1.6 CHAPTER SUMMARY 33

2 THEORETICAL BACKGROUND 35
2.1 ENTERPRISE ARCHITECTURE 35
2.1.1 A brief historical perspective of EA and EA management 35
2.1.2 Towards a working definition of EA 37
2.1.3 EA models 40
2.1.4 EA view, viewpoint, analysis concern and stakeholder 42
2.1.5 EA analysis 43
2.1.5.1 Classification models for EA analysis found in the literature 46
2.1.6 EA management 49
2.1.7 EA frameworks 51
2.2 NETWORK SCIENCE AND STRUCTURAL ANALYSIS 53
2.2.1 Introduction 53
2.2.2 Network measures 54
2.2.3 Levels of network analysis 55
2.2.4 Unimodal, bi-modal and multi-modal networks 57
2.2.5 Outputs of network analysis 59
2.2.6 Representation of network models 60
2.2.7 EA network analysis (EANA) 60
2.2.8 Network data....61
2.2.8.1 Derived data and affiliation networks 62
2.3 RELATED WORK....63
2.4 CHAPTER SUMMARY 65

3 RESEARCH DESIGN....66
3.1 RESEARCH QUESTION 66
3.2 RESEARCH CHARACTERIZATION 67
3.2.1 Research outcomes 67
3.2.2 Research logic 67

3.2.3 Research goals or purpose 68
3.2.4 Research approach or philosophy 68
3.2.5 Research process 69
3.2.6 Research procedures or methodology 70
3.2.7 Data collection and analysis methods 70
3.3 RESEARCH METHODS 70
3.3.1 SLRs 70
3.3.1.1 Basics of the method 70
3.3.1.2 Operationalization of the method 71
3.3.1.3 Limitations of use 72
3.3.1.4 Applying the method 72
3.3.2 Design Science Research (DSR) 72
3.3.2.1 Basics of the method 73
3.3.2.2 Operationalization of the method 76
3.3.2.3 Limitations of use 78
3.3.2.4 Applying the method 78
3.3.3 Thematic analysis 81
3.3.3.1 Basics of the method 81
3.3.3.2 Operationalization of the method 82
3.3.3.3 Limitations of use 84
3.3.3.4 Applying the method 84
3.4 DATA COLLECTION 85
3.5 EVALUATION METHODS 86
3.5.1 *Classification of the evaluation methods applied to the proposed artifacts* 86
3.6 THE ROLE OF THE RESEARCHER 90
3.7 CHAPTER SUMMARY 90

4 EA ANALYSIS: WHAT DID WE ANALYZE SO FAR? 93
4.1 THE SLR RESEARCH PROTOCOL 93
4.2 RESULTS OF THE SLR 97
4.3 META-DATA 97
4.2.2 EA concerns mapped 99
4.2.3 EA models mapped 111
4.2.4 EA analysis approaches mapped 119
4.2.5 Crossing concerns, analysis approaches models 130
4.2.5.1 Concerns x Analysis approaches 130
4.2.5.2 Concerns x Models 133
4.2.5.3 Analysis approaches x models 136
4.2.6 Considerations about theoretical, simulated and empirical EA analysis 138

4.2.7 Tool support for EA analysis 140
4.2.8 *Qualitative/Quantitative aspects in EA analysis* 141
4.2.9 *Data sources and data collection methods in EA analysis* 143
4.3 CONCLUSIONS 144
4.3.1 An attempt of research agenda for EA analysis 145
4.3.2 Limitations 146
4.3.3 Future works for EA analysis 146
4.4 CHAPTER SUMMARY 146

5 TOWARDS THE EANA META-MODEL AND LIBRARY 148
5.1 The EANA SLR 148
5.1.1 The research protocol 150
5.1.2 Results of the SLR 154
5.1.2.1 SLR meta-data 154
5.1.2.2 Categories of EANA concerns mapped 155
5.1.2.3 NAIs mapped 158
5.1.2.4 Data collection methods 162
5.1.2.5 Benefits of the EANA 162
5.1.2.6 An attempt of research agenda for EA analysis 166
5.1.2.7 Conclusions.....167
5.2 THE EANA META-MODEL 168
5.2.1 Overview of the artifact's design 168
5.2.2 The EANA-MM artifact 169
5.3 THE GQM-BASED EANA LIBRARY 173
5.3.1 Overview of the artifact's design 174
5.3.2 The GQM-based EANA-library artifact 176
5.4 EANA LIBRARY WEB PLATFORM 185
5.4.1 Modules and user profiles 185
5.4.2 NAI search - use case 186
5.5 CHAPTER SUMMARY 188

6 TOWARDS A PROCESS VIEW FOR EANA 190
6.1 CONTEXT AND MOTIVATION 190
6.2 EANA PROCESS RESEARCH DESIGN 191
6.3 THE EANA PROCESS ARTIFACT 192
6.3.1 Analysis setup 192
6.3.2 Define an analysis approach 192
6.3.3. Define actors 194
6.3.4 Define stakeholders 194
6.3.5 Define data sources and data collection techniques 194

6.3.6 Build the model 196
6.3.7 Design NAI 197
6.3.8 Update EANA Library 201
6.3.9 Run NAI 201
6.3.10 Analyze and communicate results 204
6.4 METHODS FOR CONCERN AND DATA ORIENTED EANA APPROACHES 205
6.4.1 Concern-oriented analysis approach 205
6.4.2 Data-oriented analysis approach 207
6.5 CHAPTER SUMMARY 210

7 DERIVATION VIEWPOINTS FOR EANA 212
7.1 CONTEXT AND MOTIVATION 212
7.2 KEY CONCEPTS 213
7.2.1 EA viewpoints 213
7.2.2 DSM and related concepts 214
7.2.3 DSM and EA 216
7.2.4 Related work 217
7.3 ARTIFACT´S RESEARCH DESIGN 218
7.3.1 Research characterization 218
7.3.2 The EA data derivation mechanism 219
7.3.3 Data collection 221
7.3.4 Analysis methods 221
7.4 DERIVATION OPERATORS AND THE EADV-MM 221
7.4.1 Derivation operators 221
7.4.2 A meta-model for derived EA viewpoints 225
7.5 EMPIRICAL EVALUATION OF THE ARTIFACTS 227
7.5.1 Using Dataset01 227
7.5.2 Using Dataset02 230
7.5.3 Using Dataset03 231
7.6 CHAPTER SUMMARY 233

8 COGNITIVE-STRUCTURAL DIAGNOSIS ANALYSIS AND ATTRIBUTE CHECK ANALYSIS METHODS 236
8.1 CONTEXT AND MOTIVATION 236
8.2 ARTIFACTS' RESEARCH DESIGN 237
8.3 COGNITIVE-STRUCTURAL DIAGNOSIS ANALYSIS AND ATTRIBUTE CHECK ANALYSIS ARTIFACTS 239
8.3.1 Cognitive-structural diagnosis analysis 240
8.3.2 Attribute check analysis 242
8.4 EMPIRICAL EVALUATION OF THE ARTIFACTS 243
8.4.1 Cognitive-structural analysis 243

8.4.2 Attribute check analysis 247
8.5 CHAPTER SUMMARY 248

9 EVALUATING THE ARTIFACTS 250
9.1 ARTIFACTS AND EVALUATION CRITERIA 250
9.2 EANA META-MODEL AND EANA LIBRARY EVALUATIONS 253
9.2.1 Evaluation results 254
9.3 THE EANA PROCESS VIEW EVALUATION 259
9.3.1 Evaluation results 261
9.4 EA NETWORK DATA DERIVATION STRATEGY EVALUATION 265
9.4.1 Evaluation results 265
9.5 COGNITIVE-STRUCTURAL DIAGNOSIS ANALYSIS AND ATTRIBUTE CHECK ANALYSIS EVALUATIONS 266
9.5.1 Evaluation results 267
9.6 CHAPTER SUMMARY 267

10 MAKING SENSE OF THE EANA FRAMEWORK 270
10.1 AN INTEGRATED VIEW OF THE ARTIFACTS 270
10.2 USING THE FRAMEWORK FROM THE BEGINNING 272
10.3 EVALUATION OF THE OVERALL FRAMEWORK 275
10.4 CHAPTER SUMMARY 275

11 FINAL CONCLUSIONS AND FUTURE WORKS 276
11.1 SUMMARY OF FINDINGS 276
11.2 RESEARCH LIMITATIONS AND FUTURE WORK 281

REFERENCES 284
APPENDIX A – NETWORK ANALYSIS BASIC CONCEPTS 299
APPENDIX B – THE EANA META-MODEL AND LIBRARY'S EVALUATION SURVEY 311
APPENDIX C – EA ANALYSIS SLR - LIST OF PRIMARY STUDIES 318
APPENDIX D – EA ANALYSIS SLR - PRIMARY STUDIES GROUPED BY CONCERN 331
APPENDIX E – EA ANALYSIS SLR - PRIMARY STUDIES GROUPED BY ANALYSIS APPROACH 349
APPENDIX F – EA ANALYSIS SLR - PRIMARY STUDIES GROUPED BY MODELS 356
APPENDIX G – EANA NAIS 358
APPENDIX H – EANA LIBRARY WEB PORTAL 365

APPENDIX I – ACADEMIC PRODUCTION 369

1

INTRODUCTION

This chapter presents the overview of the current research. The research context and motivation, research problem, research questions and contributions are presented in Sections 1.1, 1.2, 1.3 and 1.4, respectively. The chapter closes with the thesis' outline described in section 1.5.

1.1 CONTEXT AND MOTIVATION

The development of new products and services, technology changes, government regulations, and economic crisis are examples of factors pushing enterprises to change. In this regard, enterprises should effectively adapt themselves to these changing processes and structures to stay ahead of the competition (AHLEMANN, 2012). However, "if changing initiatives are adopted independently, with little or no coordination across the enterprise, they can result in a plethora of heterogeneous, incompatible and costly changes to information technology, information systems, business process and organizational structures, especially in medium and large sized enterprises, sometimes running hundreds of information systems (AHLEMANN, 2012). In these cases, the IT function may hinder the organization's ability to respond to current and future market conditions in a timely and cost-effective manner (SESSIONS, 2007). At the extreme, IT complexity growth may lead to a situation where changes can no longer be efficiently implemented (SCHMIDT, 2013).

Part of these problems is due to a lack of internal understanding of structure, components and relations in different areas of the organization (HANSCHKE, 2014). Embracing all the major business and IT structures, as well as the associations that exist between them, Zachman (ZACHMAN, 1987) defined the concept of Enterprise Architecture (EA) to model a holistic view of the entire company. EA is a description of an enterprise from an integrated business and IT perspective intended to improve business and IT alignment (SCHMIDT; BUXMANN, 2011; BRADLEY et al., 2011; TAMM et al., 2011). For WAN; CARLSSON (2012), EA is a state of the art alternative for achieving enterprise management goals such as improving business performance, decreasing resource use, controlling risk and complexity, and coping with an uncertain environment. Having an accurate and comprehensive current view of EA components is an important reference for project planning, asset management, and investment decision-making (BERNARD, 2012). Unsurprisingly, EA is practiced by the majority of large companies (AMBLER, 2010; VAN DER RAADT; SLOT; VAN VLIET, 2007) and makes a significant contribution to their success (ROSS; WEILL; ROBERTSON, 2006; WIJEGUNARATNE; FERNANDEZ; EVANS-GREENWOOD, 2014; LANKHORST, 2016).

Lately, we have seen an increase in the research on EA (SCHNEIDER; ZEC; MATTHES, 2014). Previous works dealt with EA principles (STELZER, 2009), qualitative aspects like values and benefits (BOUCHARAS et al., 2010), modeling efforts like Archimate (The Open Group, 2013), and finally standards like The Open Group Architecture Framework (TOGAF) (The Open Group, 2013) and others. While the literature on models, frameworks, and methodologies for EA implementation has many exemplars, the field is still missing mechanisms of analysis to help experts to manage EA itself (LANKHORST, 2013).

EA analysis is the application of property assessment criteria on EA models (DAVOUDI; ALIEE; MOHSENZADEH, 2009). In this thesis, we elaborate on the previous definition to say that analysis is the process which uses any technique or method to extract information about a particular concern, in order to support EA management by the experts or inform stakeholders. EA Analysis approaches can have different purposes. For example, changes in the set of strategic goals (e.g. addition, deletion or modification of enterprise goals) triggers consistency checks on the EA models and revision of the current landscape (VASCONCELOS et al., 2004). On the other hand, analysis of the current situation (e.g. current IT environment of the enterprise) can trigger changes in enterprise design, potentially makes it outdated and creates conflicts with the new goals, new situations and potential imposed regulations (VASCONCELOS et al., 2004). A possible third analysis approach may focus in performing gap analysis between an AS-IS (current) state to a TO-BE (desired) state of the EA. To sum it up, all of these analysis methods and techniques should allow experts to extract relevant information from the enterprise models to provide structured information that might be valuable to manage the EA evolution (AIER; SCHÖNHERR, 2006) while ensuring business-IT (BIT) alignment as well.

Taking a brief look at literature of EA analysis, imminent paradigms are: probabilistic relation models (BUSCHLE et al., 2011) - which analyzes mainly EA availability and EA security aspects with Bayesian networks; EA business intelligence (VENEBERG et al., 2014), which together with other minor research streams, brings operational and quantitative data of EA components to the analysis (e.g servers' uptime, cost, number of transactions per minute and so on); complexity management (SCHNEIDER et al., 2015) - which considers the amount and heterogeneity of components and their relations (in this case, a typical measure employed is entropy); ontology-based analysis (ANTUNES et al., 2013) - which together with other formalisms, brings languages and rules to reasoning and support (semi) automated EA analysis; and network-based analysis (DREYFUS; WYNER, 2011) - which focus on modeling EA as a complex network, using network metrics and clustering algorithms from graph theory to analyze EA structural aspects. Nevertheless, after almost three decades of the Zachman's seminal paper, all the previous EA analysis paradigms have a long road ahead, as discussed in chapter 4, which presents a Systematic Literature Review (SLR) about EA analysis paradigms. We decided to shed light particularly to the last one: the network-based. Thus, we look at EA as a complex network, a concept discussed in network science (SCOTT, 1992). We chose to investigate further how EA modeled as a complex network can benefit from the application of network-based techniques, under a structural analysis perspective. We label each attempt to analyze EA modeled as a network analysis initiative (NAI), which can be a simple, well-known network measure (e.g. degree centrality, eigenvector), more sophisticated techniques or methods composed of those simple network measures (e.g. the hidden structure method, (BALDWIN; MACCORMACK; RUSNAK, 2013)); or the overall analysis inspired by network models (e.g. random networks). The reasons to that choice are enumerated in the following four points:

- EA as an intertwined network of components and relations

EA as an intertwined system of strategic goals, business processes, applications and infrastructure components is subject to a variety of relationships and dependencies among its several components. All these components and relations constitute the EA model. In this sense, EA is intrinsically suitable for the network (graph) modeling approach. Once the EA network of relations is identified, network analysis initiatives can be used to identify important elements in terms of structures that take relations into account (IYER; DREYFUS; GYLLSTROM, 2007; FÜRSTENAU, 2015).

- Cognitive limitation of human beings to analyze EA complex models

Given the limitation in organizational resources, it is important to identify, at least, a subset of components in the overall architecture that are supposed to be controlled or actively managed. To some extent, experts may be able to identify intrinsically important components that may, for example, support critical business tasks, represent large investments, or have a large number of users and are thus often the focus of decision makers (SIMON; FISCHBACH, 2013). In practice, these checks are, nowadays, often completed using rough and subjective estimations (SCHMIDT, 2013). This is where network measures may offer help.

In the view of SIMON; FISCHBACH (2013), these network analysis initiatives could represent more reliable and quantitative indicators to describe EA structure, helping experts guide its analysis process and evolution, allowing them to go beyond their perceptions, reducing subjectivity.

The capability of looking at the whole structure while identifying its critical points is specific of structural analysis paradigm. EA analysis based on ontology, EA business intelligence or probabilistic relational models do not offer this interesting requirement, since they do not focus on components' relationships neither can capture the complexity that emerges from them.

- Structural analysis has been applied in correlated fields

Structural analysis matters for software architecture field, but at another level of granularity (e.g. code, methods, classes, and modules) and for different analysis goals. The research about methods and techniques to analyze structural aspects such modularity and cognitive complexity of software, for example, has been developed with some extent (JENKINS; KIRK, 2007; PAN, 2011). In contrast to these perspectives, in EA, one would like to take the perspective of the IT manager looking at a company's set of applications and their interconnections, also referred to as Application Architecture (AA) (MOCKER, 2009). One does not look at their "interior" complexity. Instead, the analyst moves one level up by viewing the applications in relation to other applications (MOCKER, 2009), for instance.

Structural analysis was also applied to analyze components and relations in other EA related fields such as product engineering (LINDEMANN; MAURER; BRAUN, 2008), complex systems (CHRISTENSEN; ALBERT, 2007), system engineering (BARTOLOMEI, 2007), with success for many years. Applying the structural analysis is somehow consolidated at micro level (KREIMEYER, 2009), especially in product engineering resulting in a useful paradigm to understand those systems. In an analogous way, EA components, now, at a medium or macro organizational perspective level, analyzing different models and components, might benefit from that too.

Unfortunately, the same advances of structural analysis in those related fields cannot be observed yet in the EA community. We believe that the EA field can benefit from and evolve structural analysis as happened in those related fields. In some cases, we expect that metrics from them might be adapted and reused to contribute to the EA analysis context.

- Complexity behavior can also be measured

Complexity issues can be studied with network models such as path dependence (FÜRSTENAU, 2015), small world networks (WATTS; STROGATZ, 1998), and architectural control points (DREYFUS; IYER, 2008). This is possible only due to the explicit choice of modeling EA as a network. EA complexity phenomena was the most prominent analysis theme in a survey with respondents of 24 industries executed by VRIES; GERBER; MERWE (2015).

As in NARANJO; SÁNCHEZ; VILLALOBOS (2014), we could not find approaches that take full advantage of the several topological properties of enterprise models seen as networks/graphs, such as the differentiation of relations between elements, discovery of paths, clusters, or graph metrics. For SIMON; FISCHBACH (2013) and ANDERSEN; CARUGATI (2014), this is an emerging field with several research opportunities.

Considering the above discussion about EA network analysis potentialities, we decided to invest our efforts in organizing the existent body of knowledge, defining its current boundaries, mapping research opportunities and collaborating to transfer knowledge to the practice. In the next section, we present the barriers found in order to achieve these goals.

1.2 Research problem statement

During our exploratory study about EA network analysis, some punctual initiatives were found. Nevertheless, it was clear that the field still lacks elementary definitions and faces problems that are often present in emergent research fields, such as:

- No common language shared by researchers is available

A taxonomy for concepts like levels of network analysis, analysis concerns, and types of outputs produced by the analysis initiatives is missing. In addition, since some authors consider different modelling decisions, in terms of semantics for the same set of components, it's hard to compare results or advance previous proposed methods.

As a consequence, it's not clear in the papers what kind of network analysis approaches were taken by the authors and if they can be compared, for example.

- There is no clarity about what concerns can be analyzed with network analysis initiatives

Centrality methods' results can have different interpretations according to the semantics of the model. Studies can combine different network measures as well. So far, the field does not have a consolidation of what can be analyzed with the help of network analysis. All in all, though, the use of network analysis in EA has not yet been made systematic and has not been detailed in terms of which network centrality measures can help gaining insights into the IT landscape and single applications (SIMON; FISCHBACH, 2013).

Thus, we do not know the analysis possibilities nor what has been analyzed so far. To the best of our knowledge, the first literature review concerning this issue is provided by this thesis.

- Techniques and methods' implementation are not clear in the papers.

A problem that we have identified is that analysis techniques or methods often lack an explicit specification of the information and structures that they require to work. Thus, it is not easy to realize if a model is suited or not to support a particular analysis function. This is also discussed in RAMOS et al. (2014).

Thus, it is not clear what kind of data is necessary to perform the analysis (e.g. if manual or automated data collection techniques can be applied). Researchers do not provide details about the steps followed to apply their methods. In addition, several questions are underestimated like the semantic of relationships, types of data sources available, efforts to collect data and so on. The end-to-end analysis processes presented on the vast majority of papers are often nebulous, generating barriers to apply them in real cases.

In general, there is no guidance to transfer those methods to the practice due to this missing information. Furthermore, many of these methods are only theoretically discussed, still requiring cross-company validation (Santana A.; Fischbach K.; Moura H., 2016a). In a few words, network initiatives to perform structural analysis in EA were not consistently explored and can benefit from both practical and theoretical perspectives if these foundational conceptual issues could be clarified. This lack of theoretical ground was observed after our initial exploratory research and confirmed by a later broad systematic literature review (SLR) presented in Chapter 5.

Given the situation described, **there is a need for a foundational classification of the main constructs of EA network analysis, a mapping of which initiatives do already exist and what are its objects of interest, in terms of analysis**. In a second moment, the concern should be **to describe how this body of knowledge could be used by researchers and practitioners** to advance its knowledge base and practice, respectively.

1.3 Research question

In this sense, our main research question is *how to perform analysis of EA components and their relationships supported by network analysis initiatives?*

In order to answer this question, the following specific research sub questions are defined:

- RQ1 -Which analysis initiatives are already available?
    - RQ1.1 What are the concerns, models and approaches analyzed in EA?
    - RQ1.2 What are the concerns, models and approaches analyzed in EANA?
- RQ2 -What are the information requirements to perform EANA?
    - RQ2.1 What are the inputs to perform EANA?
    - RQ2.2 What kind of data can be used in EANA?
- RQ3 - How can EA network analysis initiatives be classified?
- RQ4 - How can experts be guided in performing EANA?

1.4 Research scope and contributions

- The development of a **framework for EA network analysis (EANA)**

We propose **an Enterprise Architecture Network Analysis Meta-Model (EANA-MM)** to classify existent EA network analysis research, the main modeling decisions for EANA, bringing a common set of concepts to be shared by the research community.

In addition, we list and organize all of the knowledge produced about network analysis in EA so far, contextualizing it with analysis concerns to design a **Goal Question Metric (GQM)-based EANA library**.

Transferring some ideas from social network science, we design **methods to derive data** from original EA components' data, generating in turn, new analysis possibilities.

Finally, we propose some guidelines to help experts through the analysis process (**EANA process**).

Altogether, the above contributions constitute what we call **EANA framework**.

- The development of two EA network analysis methods

We also design the **cognitive-structural analysis and attribute check analysis methods** which support architects in analyzing their EA concerns. Designing both methods was a means to evaluate the EANA framework. Both methods are introduced in chapter 8.

As an important side contribution, this thesis presents a **systematic literature review about EA analysis**, which organizes the research related to EA analysis paradigms in the last two decades. This review is described in chapter 4.

1.5 Outline of the thesis

The remaining of this document is organized as follows: chapter 2 brings the main concepts and theories that serve as basis for this research; chapter 3 details the whole research design, including research typology, objectives, and methods; chapter 4 – "EA analysis -what we did analyze so far?"- recovers the state of art of EA analysis approaches. Still in chapter 4, EA network analysis is put in perspective and compared to other paradigms with our EA analysis SLR; In chapter 5, the results of a second and specific SLR about EANA is presented, together with the EANA-MM and the GQM-based EANA library. In chapter 6, a process for EA network analysis is introduced. A strategy for EA network data generation is described in chapter 7. In chapter 8, we present the cognitive-structural analysis and attribute check analysis methods. The evaluation of the six proposed artifacts is described in chapter 9; Chapter 10 presents the EANA framework. Finally, we make the conclusions of the research in chapter 11. The thesis' structure is depicted with more details in Figure 1.1.

1.6 Chapter summary

In this first chapter, an overview of the thesis was presented. We introduced the problem contextualization, research question, and the expected contributions to theory and practice. In chapter 2, the key theoretical concepts of this thesis are discussed.

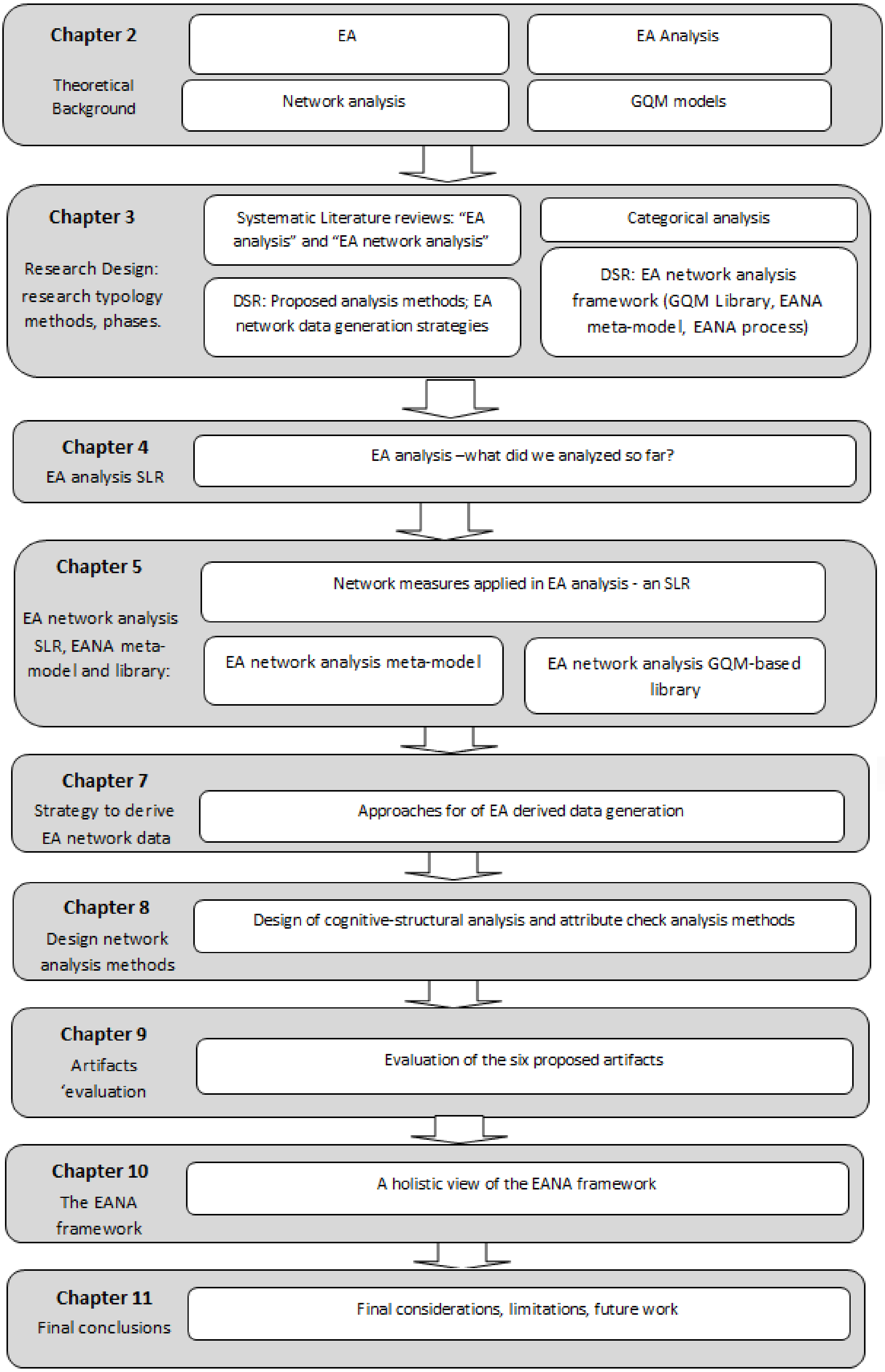

Figure 1.1: Structure of the thesis

2

THEORETICAL BACKGROUND

Once the overview of the thesis was presented, in this chapter we describe the key theoretical concepts of EA and network analysis employed in this thesis. In Section 2.1, we start with a brief review of EA history and then, we introduce the EA concept adopted in this thesis. In the sequence, we discuss EA modeling, EA analysis and EA management, three fundamental constructs in this thesis. Closing the EA discussion, we briefly review the main EA frameworks. At this point, in Section 2.2, we switch the focus to understand the network science, its main concepts, metrics and methods; and to discuss the implication of modeling EA as a complex network. Finally, we conclude the chapter with the related works in Section 2.3.

2.1 ENTERPRISE ARCHITECTURE

2.1.1 A brief historical perspective of EA and EA management

Architecture is well known from the world of construction. After years of architecture in the physical world, the term has also taken a foothold in the field of IT (CLOO et al., 2009). Generally, it is considered that the EA was first introduced by Zachman in 1987, in his paper "A Framework for Information Systems Architecture"(see Section 2.1.7). Zachman claimed that to manage a company's information systems, they needed to be specified in the same way that e.g. an airplane or a building is (PONTUS et al., 2013). Thus, he introduced the conceptualization of architectures from multiple perspectives (e.g., objectives/scope, enterprise model, system model and technical model), using different architectural descriptions (e.g., data, function and network). The framework is described as a matrix (with 30 cells) and suggests specification documents for each cell (e.g., using entity-relationship models to describe data, or using functional flow diagrams to describe processes). Despite not providing any concrete directions to construct an EA, it has influenced the subsequent frameworks over time.

Since the 1980s, EA has evolved together with the evolution of the business practice, when EA management takes place. AHLEMANN (2012) defines EA Management (EAM) as "a management practice that establishes, maintains and uses a coherent set of guidelines, architecture principles and governance regimes that provide direction for and practical help with the design and the development of an enterprise's architecture to achieve its vision and strategy". The evolution of the EAM concept through the last three decades, as it evolves from the information systems scope to a strategic practice, is depicted in Figure 2.1.

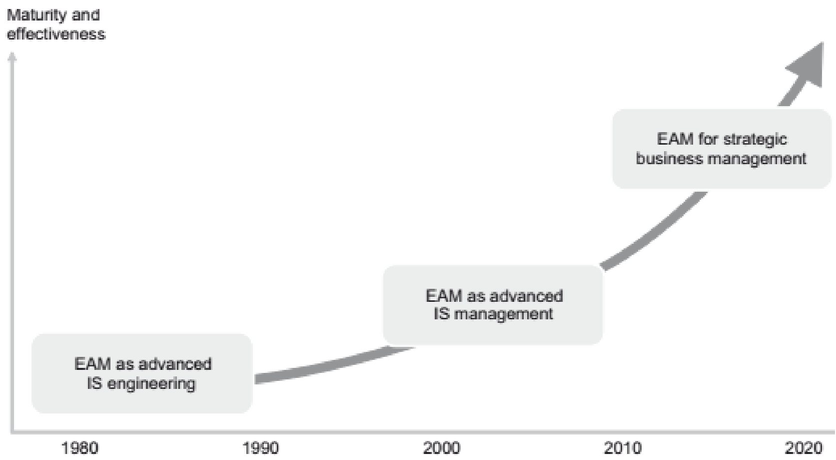

Figure 2.1: Evolution of EAM according to Ahlemann et al. (2012).

According to AHLEMANN (2012), EAM's formation phase was at the beginning of the 1980s, with IBM's 'business systems planning' concept and the subsequent development of the Zachman framework. At this time, Zachman observed that the term 'architecture' was widely used by information system professionals, but often had different meanings. Zachman's framework provided the means for a great leap forward. Zachman's idea of a multi-perspective and multi-layered enterprise modeling approach became state of the art at the beginning of the 1990s, influencing many other frameworks.

During the 1990s and 2000s (second phase), technological advances and the dissemination of desktop computing, local area networks and increased business process digitization, IT/IS landscapes became increasingly complex (AHLEMANN, 2012). This also meant that more stakeholders were involved and IT/IS spending increased. EAM professionals felt that a pure modeling approach, such as Zachman´s framework, was not enough. Then, EAM was taken to the next level incorporating information system management practices to take control of the IT/IS landscapes. Advanced EAM frameworks emerged and not only provided architectural artifacts and models but also contained guidelines for EAM planning, implementation and controlling (AHLEMANN, 2012).

The Department of Defense was one of the first federal agencies to adopt EA (BUSS; SHILLABEER, 2012). In order to speed up the delivery of information systems, lower their costs, and promote integration and flexibility, the Defense Information Systems Agency (DISA) in 1994 introduced the Technical Architecture Framework for Information Management (TAFIM). After the passage of the Clinger-Cohen Act in 1996, TAFIM was superseded by the System Command, Control, Computers, Communications, Intelligence, Surveillance, and Reconnaissance (C4ISR) framework (GROUP et al., 1997; LEVIS; WAGENHALS, 2000) and officially withdrawn in 2000. After TAFIM had been replaced, its materials were explicitly given to The Open Group and provided a basis for the creation of the TOGAF® standard initiated in 1995. Presently, TOGAF (2011) is the most cited and widely discussed publication in EA literature (SIMON; FISCHBACH, 2013). It embodies the modern understanding of EA and is even considered as a de facto industry standard in EA practice by some authors (KOTUSEV, 2016).

In the third and most recent phase, EAM is no longer understood as just an IT department job, but as a strategic function (AHLEMANN, 2012). Consequently, EAM aligns with the organization's strategy planning and strategy implementation processes. Many leading organizations already follow this broader understanding of EAM and involve highly skilled EAM specialists in these processes.

In terms of research evolution, SIMON; FISCHBACH (2013) present how studies about EA grew significantly after the 2000s (see Figure 2.2). Since then, several frameworks emerged, focusing on different perspectives of EA, and most of them present their own modeling approach. This plurality of frameworks and models resulted in the vast but heterogeneous field.

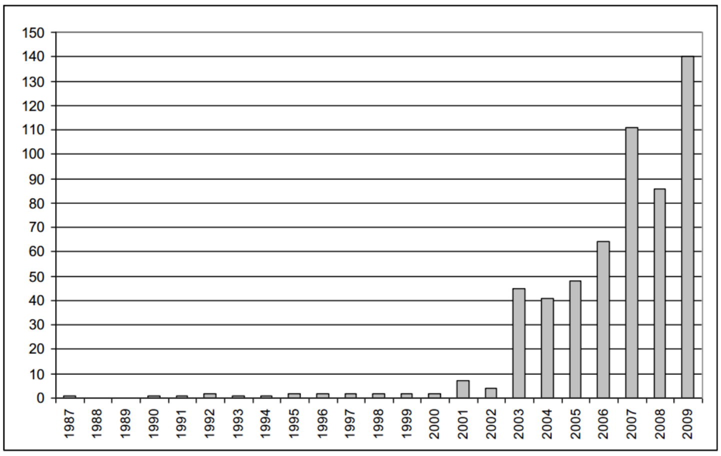

Figure 2.2: Evolution of the number of EA studies through the years (SIMON; FISCHBACH, 2013).

2.1.2 Towards a working definition of EA

To date, EA has not a standard definition, as authors have consistently enhancing the concept, highlighting different aspects and adapting the concept to the reality of various organizations. According to SCHOENHERR (2009), until 2008, 49 authors presented their own definition of EA. Despite this plurality in the literature, KOTUSEV et al. (2015) state that most of the definitions are incomplete or cannot be universally accepted. The reason is their focus on practices or artifacts used in EA or to describe it, and they can vary according to the chosen EAM approach; an EA definition also must explicitly state EA is about business and IT alignment. As LANKHORST (2004) declares, "the most important characteristic of an EA is that it provides a holistic view of the enterprise" and thus it should be clear in any description. Table 2.1 exemplifies some definitions found and the focus of each one of them.

Though the statements are not incorrect, they may not suit the reality of a different EAM approach. In this work, we adopt the general definition given by KOTUSEV et al. (2015):

Authors	EA definition	Focus
ADDICKS; APPELRATH (2010)	"An EA is a triple EA = (A,R, L) in which A is the set of all artifacts, R is the set of all relations that connect exactly two artifacts, and L is a set of layers (cf. [6]) that are used to partition the whole structure."	Alignment among artifacts by their relations; layer organization.
CAETANO; SILVA; TRIBOLET (2009)	"An enterprise architecture is the result of the continuous process of representing, integrating and keeping consistently aligned the elements that are required for managing and understanding the organization"	Representation and integration and management of elements.
LANKHORST (2004)	"A coherent whole of principles, methods, and models that are used in the design and realisation of an enterprise's organizational structure, business processes, information systems, and infrastructure.	Design and realisation of an enterprise's organizational structure, business processes, information systems, and infrastructure"
AHLEMANN (2012)	"Enterprise architecture (EA) is therefore understood as the fundamental organization of an enterprise as a socio-technical system, along with the principles governing its design and development. An EA includes all relevant components for describing an enterprise, including its business and operating model, organizational structure, business processes, data, applications and technology."	Socio-technical aspect; holistic view; layers.

Table 2.1: Common EA Definitions

"EA is a description of an enterprise from an integrated business and IT perspective". Therefore, we understand EA essentially as **a model** to describe the organization from the very up-level (where strategy is discussed) to the infrastructure level of an organization, mapping its components and relationships.

Organizations use EA for different reasons. SCHOENHERR (2009) lists the main drivers for the EA adoption, separating them in internal or external drivers. Internal drivers are related to benefits to the organizations; while external ones might be a response to an external regulation (compliance law). The main internal drivers noticed are business-IT alignment, cost reduction, standardization, and governance. All drivers are connected and are a reflection of how EA allows architects and managers to make smarter decisions about IT Investments, and thus permits firms respond to changes leaving in consideration the impact in the enterprise as a whole.

In order to describe the organization structure, EA models often comprise a huge number of **components**. The EA is inclusive if it is presented from different perspectives at different **layers** of abstraction (AHLEMANN, 2012). However, still now, there is no consensus about the definitive set of layers to be represented. For example, in Figure 2.3 is depicted the modeling proposal of LANKHORST (2004) for EA containing four layers, reflecting the real world organization and modeling approaches suggested to describe each layer.

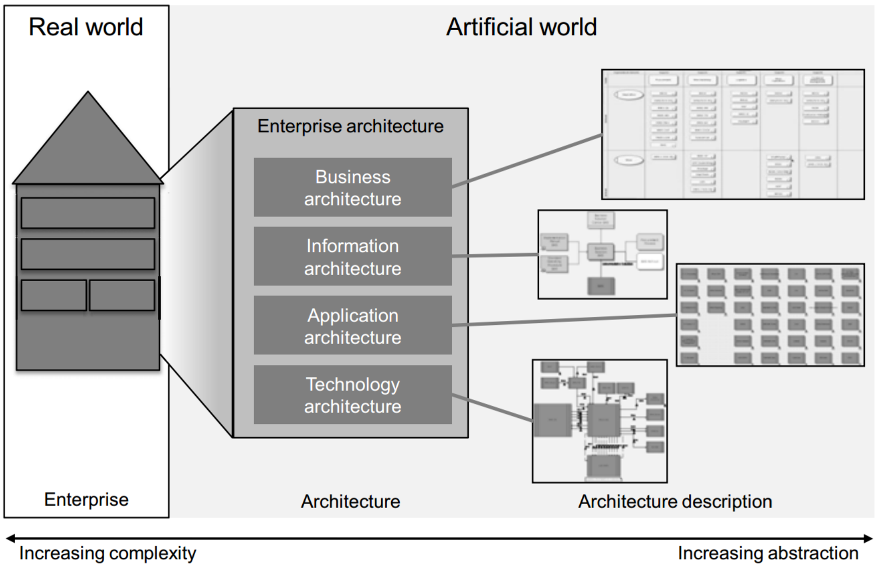

Figure 2.3: Overview of the EA modeling representation (AHLEMANN, 2012)

Thus, the number of layers varies from author to author, from 3 to 5 layers, with different (although similar) nomenclatures. The Open Group (2011), for example, presents an architecture modeled in four layers: business, application, data and technology. Lately, some extensions have been proposed as well, like motivation and strategic planning for Archimate (GROUP, 2016). For this thesis, we consider the layers described in Table 2.2 as integrating our architecture model together with their respective components. They reflect the components analyzed in the primary papers studied in the two SLRs perfomed in this thesis.

Layer	Decription	Typical EA components
Value	Business strategy elements are mapped in this layer.	Mission, enterprise goals, directives, principles, Stakeholders.
Business	A description of the structure and interaction between the organization units, business functions, business processes (The Open Group, 2011).	Business process, Business units.
Information	Describes how the enterprise data is organized and accessed (The Open Group, 2011). Business entities are described in this layer.	Business Objects.

Continued on next page

Table 2.2: EA layers and components terminologies adopted in our work

Table 2.2 – continued from previous page

Layer	Description	Typical EA components
Application	It is a high-level logical representation of a collection of software applications, usually implementing specific business functionality.	Application, Service.
Technology	A description of the structure and interaction of the platform services, logical and physical technology components. (The Open Group, 2011).	Application server, Database system, platforms

Table 2.2: EA layers and components terminologies adopted in our work

Those relations among the EA components described in Table 2.2 are now depicted in Figure 2.4:

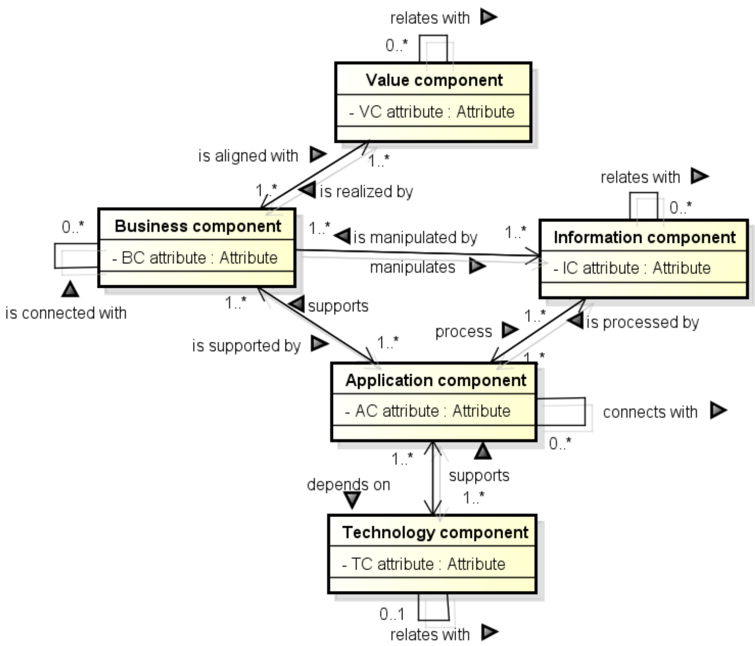

Figure 2.4: EA components and relationships considered in the thesis.

The task of building EA models is known as enterprise modeling and is discussed in the next section.

2.1.3 EA models

According to STECHER (1993), enterprise modeling is the activity of defining how an enterprise operates. EA modeling consists of the development of Enterprise Models (EMs), which are the embodiment of all the collected information about the enterprise from several perspectives (NARANJO; SÁNCHEZ; VILLALOBOS, 2014). It is done to understand the enterprise activities better before any changes are undertaken.

In most complex environments such as automotive or aircraft manufacturing, the need for modeling is commonly accepted as a means of simplification and concealment (STECHER, 1993). For example, the development of a new airplane begins first with a model -often basic- of the airplane before a complete set of engineering drawings is produced. Similarly, it is possible to apply modeling techniques to an enterprise (STECHER, 1993). At the higher levels, it helps an executive identify the functions required to run the business. Models at this level cover the entire enterprise, not just the functions that are computerized. The enterprise is defined in the business language, not in IS jargon. At this level, it is possible to consider different strategies for running the business and to assess the effect on the business processes.

One example of modeling language commonly adopted in the EA field is the Archimate (GROUP, 2016). Archimate defines a particular set of layers, components and relationships to describe an EA. For example, in Figure 2.5, we depict a fragment of its meta-model (the application layer meta-model). The application layer is typically used to model the information systems architectures of the enterprise (The Open Group, 2013). Other different modeling approaches can also be used. For instance, in chapter 4, based on a SLR made with a set of 120 selected primary studies, we describe other modeling choices identified in our EA analysis SLR such as formal specification-based (present in 10 of 120 studies), graph-based (20/120), intentional modeling languages(3/120), probabilistic networks- based (24/120) and so on. Archimate-based models, which are the standard models of TOGAF, were present in 25 of 120 papers.

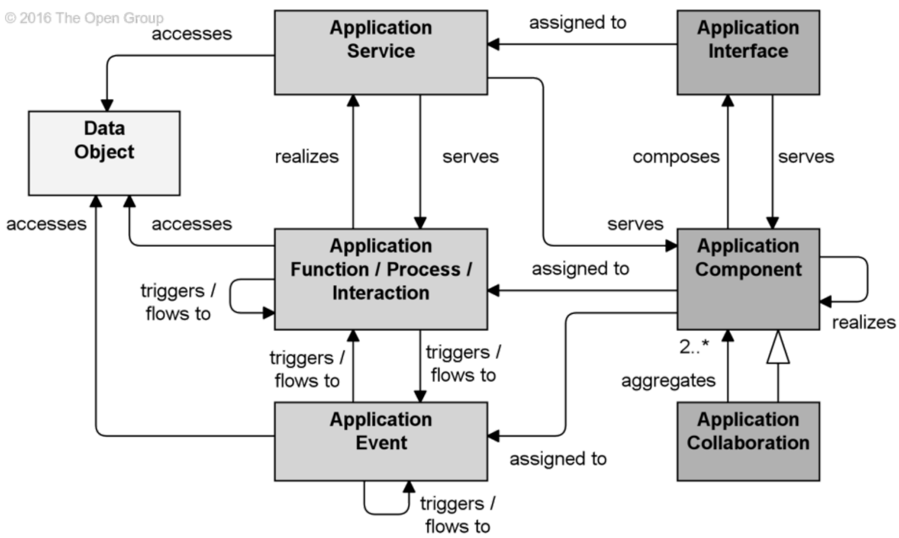

Figure 2.5: Application layer meta-model of Archimate

EA models are used to abstract the structure of the enterprise in its current state (as-is models), showing possible alignment issues, easing communication and to aid in decision-making, by being used to predict the behavior of future states (to-be models) rather than modifying directly the systems in the current architecture (BUSCHLE et al., 2011). PONTUS et al. (2013) emphasize two important aspects of using models: objectivity and communication. Models permit focus on the important aspects of a problem, enabling an analysis of a specific concern or perspective of a situation.

Furthermore, models provide a common language that permits a better communication among people from different backgrounds and focuses, improving the possibility of gaining consensus of understanding of the as-is status, to-be architecture design and the many options to implement EA transition. EA models are, thus, "effective tools for planning, communicating, and of course, also for documenting remembering”(PONTUS et al., 2013). Yet, EA modeling provides a foundation for enterprise analysis and implementation. It enables enterprise architects to do a gap analysis, capability analysis, make roadmaps for enterprise transition and finally implement changes based on the roadmaps (DAM; LÊ; GHOSE, 2015).

EA modeling combined with EA analysis are two of the subjects where EA can deliver real value to the organization. Building EMs typically comes with a high price tag that is paid for when they are analyzed, i.e. when additional knowledge is created by processing and reworking previously defined facts (Buckl et al., 2010). This topic is discussed in the next section.

2.1.4 EA view, viewpoint, analysis concern and stakeholder

The complexity of the execution of an enterprise’s strategy is likely to be immense because many processes, departments, and information systems are involved. When using enterprise architecture as a planning and steering instrument, then this instrument should reflect this complexity. As a result, it is almost undoable to make one single univocal and comprehensive set of models that can be used for all people concerned (CLOO et al., 2009). Besides, very often, no stakeholder, apart from perhaps the architect, is interested in the architecture in its full scope and detail. Therefore, the notions of **view and viewpoint** become central to the communication of architectures.

The IEEE 1471 standard (HILLIARD, 2000) defines views as a representation of a system from the perspective of a related set of concerns. A viewpoint is a specification of the conventions for constructing and using a view; a pattern or template from which to develop individual views by establishing the purposes and audience for a view and the techniques for its creation and analysis. In other words, TOGAF (The Open Group, 2011) describe an architecture view as a representation of an overall architecture with meaning to one or more stakeholders in the system. A viewpoint defines the perspective from which a view is taken. The latter defines: 1-How to construct and use a view; 2- the information needed; 3- the modeling techniques for expressing and analyzing it; 4- a rationale for these choices (e.g., by describing the purpose and intended stakeholders of the view).

A stakeholder is an individual, team, or organization (or classes thereof) with interest in, or concerns relative to, a system (such as an enterprise). **Concerns** are those interests, which pertain to the system's development, its operation or any other aspect that is critical or otherwise important to one or more stakeholders (CLOO et al., 2009).

2.1.5 EA analysis

Through our research, we found that just a small portion of the papers took the time to define the concept of EA Analysis, reflecting a lack of theoretical foundation in the topic. Many works reason about the importance of Analysis in EA, its benefits, the advantages to using models as a base for evaluation techniques, however, they do not present a clear definition to what they consider EA Analysis, as exemplified in Table 2.3:

Author	Definition	Aspect presented
JOHNSON; NORDSTRÖM; LAGERSTRÖM (2007)	EA analysis is the application of property assessment criteria on EA models	Analysis based on models.
BUCKL et al. (2008)	EA analysis is a means of providing decision support throughout the management process by making the impact of planned projects explicit.	Analysis influence in decision-making/change management.
BUSCHLE et al. (2011)	EA analysis is at the core of making rational decisions about information systems.	Analysis as the core of decision- making process about IS.
FASANGHARI et al. (2015)	"EA analysis is an essential tool in achieving organizational efficiency and effectiveness . Although there are many EA analysis methods in the literature of the context, there is a lack of an adequate group decision-making model for analyzing the proposed EA scenarios in conditions of uncertainty ."	EA analysis as a tool

Continued on next page

Table 2.3: Main EA analysis definitions found in the literature

Table 2.3 – continued from previous page

ID	Category	Definition
BUCKL et al. (2009)	"Using EA analysis, system properties (quality attributes) such as availability, performance, flexibility, or operational risk, etc., can be operationalized as metrics. These metrics can then be used to compare different scenarios (Lankes, 2008)."	Benefits of EA analysis
RAZAVI; ALIEE; TAFRESH (2009)	"Because the risk and impact of EA are pervasive across the enterprise, it is critical to perform an architecture assessment before any decision about choosing a scenario. EA analysis is the application of property assessment criteria on EA models (Johnson 2007c)."	EA analysis importance

Table 2.3: Main EA analysis definitions found in the literature

In this thesis, grounded in our SLR performed in chapter 4, we elaborate on the previous definition to say that EA analysis as the assessment of any EA´s property, based on models or other EA related data, in other to inform or bring rationality to decision support of stakeholders. Compared to the previous definitions in Table 2.3, ours adds the possibility of using also raw EA data not organized in models (e.g. cost of IT applications) to the analysis. Besides, our definition explicitly brings the goals for the analysis process: just inform EA stakeholders or add rationality to their decision process. Finally, any EA components such as information systems, business process or technologies can be considered in the EA analysis.

EA changes may create a ripple effect among other elements of architecture, so every decision must be well evaluated before actual implementation. This is where analysis is helpful: it allows the assessment of the scenarios before taking any serious risks. Almost all EA analysis definitions presented share the focus on decision making to future states. Evaluate the current state of the architecture is also important to detect bottlenecks in the current state, correct misalignments and improve the organization's performance. Therefore, an analysis concern is an aspect of the EA component or of the whole EA, which is investigated/measured by the experts in order to help EA evolution, to communicate with stakeholders and/or to ensure business-IT alignment.

EA analysis can have different audiences. In his book, LANKHORST (2013) presents a classification for EA levels of representation abstraction and respective, audiences and purposes of analysis which are described in Table 2.4:

Level of abstraction	Typical stakeholders	Purpose	Examples
Details	Software engineer, process owner	Design, manage	UML class diagram, Testbed process diagram
Coherence	Operational managers	Analyze dependencies, impact of change	Views expressing relations like "use", "realize" and "assign."
Overview	Enterprise architect, CIO, CEO	Change management	Landscape map (GROUP, 2016)

Table 2.4: EA representation and abstraction levels (adapted from LANKHORST (2004)

According to our EA analysis SLR (see chapter 4), the EA analysis may have a **qualitative or quantitative** character. Qualitative analysis usually relates to strategic concerns, evaluating alignment, goal and requirements compliance or decisions (upper layers of EA). Quantitative approaches provide numeric values, usually from quantitative metrics, and help to evaluate system properties (quality attributes) such as availability, performance, flexibility, operational risk, etc. (BUCKL et al., 2009).

Following our EA analysis classification, some approaches perform analysis based on other data sources than EA models (**model-based analysis**), e.g. stakeholders' interviews. We classify those approaches as EA **functional analysis**. For example, RICO (2006) presents metrics to evaluate the return of investment of an EA using cost related data from the enterprise. Simirlaly, GAMMELGÅRD; EKSTEDT; NÄRMAN (2007) use stakeholders' interviews in the decision-making about TO-BE scenarios.

The vast majority of EA analysis approaches are based on models (**EA model-based analysis**). Models are useful to analyze the as-is state of an organization and to model proposed scenarios in an EA (to-be states), for projects and change initiatives, for instance, as depicted in Figure 2.6.

NARANJO; SANCHEZ; VILLALOBOS (2012) classify three types of models used to EA analysis: queries, views, and visualizations. Queries define questions expressed in a formal language (e.g. ontologies). Views express a perspective of the models, as they cover only a specific stakeholder concern, i.e. they do not represent the "big picture". Using different views in analysis allows investigating a combined influence of factors from different views on the organizational behavior, and provides a more informed decision-making process than approaches that take only one view in consideration (POPOVA; SHARPANSKYKH, 2007). The last type is visualizations that use a visual language, "visual metaphors", to deliver key information, e.g. diagrams or images. Though they have a significant cognitive importance, they lose comprehensibility and communication as complexity increases. We present our own wider

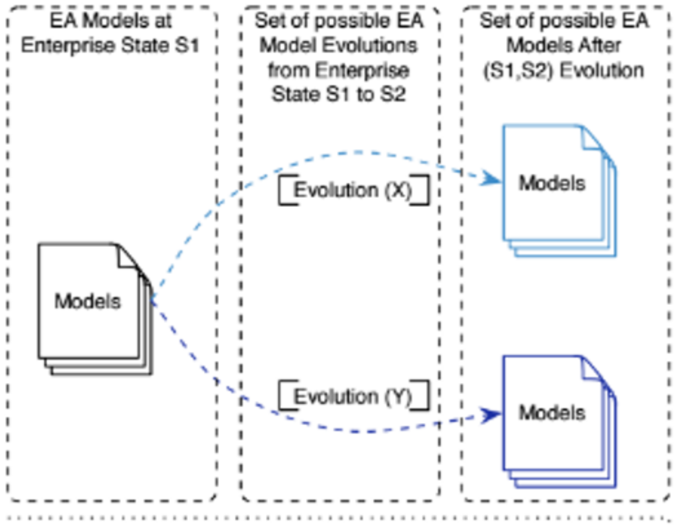

Figure 2.6: EA model-based analysis (GAALOUL; GUERREIRO, 2015)

categorization of models used in EA analysis in Section 4.2.4.

Finally, some approaches evaluate the impact and benefits of an EA in the organization, analyzing EA as whole and abstract phenomena. For instance, these techniques may cover the overall organization performance after the adoption of EA. We classify these approaches as **EA as a phenomenon analysis**. They are not covered in the scope of our research.

2.1.5.1 Classification models for EA analysis found in the literature

With the evolution of research regarding EA analysis, some classification schemas emerged, such as the classifications of BUCKL et al. (2009); LANKHORST (2004), and the concern-based classification model presented by NIEMANN (2006). All of them are detailed in the following.

- Lankhorst et al.(2004)

LANKHORST (2004) shows the variety present in techniques and methods, analyzing them according to the type of the employed technique (analytical x simulation) and type of produced result (quantitative x functional). Functional analysis is performed to gain insight into the functional aspects of architecture. Quantitative analysis techniques answer quantitative questions like "how quick" or "how cheap" a process is. For both functional and quantitative analysis, LANKHORST (2004) distinguish two main types of techniques: analytical techniques and simulation.

Figure 2.7 depicts the four dimensions of analysis described by LANKHORST (2004). These categories can be enhanced by other detailed analysis aspects. For instance, it does not details what types of functional and quantitative analysis concerns can be targeted, as we do in chapter 4.

- Buckl et al. (2009)

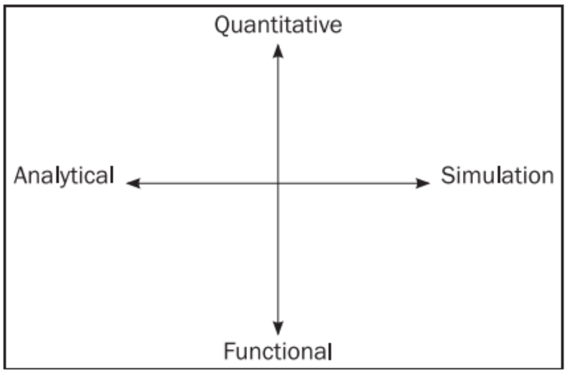

Figure 2.7: Lankhorst' EA analysis classification. Extracted from LANKHORST (2004)

BUCKL et al. (2009) present the following EA analysis dimensions: body of analysis, time reference, analysis technique, analysis concern and self-referentiality, as depicted in Table 2.5.

Dimension	Possibilities		
Body of analysis	structure	behaviour statistics	dynamic behaviour
Time reference	ex-post	ex-ante	
Analysis technique	expert-based	rule-based	indicator-based
Analysis concern	funcional	non-functional	
Self-referentiality	none	single-level	multi-level

Table 2.5: Lankhorst' classification in four dimensions. Extracted from LANKHORST (2004)

The "body of analysis" dimension can be classified into three types: 1- structure, the focus of the analysis technique is on connections and elements, i.e. the structure of the EA; 2 - behaviour statistics, covers analysis methods that provide statistic information about the system's behavior; and 3 - dynamic behaviour, techniques that inform about the behavior of system's constituents through time, e.g. the impact of a system failure propagating over time.

The second dimension, "time reference", divides methods regarding the time position in which the analysis takes place. The work considers only model-based EA analysis and if the models are used to analyze the existing architecture (ex-post) or analyze planned architectures (ex-ante).

The "analysis technique" dimension covers the level of formalization of EA analysis techniques. Expert-based analysis approaches are the most flexible but also most time-consuming ones and depend on the experience and expertise of the executing person (BUCKL et al., 2009). Thus their results can be considered subjective. Rule-based techniques are more formal and allow detect patterns or anti-patterns to desired states of an EA. Indicator-based analysis approaches are the most formal techniques and perform a quantitative analysis using values obtained by assessment of observable properties or mathematical formulas.

The "analysis concern" dimension corresponds to the needs to fulfill functional requirements of the organization (e.g. manage customers) or may comprise non-functional requirements such as availability.

The last dimension, "self-referentiality", analyzes if the model includes concepts from EA management processes or not. It concerns the complexity of the model and classifies it in none, if the architectural aspects of the activities of EAM are not considered, single-level, if it includes EAM activities, and multi-level, when it considers meta-processes activities of EA management as part of the model.

In this research, we elaborate on the "analysis technique" and "analysis concern" dimensions, detailing them to classify the recent development of EA analysis research.

- Niemann (2006)

In his work, NIEMANN (2006) describes different types of analysis according to analysis concerns under investigation (dependency, coverage, interface, heterogeneity, complexity, compliance, cost and benefit) and discusses each one separately. Table 2.6 shows the description of each procedure.

Object under investigation	Description of procedure
Dependency	Directly or indirectly (i.e. cross level) linked elements in the EA are selected. Relationships and their impact are shown.
Coverage	The coverage of departments (e.g. units in a process-product matrix) by application system is analyzed.
Interfaces	The interfaces between the application systems are analyzed in terms of their type, number, complexity, frequency/currency, performance, stability, and availability.
Heterogeneity	The heterogeneity of ones IT assets in defined areas of deployment is analyzed.
Complexity	An analysis is run to determine how many components there are in the EA and how many relationships they have.
Conformity	Adherence to standards and ascertainment of the degree of variance (e.g. as a % of the application systems or infrastructure components). Compliance rules.
Costs	Reporting on accumulated production, operation and maintenance costs.
Benefits	Benefits calculation, e.g. as percentaged contributions to the achievement of enterprise goals or via defined key performance indicators (KPIs).

Table 2.6: EA Analysis types according Niemann - Adapted from NIEMANN (2006)

Finally, NIEMANN (2006) presents EA analysis concerns as we do, however, his work does not perform any validation of the resulted concern categories. In addition, a much higher number of analysis concern was identified in chapter 4.

- Andersen and Carugati (2014)

The classification proposals by BUCKL et al. (2009) and LANKHORST (2004) are evaluated ad hoc, by framing a couple of existent works (published until 2009) in their classification models. ANDERSEN; CARUGATI (2014) built their classification supported by a more recent SLR. Though more extensive than the previously cited ones, his work is limited regarding coverage when compared to our work presented in chapter 4. We also introduce a different discussion regarding the SLR results. ANDERSEN; CARUGATI (2014), which has a final set of 45 primary studies, categorizes the concerns of the EA analysis approaches in layers used, and define macro categories to them (e.g. business, technical or financial). The work also presents the approaches' outcomes (model, measurement, method) and which elements their techniques are evaluating (architecture, IT projects, and IT initiatives; services and applications; business elements). However, this classification is still superficial in light of the plurality of methods, techniques, and concerns related to EA Analysis. His study presents ten metrics to demonstrate quantitative measurement approaches (e.g. data accuracy, usage of applications, complexity of EA), while the authors state that only five papers among primary studies employ empirical and qualitative research approach to EA analysis, and claim that qualitative research in EA analysis is scarce. Our results reveal a different reality for that, especially regarding the value and business layers.

Overall, although useful, all the classifications presented by those authors do not provide a broad view of existing analysis techniques and analyzed concerns. In addition, they still lack a representative set of papers to ground them. And this is precisely the gap we aim to solve in chapter 4. We seek to perform a thorough SLR on EA analysis approaches, discuss the concerns studied, techniques and methods applied, models used, research opportunities and gaps, to better understand the field and its potentialities.

2.1.6 EA management

Though frameworks and modeling approaches are important to represent an EA, they are not enough to support the governance of the practice. Enterprise Architecture Management can be defined as the management practice of using EA in a continuous process. EAM defines practices and artifacts used in the process of achieving business-IT alignment. It is a constant and iterative process controlling and improving the existing and planned IT support for an organization (MATTHES et al., 2008).

AHLEMANN (2012) define EAM as "a management practice that establishes, maintains and uses a coherent set of guidelines, architecture principles and governance regimes that provide direction for and practical help with the design and the development of an enterprise's architecture in order to achieve its vision and strategy".

KOTUSEV et al. (2015) perform an extensive literature review aiming to consolidate a definition for EAM. According to the authors, popular EA literature describes the traditional EAM approach as a four-step iterative process: (1) document a current state of an enterprise,

(2) develop a desired future state of an enterprise, (3) develop a transition plan describing how to migrate from the current state to the future state and (4) implement the plan and repeat the process all over again. Nevertheless, significantly different descriptions of EAM also exist in the literature. Those authors also identified two other schools of EAM thinking, besides the traditional one:

- The MIT EAM approach

The MIT approach to EAM was developed in Massachusetts Institute of Technology (MIT) by ROSS; WEILL; ROBERTSON (2006). The MIT approach advocates the development of a core diagram reflecting a long-term enterprise-level architectural vision. The core diagram represents the essence of EA in the MIT approach. The MIT approach relies on the top management setting the architectural direction and the subsequent translation of this direction into concrete project-level decisions. EAM in the MIT approach is an integral part of organizational decision-making processes.

- DYA (dynamic architecture) EAM approach

DYA advocates "just enough, just in time" architecture, no EA is designed until there is a need for it. EAM activities in the DYA approach are triggered by concrete business initiatives appearing in the process of a strategic dialogue. As a response to a new business initiative, architectural services update EA if necessary and prepare a project-start architecture for a new project in order to ensure that this new project fits nicely into existing EA and larger picture. EA in the DYA approach is represented mostly by a set of architectural principles. Detailed architectural diagrams play only a secondary temporary role in DYA.

Following their rationale, KOTUSEV et al. (2015) say that the EAM described in literature do not represent the only possible three stable discrete states of EAM but rather three considerable points in a continuum of possible approaches to EAM ranging from rigid and heavyweight approaches to flexible and lightweight ones, considering three essential elements: development, description and usage, as depicted in Figure 2.8:

Since the EA analysis approach developed in this thesis strongly depends on models, after reflecting about these three EAM schools, we conclude that present research dialogues closely to the traditional EAM approach, which in turn, is aligned with open standard frameworks such TOGAF (The Open Group, 2013), widely accepted in the EA industry community. In this sense, we consider EA analysis as a transversal activity applied throughout the traditional four-step iterative process mentioned earlier. Moreover, EA analysis can be also considered a continuous activity in EAM lifecycle, once it is applied to achieve permanent EA goals such as: (1) to check the EA alignment (between business and IT), one of the main motivations of EA adoption (AHLEMANN, 2012); (2) to analyze and compare scenarios to achieve a target state (TO-BE scenarios) (AHLEMANN, 2012); and (3) to analyze impacts of failure or changes of components in the EA.

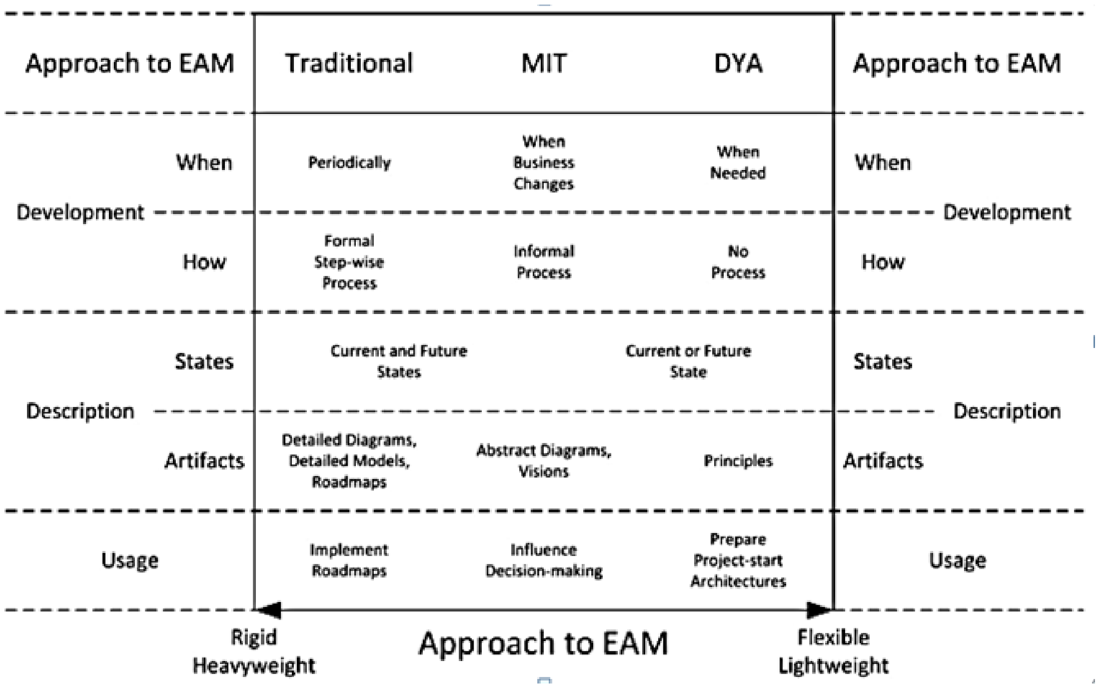

Figure 2.8: Consolidated view of EAM according to the findings of KOTUSEV et al. (2015)

2.1.7 EA frameworks

An EA framework characterizes the organization of the structure (i.e. the concepts and their relationships) and the views associated with the architecture (CAETANO; SILVA; TRIBOLET, 2009). They facilitate the communication since they provide a common vocabulary and are used for many analysis approaches.

Design and maintain coherent EA is a complex effort since it involves many different people with differing backgrounds using various notations (STEEN et al., 2004). To assist architects and managers in the EA practice, several frameworks were created e.g. ZACHMAN (1987), the Department of Defense Architecture Framework (DoDAF) (The DoDAF Architecture Framework Version 2.02, 2010), the Federal Enterprise Architecture Framework (FEAF) (FRAMEWORK, 2013), TOGAF (The Open Group, 2011) and Gartner Framework (JAMES et al., 2005). In the following, we describe briefly three of the most well known ones (adapted from LEIST; ZELLNER (2008)):

- Zachman Architecture Framework (ZAF)

Previously discussed in Section 2.1.1, the ZAF allows identifying models for each viewpoint described in its matrix of 30 cells as described in Table 2.7. His framework provides a model that describes an enterprise holistically, despite being quite idealistic. Furthermore, it is hard to apply because it does not offer any definition of specific products or templates. An additional disadvantage is that there is no process for application of the architecture, so it is difficult to develop architectures. Generally speaking, the framework does not actually comprise any specific method, nor support tool or guidelines for designing and managing an enterprise architecture real case. It is more a taxonomy to describe enterprise architecture components through the use of models/artifacts. The framework´s current version is 3.0 and was edited in 2011 (ZACHMAN, 2011)

Perspectives/ Aspects	Data What	Function How	Network Where	People Who	Time When	Motivation Why
Scope (Contextual) Planner	List of things important to the business	List of process the business performs	List of locations in which the business operates	List of business responsibilities	List of events of significant to the business	List of business goals/strateg y
Business model (Conceptual) Owner	Semantic model	Business process model	Logistics network	Workflow model	Master schedule	Business plan
System model (Logical) Designer	Logical data model	Application architecture	Distributed system architecture	Human interface architecture	Processing structure	Business rule model
Technology model (Physical) Builder	Physical data model	System design	Configuration design	Presentation architecture	Control structure	Rule design
Detailed representations (Out of context) Subcontractor	Data definition	Program	Network architecture	Security architecture	Timing definition	Rule specification

Table 2.7: Zachman framework (ZACHMAN, 1987).

- The Department of Defense Architcture Framework (Department of Defense Architecture Framework (DoDAF))

The DoDAF was developed specifically for the United States Department of Defense to support its war-fighting operations, business operations and processes. It grew from and replaced the previous architecture framework, C4ISRAF. The DoDAF includes guidelines on determining architecture content based on intended use. Architecture development techniques have been provided in DoDAF to specify processes for scope definition, data requirements definition, data collection, architecture objectives analysis and documentation. However, a role model for the development process is also missing in the DoDAF.

- The Open Group Architecture Framework (TOGAF)

The TOGAF is an industry standard architecture framework that may be used freely by any organization wishing to develop enterprise architecture descriptions for the use within that organization. It enables designing, evaluating, and building the right architecture for any organization. The key difference of TOGAF to the previous frameworks is the TOGAF Architecture Development Method (ADM) – a reliable, proven approach for developing enterprise architecture descriptions that meet the needs of the specific business. TOGAF also adopted the Archimate as its default modeling language.

Looking closely at the previous frameworks, we can find many differences in terms of number of layers and their compositions, nomenclatures adopted, among others. A comparison among the three and even other frameworks is out of the scope of this thesis and can be found in SESSIONS (2007).

So far, we discussed the EA related concepts used in this thesis. In the next section, we proceed to study our second theoretical support: the network science.

2.2 Network science and structural analysis

2.2.1 Introduction

Social structure has long been an important concept in sociology and has a strong base in mathematics found in graph theory (Degenne, A. & Forse, 1999). Graphs are the core resource to describe social networks. Basically, a graph consists of a finite set of vertices x1,x2,. . . xn plus the set of edges or arcs that connect them.

In the context of the social network analysis (related to network theory, SCOTT (1992)) graphs are translated to sociograms or social networks, the vertices of a graph are called nodes, and graph´s edges are called ties, connections or relationships. These relationships can be weighed or not (binary); they can be directed or undirected relations; and may have a positive meaning or a negative one (signed). In Figure 2.9, a sociogram with 234 (two hundred and thirty-four) nodes and their relations is depicted. In this case, the relations are binary, undirected and unsigned.

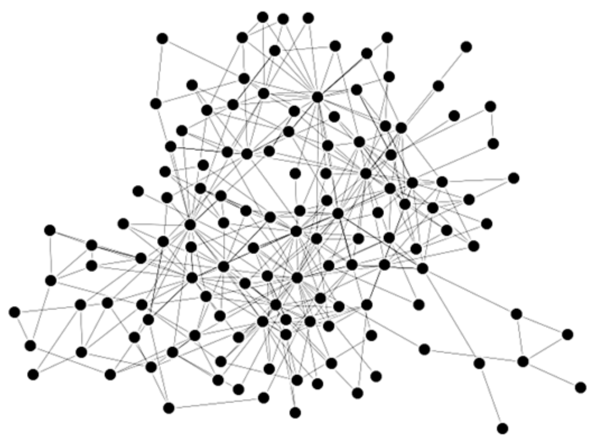

Figure 2.9: Undirected network generated from the Datatset02, detailed in chapter 3

Network analysis has moved from being a suggestive metaphor to an analytic approach to a paradigm, with its own theoretical statements, methods, network analysis software, and researchers (WASSERMAN; FAUST, 1994). Several real world systems – such as biology, power networks, the internet - are modeled and studied as networks (EROL; SAUSER; MANSOURI, 2010; SIMON; FISCHBACH, 2013; WANG et al., 2011). They typically either study entire networks (also known as complete networks), all of the ties containing specified relations in a defined population, or personal networks (also known as egocentric networks). The distinction between whole/complete networks and personal/egocentric networks has depended largely on how analysts were able to gather data (WASSERMAN; FAUST, 1994).

Bringing these concepts from network analysis and graph theory to the information systems field, we can model an application module, a single application, a technology, or a business process as a node. Also, some meanings for the relationships can be established: one can model a simple communication or information flow, a dependence relation (consumer/producer), or a common shared technology, for example. Therefore, possibly modeled relations are: application A1 "sends data to" application A2, meaning that there is a directed edge from node A1 pointing to node A2; technology T1 "realizes" application A1; or application A1 "supports" business process B1. In those cases, a typical analysis is to investigate which systems occupy significant structural positions on a network or how they are embedded in the overall structure.

In the next sections, we discuss the fundamental concepts in this thesis regarding network analysis.

2.2.2 *Network measures*

There are several network measures, methods or algorithms that can be applied to extract potential useful information from networks. They identify important structural positions, group formations, the degree of centralization network among others structural indicators (HANNEMAN; RIDDLE, 2005). FREEMAN (1978) argues that centrality degree, betweenness, eigenvector and closeness are related to different aspects of centrality such as control, brokerage, or independence. Hence, those centrality metric offer four different ways of identifying how network structure might differentiate the roles of components. In Table 2.8, we illustrate possible interpretations for network measures in EA context.

Now that we described some of the main structural analysis measures, we need to think about the importance of those measures for the context of analysis. The main rationale behind the centrality analysis is that, when nodes occupy a central position they have the ability to influence the entire network or at least part of it. As a result, they gain importance within the architectural context. Such might have to be managed closely or they could adversely impact the network as a whole (IYER; DREYFUS; GYLLSTROM, 2007). In another level of analysis, the ability to detect groups and sub-structures could have vital practical importance: knowing how a node is embedded in the structure of groups within architecture may also be critical to understanding its behavior. For example, some systems or web-services may act as "bridges" between groups of other systems. Others may have all of their relationships within a single group (locals). Some nodes may be part of a tightly connected and closed module, while others are completely isolated from this group acting as a satellite system. Such differences in the

Network measure	Contextualization of the measure
Degree	In directed networks, connecting the provider of information with the consumer can allow us to analyze in-degree and out-degree centralities. In-degree centrality indicates the degree to which consumers depend on data provided by others while out-degree centrality indicates the degree to which producers generate the information flow.
Betweenness	Betweenness can identify components that work as bridges and connectors (SCOTT, 1992) in the components network.
Out-closeness	Closeness has a significant effect on how coupled a component is to all other components, not just to those that integrate with it directly (DREYFUS; WYNER, 2011). Out-closeness, in a directed network, can indicate the range of reachability of a component to the entire network. We apply out-closeness in our datasets.
Eigenvector	Eigenvector centrality can be considered an extension of degree centrality that privileges components connected to other well-connected components (SCOTT, 1992). Thus, this measure can identify components that together constitute global structural points.

Table 2.8: Network measures and their contextualization in EA. Based on SCOTT (1992); SIMON; FISCHBACH (2013)

ways that nodes are embedded in the structure of groups within in a network can have profound consequences for the ways that these nodes impact their "network," and how they are perceived (HANNEMAN; RIDDLE, 2005).

For the sake of organization of this document, we provide a detailed description of network metrics and other related concepts used in this thesis in Appendix A. For an even more detailed study of network measures, see WASSERMAN; FAUST (1994).

2.2.3 Levels of network analysis

According to SCOTT (1992) and HANNEMAN; RIDDLE (2005), the network analysis can be done at different levels of abstraction: component level, group or modular level and network level. First, the component level analysis takes components individually. According to our SLR detailed in Chapter 5, this is the most frequent type of network analysis applied in EA context. Expanding the level of analysis, methods and metrics in group analysis category take a set of components into account. The overall network can be considered in the analysis: this is the network level analysis. Those analysis levels are explained as follows:

- Component level

Description: Measures applied at this level calculate individual values for each component in the network.

Example of analysis initiative: Degree centrality (SCOTT, 2012) calculates a value for a node in the network, counting the edges that connect that node to other ones. For instance, Figure 2.10 depicts a central node as a bigger one in the network.

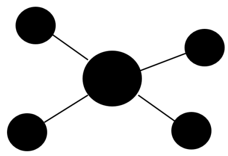

Figure 2.10: The biggest node has the biggest centrality value of network, since it is connected with other four nodes

- Dyad/triad level

Description: The smallest social structure in which an individual can be embedded is a dyad (that is, a pair of actors) (HANNEMAN; RIDDLE, 2005). An expansion of this concept, triad is the smallest social structure that has the true character of a "society" - any "triple" A, B, C of actors (HANNEMAN; RIDDLE, 2005).

Example of analysis initiative: Measures applied at this level calculate minimal structures in the network represented by pairs or triples of components, which can be directed or not. We can characterize the whole population in terms of the prevalence of these dyadic "structures" (HANNEMAN; RIDDLE, 2005). Triads are also the simplest structures in which we can see the emergence of hierarchy among components. We show triads and a dyad in Figure 2.11.

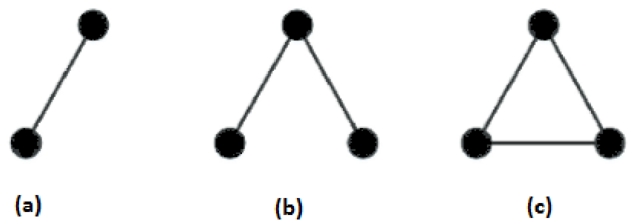

Figure 2.11: Dyad and triads: (a) - dyad (b) - triad with two relations (c) - triad with three relations

- Module, group or cluster level

Description: Measures or methods applied at this level deal with several components at the same time. They can find clusters or groups of nodes that are well interconnected or evaluate those clusters regarding modularity, for example.

Example of analysis initiative: Clustering algorithms in general (e.g GIRVAN; NEWMAN (2002)).

- Network level

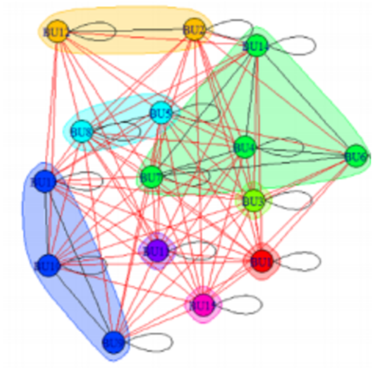

Figure 2.12: Clusters detected in the network are represented by different colors.

Description: At this level, measures or methods generate values considering the overall network of nodes/components.

Example of analysis initiative: Network density (SCOTT, 1992).

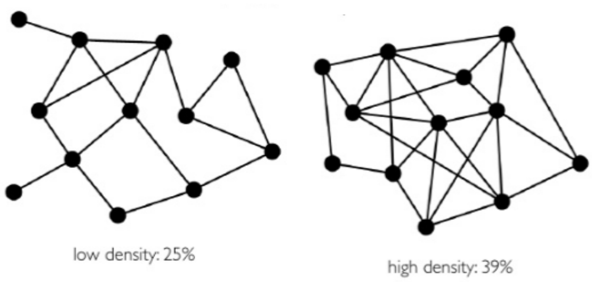

Figure 2.13: Two different networks with different values for density.

2.2.4 Unimodal,bi-modal and multi-modal networks

According to HANNEMAN; RIDDLE (2005), modelers can decide to build networks formed by just one type of component, called uni-modal networks (e.g. application network or layer) or networks composed of two types of components,called bi-modal networks (e.g application x business process network or layers). There is yet, the possibility of use several types of components (more than two) in a multi-modal network (e.g. business process x application x technology network). The term “mode” refers to a class of components of the same kind (e.g., applications, routers, business process). Those concepts are summarized next (HANNEMAN; RIDDLE, 2005):

- Uni-modal network

Description: The analyzed network is composed of just one type of component.

Example: One example of those is the “application network”, whose nodes are all applications.

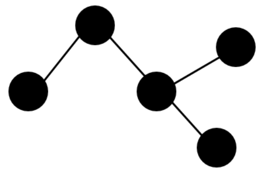

Figure 2.14: Uni-modal network

- Bi-modal network

Description: Networks composed of two types of components or nodes.

Example: One example of a bi-modal network is the “Application x Business process” network. (relations between applications which support business processes are mapped).

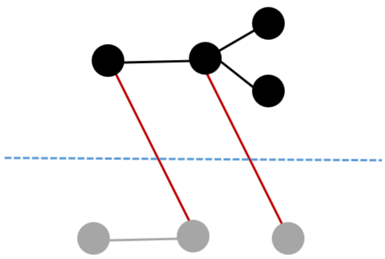

Figure 2.15: bi-modal network

- Multi-modal network

Description: Networks composed of more than two types of components or nodes.

Example: One example of a multi-modal network is the “Technology x Application x Business process” network. (relations between technologies which applications, which in turn, support business processes are mapped). In Figure 2.16, we show three diferent types of nodes represented by different colors.

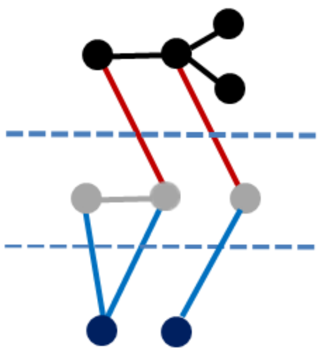

Figure 2.16: Multi-modal network

2.2.5 *Outputs of network analysis*

Regarding outputs, network analysis initiatives can either generate one of the options described next.

- Graph

Description: Metrics or methods' outputs can be plotted as graphs, with different sizes and colors of nodes, edges, for example.

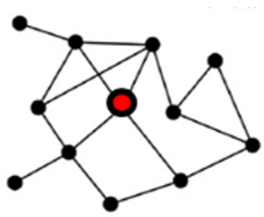

Figure 2.17: Example of a graph.

- Heat maps

Description: It represents the network information with a colored matrix. Each cell is colored according to the interaction level between the components involved.

Figure 2.18: Example of a heat map.

- Ranking lists

Description: As results, network analysis initiatives can produce a ranking with the ten components with the highest values for degree centrality, for example.

Order	Overall value
R_5	0.1082
R_{14}	0.0862
R_{11}	0.0825
R_1	0.0798
R_{15}	0.0794
R_8	0.0726
R_4	0.0688
R_{13}	0.0676
R_3	0.0658
R_2	0.0647

Figure 2.19: Example of a ranking list.

2.2.6 Representation of network models

Network analysts use two kinds of tools from mathematics to represent information about patterns of ties among social actors: graphs and matrices (HANNEMAN; RIDDLE, 2005). The traditional visual representation of a network is a graph. Analysts using matrices to represent social networks often dispense with the mathematical conventions, and simply show their data as an array of labeled rows and columns. For instance, in Table 2.9, we depict a 4-node network and their relations.

	A	B	C	D
A		1	0	0
B	1		1	0
C	1	1		1
D	0	0	1	

Table 2.9: Matrix representation of a network

This modeling representation is very explored in the design structure matrix (DSM) research (BROWNING, 2001). We discuss in details this modeling approach in chapter 7.

2.2.7 EA network analysis (EANA)

Networks provide two benefits to understand architecture. First, they provide a good metaphor to communicate architectural issues. Second, networks are amenable to certain analytics that can help the architect compute relevant metrics (IYER; DREYFUS; GYLLSTROM, 2007). The architect that is interested in such variables as system flexibility, robustness, adaptability, and performance should be interested in the ongoing evolution of the information system architecture as an emergent phenomenon typically studied with network analysis (DREYFUS; IYER, 2006).

To the best of our knowledge, DREYFUS; IYER (2006) was the first work to explicitly advocate a network-based view for EA (dependency) analysis. Those authors wanted to provide architects with support tools to communicate and make decisions about architecture, describing the data requirements and algorithms that could be used to build a decision support system that enables enterprises to incorporate a network perspective in their decision-making process.

In our research, as in DREYFUS; IYER (2006); IYER; DREYFUS; GYLLSTROM (2007); DREYFUS; WYNER (2011); SIMON; FISCHBACH (2013), we dedicate special attention to the EA modeled as a complex network to investigate the possibilities for network based metrics and methods as EA analysis techniques.

With constructs from network analysis, researchers can model EA components as networks/graphs and apply measures and/or methods grounded in network analysis (SCOTT, 1992) to identify critical structural components, considering only the relations among them. The complexity of analysis arises when we consider that changes within one architecture or layer may affect the overall enterprise or may be subject to interdependencies with other architectures. In this sense, we argue here, as proposed by HANSCHKE (2010), that it is possible to establish the links among the various sub-architectures or architectures to create a complete view of business and IT structures.

2.2.8 EA network data

Network data are defined by actors and by relations (or nodes and ties, respectively). Network analysis focuses on the relationships among nodes, and not individual nodes and their attributes (HANNEMAN; RIDDLE, 2005). Therefore, under the EA network paradigm, the data is essentially about **EA components** (nodes) and **connections among those components** (relations). In this sense, EA components like strategy goals, business process, information entities, applications and infrastructure elements can be represented individually as nodes. For instance, in our EANA SLR, we found EA components such as stakeholder, business unit, data object, application, technology mapped as network nodes.

The connections or relations among EA components can assume several meanings in the data modeling. Examples also found in the review were software component "is dependent on" software component, application "supports" product, application “communicates with" application, application “runs on" infrastructure component and so on. **Each relation meaning might be associated with a particular goal or concern of analysis**. For instance, modeling choices of relations similar to application "depends on" application might be associated with risk of failures or dependability concerns. Relations can be made by components of the same domain **(intra-domain relation)** or different domains **(inter-domain relation)**.

Regarding the kind of data used in the analysis process, we classify it into two types. The first one is named **primary data** and represents the data gathered directly from documents, models, interviews with stakeholders in the enterprise. This kind of data may also be automatically generated through the analysis of the data flow between applications with some tool support. The second data type is the **derived data**, explained in details in the next section, since we will actively work on this type data to propose derivation mechanisms in Chapter 7.

2.2.8.1 Derived data and affiliation networks

Derived data comes out from the process of taking components and relationships modeled from primary data sources and create new artificial "relations" which in turn, form new derived networks. This data derivation process is supported theoretically by the concept of **"co-affiliation network"** (BORGATTI; HALGIN, 2011). In social network analysis, the term "affiliation" usually refers to membership or participation data, such as when we have data on which actors have participated in which events. Often, the assumption is that co-membership in groups or events is an indicator of an underlying social tie (BORGATTI; HALGIN, 2011). We can represent affiliations as graphs in which nodes correspond to entities (such as women and events) and lines correspond to ties of affiliation among the entities. Affiliation networks are necessarily bi-modal networks. In Figure 2.20, we depict the affiliation mechanism composed of the thicker and directed edges ("supports") between APP1 and P1, APP2 and P1. In this case, we say that APP1 and APP2 are co-affiliated with P1.

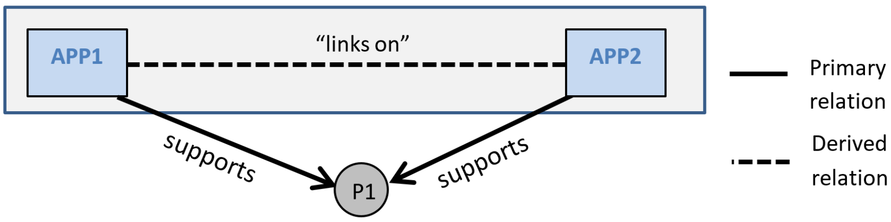

Figure 2.20: Visual representation of the affiliation network and data derivation process.

Affiliations graphs are distinctive in having the property of bipartiteness, which means that the graph's nodes can be partitioned into two classes such that all relations occur only between classes and never within classes (BORGATTI; HALGIN, 2011). In our example, we only considered relations between the classes "Business" and "Application" but not within each of those classes of nodes.

Once can construct a relation among members of a node set simply by defining **co-affiliation** (e.g., attendance at the same events, membership on the same corporate board) as a relation. In our example, the co-affiliation relation is "be supported by the same application". We thus construct a derived relation between the two applications based on the co-affiliation with the business process "P1". This new relation is depicted in Figure 2.20 with the dashed edge ("links on"). For BORGATTI; HALGIN (2011), one justification for relying on co-affiliation is the idea that co-affiliation provides the conditions for the development of social ties of various kinds. For example, the more often people attend the same events, the more likely it is they will interact and develop some sort of relationship. In our example, the co-affiliation may be taken as a proxy for the intensity of work collaboration of the people involved with APP1 and APP2, regarding the business process execution of P1. We ellaborate on this mechanism to propose a set of derivation operators for the EA context, presented in chapter 7.

2.3 Related work

As for related works, first, we consider the research which applied the network paradigm in correlated fields such as systems engineering, product engineering and project management. We thus select a few works which consolidated structural analysis combining them with architectural concepts in these specific contexts.

HOLLAUER; WILBERG; OMER (2015) reunited conceptual elements extracted from agent-based and dynamic models already used in companies to select key conceptual elements and build a matrix-based framework with them. Next, those authors analyze how the model´s elements relate to each other to suggest analysis questions to support the management and planning of dynamic modeling activities during product development. In his Ph.D. thesis, BARTOLOMEI (2007) examined in detail several frameworks for system engineering modeling and found out that, all of them, taken individually, were limited in scope. Thus, he presented an improved framework called Engineering Systems Multiple-Domain Matrix (ES-MDM), bringing a broader scope to the system analysis. His framework covers not only technical, organizational or component aspects of the systems, but also time and environmental ones, composing six domains which can model any system. Similarly to the model of HOLLAUER; WILBERG; OMER (2015), the ES-MDM is an adjacency matrix with identical row and column headings. Each row/column of the ES-MDM represents a "view", a particular aspect or a set of concerns. Moreover, the ES-MDM establishes a basis for view-view interactions, giving a more formalized, theoretically grounded perspective of an engineering system. Bartolomei's framework is depicted in Figure 2.21:

In his Ph.D. thesis, KREIMEYER (2009) worked with business process modeled as multiple-domain-matrices (MDMs) and focused on the process architecture at a detailed level of granularity. Similar to what we do in the present thesis, he used DSMs and structural metrics as a management toolbox to find structural weak spots (points of improvement) and outliers (only) in the process architecture.

BROWNING (2009) proposes the application of architectural framework concepts to the management of the work done to develop a complex system product. According to Browning, the architecture framework provides a portfolio of views of a complex system, each of which describes it partially and in a format meaningful to its users and their particular needs. Similarly, in BROWNING (2010), a more specific architecture framework (its views, analysis purposes, information attributes etc.) is proposed specifically for the for project management context.

Back to the EA analysis context, it is worthy of mentioning the work of NARANJO; SÁNCHEZ; VILLALOBOS (2014), which presents a general catalog for EA analysis metrics and methods. The difference to our EANA library is that the latter is structural-based only while the former is generic.

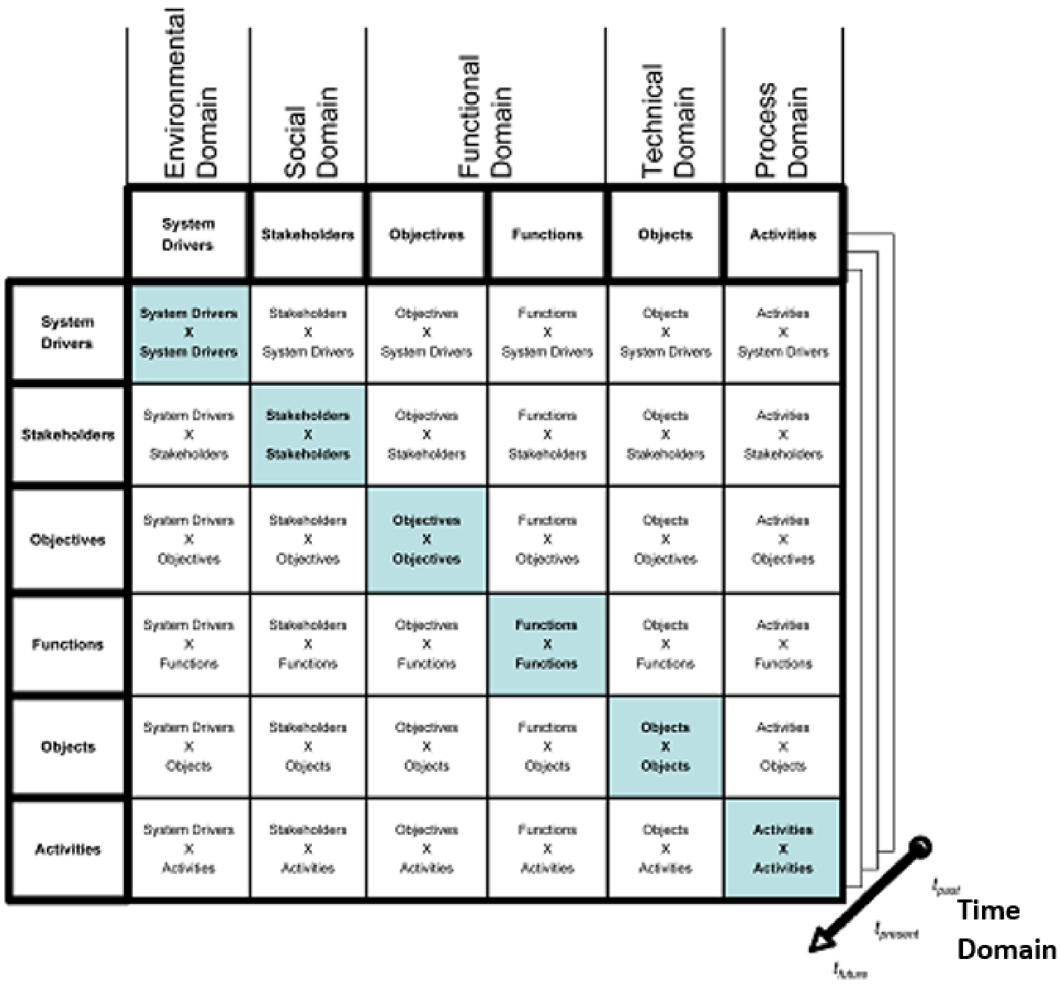

Figure 2.21: The Engineering Systems Multiple-Domain Matrix (ES-MDM) (Bartolomei,2007).

Specifically in the EANA context, before the work of DREYFUS; IYER (2006), punctual research initiatives already have applied structural analysis to the application or business process layers, for instance, without a holistic perspective awareness. In those papers, the notion of EA was not explicitly discussed. However, it was possible to identify architectural analysis at some levels on them. Since 2006 to date, a few other research endeavors also applied network analysis in EA. All of them are organized in the EANA library described in chapter 5. VAKKURI (2013), in his master thesis, examines how the design structure matrix (BROWNING, 2001) could be used in managing an enterprise's information systems portfolio, and offers a practical tool for structural analysis supported by clustering algorithms.

Some of the previous works´ ideas touched some points approached in our thesis. For instance, the combination of architectural principles and structural analysis was aimed in BARTOLOMEI (2007); KREIMEYER (2009); BROWNING (2010) in their specific fields.

The idea of building an EA analysis library was already presented in NARANJO; SÁNCHEZ; VILLALOBOS (2014) without any specific analysis focus, and in KREIMEYER (2009) with a structural focus, but in product engineering field.

Nevertheless, the EA structural analysis state of art did not count with any broad and consolidating research approach so far. The elements which compose our EANA framework like the EANA meta-model, the EANA library and the EANA analysis process, were not designed before, thus, are innovative for their context.

2.4 Chapter Summary

In this chapter, we described how the concept of EA evolved since its early studies, despite presenting conceptual gaps, yet. We introduced the main theoretical concepts from EA and network science that underpin this research. Moreover, we discussed how this research proposal is positioned regarding the related literature. We take the network analysis conceptual elements presented in this chapter to design the EANA-MM in chapter 5. In the next chapter, we introduce the methodological choices for this thesis.

3

RESEARCH DESIGN

Chapter 2 presented the theoretical foundations of the present research. This chapter aims to build on the previous chapters to describe our general research strategy. In section 3.1, the main research and secondary questions are reviewed. In section 3.2, the research is qualified according to its nature, goals, methods and so on. Next, the research methods are explained in section 3.3. The data collection methods and data sources are described in section 3.4. In section 3.5, the attention is directed to explain the evaluation methods applied to the proposed artifacts. Finally, the research phases are described in the summary of the chapter.

3.1 RESEARCH QUESTION

As stated before, in the EANA research context, **there is a need for a foundational classification of the primary constructs of EANA, a mapping of which initiatives do already exist and what are its objects of interest in terms of analysis**. In a second moment, the concern should be **to describe how this body of knowledge could be used by researchers and practitioners** to advance its knowledge base and practice, perhaps, prescribing an EANA process in a step-by-step fashion.

The present work aims to explore how network measures can contribute to the EA analysis function. In that direction, we investigate past research and also suggest constructs to organize existent knowledge and set the foundations for these two overlapped research fields. The general purpose of this research is to answer the following question:

How to perform analysis of EA components and their relationships supported by network measures?

In order to answer this general question, we divided it in secondary questions that allow us to plan more objective, organized and coordinated research effort.

- RQ1 Which EANA initiatives are already available?
  - RQ1.1 What are the concerns, models and approaches analyzed in EA?
  - RQ1.2 What are the concerns, models and approaches analyzed in EANA?

- RQ2 What are the information requirements to perform EANA?
    - RQ2.1 What are the inputs needed to perform EANA?
    - RQ2.2 what kind of data can be used in the EANA?
- RQ3 How EA network analysis initiatives can be classified?
- RQ4 How can experts be guided in performing EANA?

We explain how we achieve the answers for those questions in the next sections.

3.2 RESEARCH CHARACTERIZATION

The choice of the research design that guides the researcher must be aligned with the type of the research and its goals (REMENYI; WILLIAMS, 1998). According to CRESWELL (2013), a researcher should make use of a methodology to guide the research project from identifying the epistemological position underlying the researcher's philosophical attitude against the object of research until the procedures for collecting and analyzing the data. In the following, we use a combination of taxonomies of GIL (2010), COLLIS; HUSSEY (2009) and WOHLIN; AURUM (2015) to classify either our intermediary research outcomes or the research as a whole unit.

3.2.1 Research outcomes

Considering the **nature of research** (GIL, 2010) or **research outcomes** (WOHLIN; AURUM, 2015), **basic research** (a.k.a. pure research) is applied to a problem where the emphasis is the understanding of the problem rather than providing a solution to a problem, hence the main contribution is the knowledge generated from the research (WOHLIN; AURUM, 2015). It does not present an immediate goal or application and produces knowledge to be used by other researchers. Therefore our "what" and "which" research questions (RQ1, RQ2) suggest a basic research is suitable for our case.

Our research also has characteristics of **applied research** (GIL, 2010), since it aims to generate knowledge to be applied in a specific real problem (i.e, the EA analysis in real cases). Our research involves the application of NAIs through the creation, application, and evaluation of artifacts and methods for a specific context - to support EA analysis. For HEVNER et al. (2004), the goal of the applied research is to produce utility, which can be evaluated by its application in the specific environment, expanding the knowledge frontiers about the problem and finally disseminate it among the community.

3.2.2 Research logic

Research logic refers to in which direction the research proceeds in terms of whether it moves from general to specific or vice versa and can assume two values: inductive versus deductive research (CRESWELL, 2009). Inductive research is based on inductive arguments. It moves from the specific to the general. The researcher infers theoretical concepts and patterns from observed data. Considering the **research logic**, our study is partly **inductive** since we create knowledge after a deep study of the literature. Deductive research approach explores a known theory or phenomenon and tests if that theory is valid in a given circumstances. We follow a **deductive** approach when we try to apply our generated knowledge (DSR artifacts) and evaluate them (RQ3 and RQ4). The induction is more present than deduction in our research activities.

3.2.3 Research goals or purpose

With regard to the research goal or purpose, GIL (2010) argues that the nature of research can be classified into three groups: exploratory, descriptive and explanatory studies. This research is **exploratory**, due to the relatively recent development of the field (enterprise architecture analysis with network metrics), happened mainly in the last ten years. The exploratory research normally aims to clarify a phenomenon or a situation In order to develop the awareness of it, allowing defining research hypothesis and look for causes and consequences (CHIZZOTTI, 1995). Our research is also **descriptive** because it defines a particular phenomenon using specific and standard techniques of data collection and analysis (RQ1, RQ2). This research is not explanatory since we are not concerned about present reasons or “whys” that explain phenomena.

COLLIS; HUSSEY (2009) adds a fourth research goal type: **evaluation**. Evaluation research aims to determine the impact of methods, tools, or frameworks that may encompass the other three research purposes: exploratory, descriptive and explanatory research (ENGEL; SCHUTT, 2009). RUNESON; HÖST (2009) apud WOHLIN; AURUM (2015) denote this type of research as “improving” in the engineering context. Therefore, our research can also be classified as evaluation research, since a set of important contributions is represented by the artifacts and their evaluations(RQ2, RQ3 and RQ4).

3.2.4 Research approach or philosophy

EASTERBROOK et al. (2008) point out that even though many researchers (at least, in software engineering) avoid addressing underlying philosophies in their **research approach**; they tend to use one of the following four: positivism, constructivism (or interpretivist), critical research or pragmatism. CRESWELL (2009) refers to those approaches as knowledge claims positions. Our research is based on **positivisty** philosophy. The desire to characterize an entire population (the literature) via sampling techniques requires a belief in reductionism, and a concern with generalizable theories(EASTERBROOK et al., 2008). Those goals are valid for our two systematic literature views (RQ1.1 and RQ1.2).

We also have **pragmatic or design** knowledge claims. For CRESWELL (2009), in that case, there is a concern with applications -"what works” -and solutions to problems. Instead of methods being important, the problem is most important, and researchers use all approaches to understanding the problem. Investigators use both quantitative and qualitative data because they work to provide the best understanding of a research problem. Pragmatism is less dogmatic than the other three stances described above, as pragmatists tend to think the researcher should be free to use whatever research methods shed light on the research problem (CRESWELL, 2009). In this line, our proposed artifacts represent pragmatic knowledge contributions to the EANA knowledge base.

Finally, VAISHNAVI; KUECHLER (2005) present a comparison among three philosophical assumptions of the three research perspectives described in Table 3.1:

	Research perspective		
Basic belief	**Positivist**	**Interpretive**	**Design**
Ontology	A single reality. Knowable, probabilistic	Multiple realities, socially constructed.	Multiple, contextually situated alternative world-states. Socio-technologically enabled.
Epistemology	Objective. Dispassionate. Detached observer of truth.	Subjective, i.e. values and knowledge emerge from the researcher-participant interaction.	Knowing through making: objectively constrained construction within a context. Iterative circumscription reveals meaning.
Methodology	Observation; quantitative; statistical	Participation; qualitative. Hermeneutical, dialectical.	Developmental. Measure artefactual impacts on the composite system.
Axiology: what is of value	Truth: universal and beautiful; prediction	Understanding: situated and description	Control; Creation; progress (i.e. improvement); Understanding.

Table 3.1: Philosophical assumptions of the three research perspectives

3.2.5 Research process

In terms of research process (WOHLIN; AURUM, 2015), this research is **predominantly qualitative** due to its intense use of qualitative data: (1) through surveys on literature in an attempt to raise a better understanding of the phenomena and to build conceptual constructs and models; (2) to design and propose analysis methods for EA; or (3) to analyze experts ´perception about the artifacts. According to SEAMAN (1999), the principal advantage of using qualitative methods is that they force the researcher to delve into the complexity of the problem rather than abstract it away. Although, the research does have a **quantitative component** since it uses network measures to produce essentially numbers as outputs. Despite not applying statistical methods, we look at those measures ´outputs with **qualitative lens** to interpret their meaning in the context of EA.

3.2.6 Research procedures or methodology

We consider the terminology of GIL (2010) to say that part of this research is **a bibliographic**. This kind of research is performed through the consolidation of theoretical references published in journals, proceedings, books, websites and so on. Part of our research is also **empirical** (GIL, 2010). The empirical research is characterized by data collection in additional information sources like people, documents from organizations, electronic sensors etc. If we consider the taxonomy adopted by WOHLIN; AURUM (2015), we can also say that this research has **design science** as a research methodology. HEVNER et al. (2004) introduce design science research as a problem-solving process, which requires a creation of an artifact for a specific problem in which the artifact needs to be innovative, be effective and needs to be evaluated by applying rigorous approaches.OSTROWSKI; HELFERT (2011) apud WOHLIN; AURUM (2015) found in their research that 78% of researchers constructed artifacts based on literature reviews and working with practitioners while 22% of researchers constructed artifacts from the literature review only. In our case, we use theoretical findings combined with empirical insights to design methods to evaluate EA analysis and validate them with practitioners.

3.2.7 Data collection and analysis methods

We use **SLR guidelines** from KITCHENHAM (2007) to collect and organize data from the literature. Based on **thematic analysis**, we intend to define our initial set of concepts and make some relations among them. We also perform **documental analysis** in 3 different organizations in order to model their EA. Experts provided access to additional documentation when necessary.

3.3 RESEARCH METHODS

The three research methods used in this thesis are described in the following.

3.3.1 SLRs

3.3.1.1 Basics of the method

According to PETTICREW; ROBERTS (2006), SLR are reviews that adhere closely to a set of scientific methods that explicitly aim to limit systematic error (bias), mainly by attempting to identify, appraise and synthesize all relevant papers in order to answer a particular question (or set of questions). SLRs adopt a particular methodology with the overall aim of producing a scientific summary of the evidence in any area. In this respect, systematic reviews are simply another research method, and in many respects they are very similar to a survey – though in this case, they involve a survey of the literature, not of people (PETTICREW; ROBERTS, 2006).

KITCHENHAM (2012) affirms that an SLR is characterized by a set of special structural elements: a review protocol, a predefined search strategy, explicit criteria for the inclusion and exclusion of studies, a quality assessment procedure, and the documentation of the whole process execution. Nevertheless, there are some differences between the several types of systematic reviews, such as narrative review, conceptual review, rapid review, realistic review, scoping review, traditional review, critical review and others (PETTICREW; ROBERTS, 2006).

KITCHENHAM (2007) typifies two literature reviews that are of our particular interest:

- Conventional SLRs. Aggregate results about the effectiveness of a treatment, intervention, or technology, and are related to specific research questions. When sufficient quantitative experiments are available to answer the research question, a meta-analysis can be used to integrate their results;

- Mapping studies (MS). Aim to identify all research related to a specific topic, i.e., to answer broader questions related to trends in research. They are recommended when there is little evidence or that the topic is very broad. A systematic mapping study allows the evidence in a domain to be plotted at a high level of granularity.

The main differences between a mapping study and a (conventional) systematic review are that first one has broader research questions driving them and often ask multiple research questions. The search terms for mapping studies will be less highly focused than for systematic reviews and are likely to return a large number of studies. For a mapping study, however, this is less of a problem than with large numbers of results during the search phase of the systematic review as the aim here is for broad coverage rather than narrow focus (KITCHENHAM, 2007).

3.3.1.2 Operationalization of the method

According to the definitions previously presented, we opted for carrying out one mapping study and one conventional SLR, on EA analysis approaches and EANA approaches, respectively. Our goal is to collect evidence that can be used to answer our research question RQ1("which EANA initiatives are already available") and RQ2 ("what are the information requirements to perform EANA"). For our both literature reviews, we follow rigorously the stages suggested in the guidelines of KITCHENHAM (2007), to define analytic categories and classify the reviewed literature. Its three macro-stages are described in Figure 3.1.

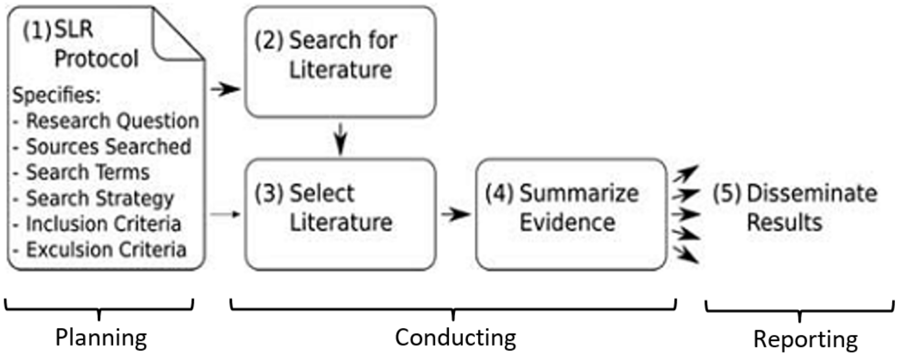

Figure 3.1: Overall SLR process adapted from BOELL; CECEZ-KECMANOVIC (2015) and KITCHENHAM (2007)

3.3.1.3 Limitations of use

The most common limitations in a systematic review are related to the possible biases produced during the selection process and inaccuracies of the data extraction. We strongly based our work on the guidelines of KITCHENHAM (2007) and we were as careful as possible regarding the coverage of the review. All reasons for inclusion and exclusion of studies at each stage were recorded, as recommended by KITCHENHAM (2007). To minimize the bias in inclusion and exclusion checks and data extraction a second researcher was introduced: a bachelor student who was trained regarding the SLR guidelines, before taking part in the research itself. The research protocol was also developed and constantly checked by the researchers. We present the main information of both SLR protocols and the results in chapter 4 and chapter 5.

3.3.1.4 Applying the method

We perform two bibliometric studies to get to know and describe the state of the art of our research field.

Our bibliometric study was divided into two separate research efforts. One bibliometric study (a mapping study) was executed to get to know more about what has been studied in the overall EA analysis field. Its focus thus is on studies which analyze EA, with their different approaches (not only with network measures), constituting a broader perspective of EA analysis field. However, in this study, we apply a more abstract analysis if compared to our second literature review. More details can be found in chapter 4.

The second bibliometric study,a SLR, was conducted to identify network measures applied in EANA context. This research initiative resulted in 74 network measure initiatives found in 29 selected papers, which are detailed in chapter 5.

3.3.2 Design Science Research (DSR)

According to CLEVEN; GUBLER; HÜNER (2009), Information Systems (IS) research is largely characterized by two paradigms, namely behavioral science and design science. While behavioral science concentrates on the development and verification of theories, design science focuses on the development of solutions for practical problems and, thereby, on accomplishing utility.

Since the publication of HEVNER et al. (2004), DSR has been the subject of growing attention within the IS community. It is now firmly established as a research method while still being at the maturation stage (FISCHER (2011) apud PRAT; COMYN-WATTIAU; AKOKA (2014)).

Design science research is a fast evolving field. Within the last decade even the most commonly accepted name for the field has changed - from Design research (DR) to Design Science Research (DSR) (VAISHNAVI; KUECHLER, 2005). The DSR paradigm has its roots in engineering and what Herbert Simon termed “the sciences of the artificial. To bring the design activity into focus at an intellectual level, Simon makes a clear distinction between "natural science" and "science of the artificial" (also known as design science): A natural science is a body of knowledge about some class of things - objects or phenomenon - in the world (nature or society) that describes and explains how they behave and interact with each other. A science of the artificial, on the other hand, is a body of knowledge about artificial (human-made) objects and phenomena designed to meet certain desired goals(VAISHNAVI; KUECHLER, 2005).

Therefore, DSR focuses on creations of artificial systems and seeks to create innovations, or artifacts, that embody the ideas, practices, technical capabilities and products required to efficiently accomplish the analysis, design, implementation, and use of information systems.

For OSTROWSKI (2012), DSR does not limit itself to the understanding, but also aims to develop knowledge on the advantages and disadvantages of alternative solutions. Literature reflects healthy discussion around the balance of rigor and relevance in DSR, which reflects it as a still shaping field.

3.3.2.1 Basics of the method

- Artifact and context

This is a fundamental concept for the understanding of the thesis. The term **artifact** is used to describe something that is artificial or constructed by humans, as opposed to something that occurs naturally (SIMON, 1996). Such artifacts must improve upon existing solutions to a problem or perhaps provide a first solution to an important problem in a certain context (HEVNER; CHATTERJEE, 2010). Examples of artifacts designed and studied in information systems and software engineering research are algorithms, methods, notations, techniques, and even conceptual frameworks (WIERINGA, 2010).

The **context or environment** defines the problem space (SIMON, 1996) in which reside the phenomena of interest. For IS research, it is composed of people, (business) organizations, and their existing or planned technologies (SILVER; MARKUS; BEATH, 1995). In it are the goals, tasks, problems, and opportunities that define business needs as they are perceived by people within the organization (HEVNER; MARCH; RAM, 2008). Business needs are assessed and evaluated within the context of organizational strategies, structure, culture, and existing business processes. They are positioned relative to existing technology infrastructure, applications, communication architectures, and development capabilities. Together these define the business need or "problem" as perceived by the researcher (HEVNER; MARCH; RAM, 2008). Figure 3.2 illustrates the relation between artifacts and their contexts.

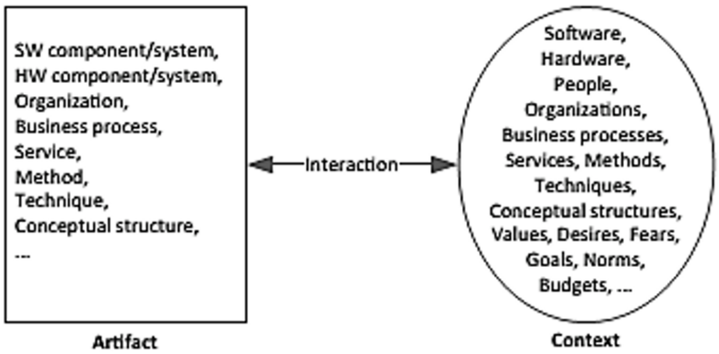

Figure 3.2: Relation between artifact and context (WIERINGA, 2010)

- Types of artifacts (HEVNER et al., 2004)

Instantiation: it is a realization or implementation of an artifact on its environment. It operationalizes constructs, models, and methods. In some cases, an instantiation can even precede their conception. Examples are software or hardware implementation.

Construct: they form the conceptual base used to characterize an investigated phenomena. They describe problems from a specific domain and their solutions as well. They can be used as a shared vocabulary of a stakeholder community (research or practitioners). Examples: the notion of entities, objects, data types.

Model: it is a set of propositions or declarations which express the relations among the constructs. Models can be understood as a description or a representation of how things are. Examples: UML diagrams, use case scenarios, storyboards.

Method: it is a set of steps (an orientation or algorithm) used to execute a task. Methods are based on a set of basic constructs and a representation (model) of the problem space (SIMON, 1996). Examples: algorithm or manual.

Design theory: it can be reached through the reflection on the construction process and on the implications the gathered data and evaluation. In the end, this process must inform insight(s) or theory(s) (KUECHLER; VAISHNAVI, 2012). These theories belong to the middle range theories because they do not aim at universal but at existential generalizations, and they do not make unrealistic idealizations in order to acquire knowledge, but aim to make only realistic assumptions about their object of study (WIERINGA, 2010). For example, the work of (KUECHLER; VAISHNAVI, 2012) defines two types of theories in DSR: design-relevant explanatory/predictive theory (DREPT) and Information Systems Directive Design Theory (ISDT). However, this discussion is beyond the scope of this thesis.

- Knowledge base

Another important component in the DSR is the knowledge base that consists of existing theories from science and engineering, specifications of currently known designs, useful facts about currently available products, lessons learned from the experience of researchers in earlier design science projects, and plain common sense (WIERINGA, 2010). The design science project uses this knowledge and may add to it by producing new designs or answering knowledge questions.

- DSR framework and research cycles (HEVNER; CHATTERJEE, 2010)

Altogether, a common framework is necessary for DSR in IS and a mental model or template for readers and reviewers to recognize and evaluate the results of such research. Without one, it may be difficult for researchers to evaluate it or even to distinguish it from practice activities, such as consulting (PEFFERS et al., 2007a). Figure 3.3 depicts the main constructs from the DSR and their relations, representing a DSR framework for information systems research.

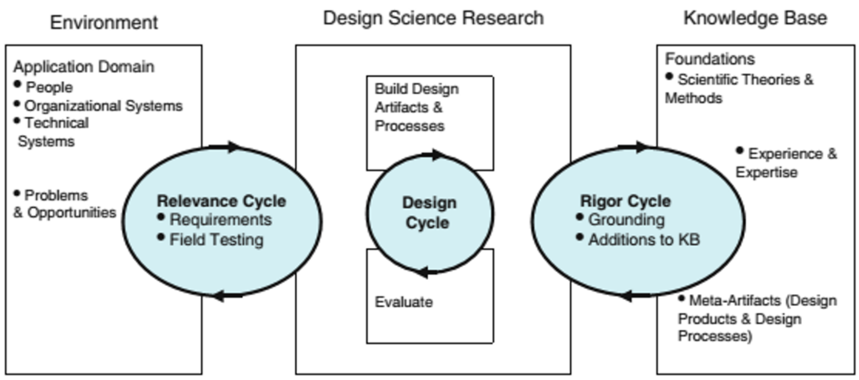

Figure 3.3: Information systems DSR framework (HEVNER; CHATTERJEE, 2010)

There are three design science research cycles in any design research project as shown in Fig. 3.3 (HEVNER; CHATTERJEE, 2010). The **relevance cycle** bridges the contextual environment of the research project with the design science activities. According to HEVNER; CHATTERJEE (2010), the relevance cycle initiates design science research with an application context that not only provides the requirements for the research (e.g., the opportunity/problem to be addressed) as inputs but also defines acceptance criteria for the ultimate evaluation of the research results. Yet, following HEVNER; CHATTERJEE (2010), the **rigor cycle** connects the design science activities with the knowledge base of scientific foundations, experience, and expertise that informs the research project. Finally, HEVNER; CHATTERJEE (2010) say that the **design cycle** is the heart of any design science research project and iterates between the core activities of building and evaluating the design artifacts and processes of the research. These cycles dialogue with the activities described in the next Section.

3.3.2.2 Operationalization of the method

There is a lack of a commonly accepted reference process model for design research: proposals include 'build – evaluate – theorize – justify' (MARCH; SMITH, 1995), 'identify a need – build – evaluate – learn and theorize' (ROSSI; SEIN, 2003), 'develop/build – justify/evaluate' (HEVNER et al., 2004), or 'problem identification and motivation – objectives of a solution – design and development – demonstration –evaluation –communication' (PEFFERS et al., 2007a). The process iteration from PEFFERS et al. (2007a) is adopted in this thesis and depicted in Figure 3.4. The method is explained in the following.

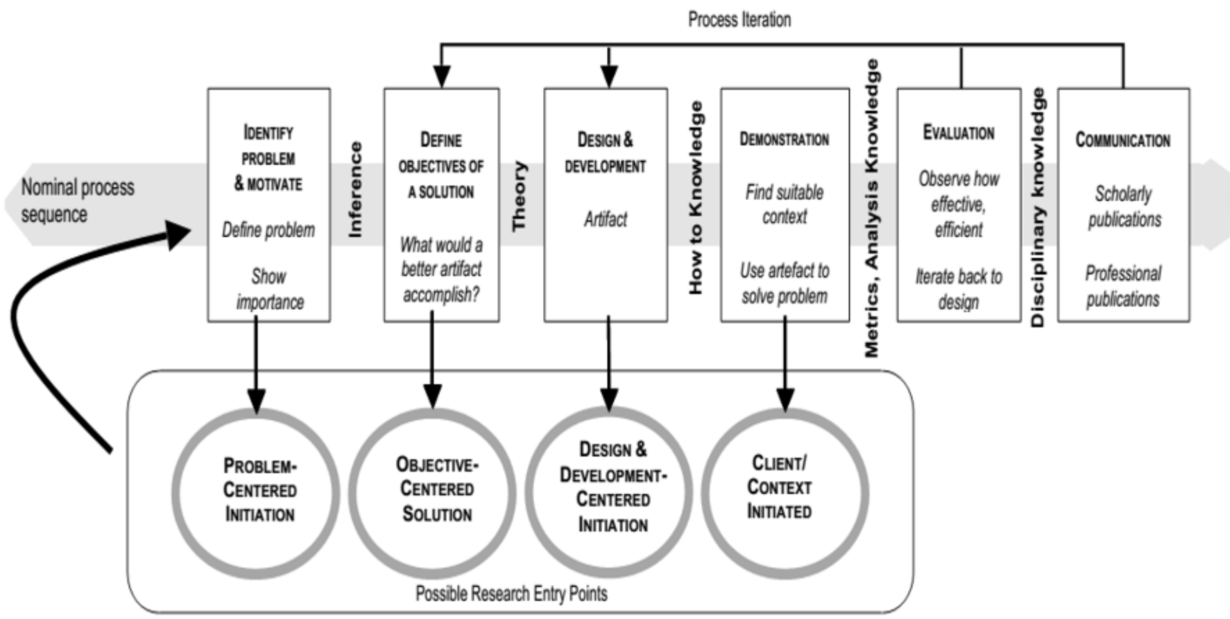

Figure 3.4: Design science research methodology (DSRM) process model (PEFFERS et al., 2007a)

- DSR activities

For PEFFERS et al. (2007a), the DSR process includes six steps: problem identification and motivation, definition of the objectives for a solution, design and development, demonstration, evaluation and communication. They are detailed as follows:

Activity 1: Problem identification and motivation. Define the specific research problem and justify the value of a solution. Since the problem definition will be used to develop an artifact that can effectively provide a solution, it may be useful to atomize the problem conceptually so that the solution can capture its complexity (HEVNER; CHATTERJEE, 2010). Resources required for this activity include knowledge of the state of the problem and the importance of its solution.

Activity 2: Define objectives of a solution. Infer the objectives of a solution from the problem definition and knowledge of what is possible and feasible. The objectives should be inferred rationally from the problem specification. Resources required for this include knowledge of the state of problems and current solutions, if any, and their efficacy (HEVNER; CHATTERJEE, 2010).

Activity 3: Design and development. Create the artifact. Such artifacts are potentially constructs, models, methods, or instantiations (HEVNER et al., 2004). This activity includes determining the artifact's desired functionality and its architecture and then creating the actual artifact. Resources required moving from objectives to design and development include knowledge of theory that can be brought to bear in a solution (HEVNER; CHATTERJEE, 2010).

Activity 4: Demonstration. Demonstrate the use of the artifact to solve one or more instances of the problem. This could involve its use in experimentation, simulation, case study, proof, or other appropriate activity. Resources required for the demonstration include effective knowledge of how to use the artifact to solve the problem (HEVNER; CHATTERJEE, 2010).

Activity 5: Evaluation. Observe and measure how well the artifact supports a solution to the problem. This activity involves comparing the objectives of a solution to actual observed results from use of the artifact in the demonstration (HEVNER; CHATTERJEE, 2010).

Activity 6: Communication. Communicate the problem and its importance, the artifact, its utility and novelty, the rigor of its design, and its effectiveness to researchers and other relevant audiences, such as practicing professionals, when appropriate (HEVNER; CHATTERJEE, 2010).

Although the activities are presented in a sequential flow, there is no exact entry point for the process. For instance, PEFFERS et al. (2007a) define four entry points for their DSRM also showed in Figure 3.4.

Problem-centered approach: is the basis of the nominal sequence. Researchers might proceed in this sequence if the idea for the research resulted from observation of the problem or suggested future research in a paper from a prior project.

Objective-centered solution: this could be triggered by an industry or research need that can be addressed by developing an artifact.

Design- and development-centered approach: It would result from the existence of an artifact that has not yet been formally introduced as a solution for the explicit problem domain in which it will be used. Such an artifact might have come from another research domain; it might have already been used to solve a different problem, or it might have appeared as an analogical idea.

Client-/context-initiated solution: It may be based on observing a practical solution that worked resulting in a DS solution if researchers work backward to apply rigor to the process retroactively.

3.3.2.3 Limitations of use

One first limitation that can be cited is the difficulty in generalizing results. However, this is an inherent choice since the focus of DSR is utility in a context rather than produce universal knowledge. As an applied research approach, the success of a DSR initiative may depend on researcher´s immersion in the real context, sometimes even physical presence in the industry is needed, which may require strong cooperation and trust between researcher and stakeholders. In our case, we have been working in a research partnership within an industry environment which was "research-friendly". The physical presence of the main researcher was required at least two times for each case analyzed. Even though, to get the expert´s availability to corroborate design choices and evaluate results was quite challenging.

3.3.2.4 Applying the method

The initial motivation to investigate the current research problem (or opportunity) came from the EA industry, through a current research partner. Our literature reviews indicated that the EANA field is relatively recent (around ten years of development) and confirmed that it still lacks some theoretical foundations. This was an opportunity to propose a conceptual meta-model for EANA, our first artifact (problem-centered research). Using the findings from our SLRs as a baseline, we extracted the common attributes of analysis initiatives found in the primary papers, composing our EANA library for the field (and answering the RQ1) and a conceptual meta-model (EANA-MM) with its respective key elements (answering the RQ2). Also, the same EANA-MM was designed to be used as a classification schema for NAIs, answering the RQ3.

Inspired in a related domain application, we designed the EA data derivation strategy to extend the EANA data possibilities. Furthermore, combining insights from the two SLRs, together with expert evaluation, it was possible to design a proposal for a process to perform EANA (RQ4), a research gap identified in our literature reviews. These four artifacts compose our EANA framework.

Finally, we designed two new methods to perform EANA analysis (cognitive-structural analysis and attribute check analysis). Our motivation was to apply theoretical knowledge to design these new artifacts as a means to empirically evaluate our EANA framework (objective centered solution).

As in any research approach, we need to ensure our adherence to the method (DSRM), transparency in the procedures and decisions and specify clearly the evaluation criteria, in order to enhance our conclusions' validity and artifact's utility. In this research, we adopted the constructs and the process for the DSRM proposed by (PEFFERS et al., 2007a) to frame our DSR initiatives. Table 3.2 describes the design of our six proposed artifacts along the phases of DSRM adopted. Evaluation criteria for the artifacts are discussed in section 3.4.

DSR activities	**EANA meta-model**	**EANA library**	**Data derivation strategy**	**EANA process**	**Two methods for EANA**
Identification of problem and motivation.	Problem-centered initiation based on our SLR.	Design- and development-centered approach to develop a GQM-based library for EANA.	Design and development centered approach to support EA experts generating new types of EA data based on a similar artifact from DSM field.	Problem-centered initiation based on the research gap pointed in our SLR.	Objective-centered solution triggered by an industry need to support EA experts combined their knowledge with network analysis metrics.
Goals definition	To define a conceptual meta-model for the EANA field.	To develop a GQM model together with a web repository to guide EA experts in choosing NAIs according to a selected EA concern.	To design an approach to derive EA network data based on data already available, providing new viewpoints for EA analysis.	To define a set of activities, inputs, outputs, actors to perform EANA.	To design two methods for support EA analysis combining expert knowledge with network analysis.

Continued on next page

Table 3.2: Distribution of activities of DSRM for each proposed artifact

Table 3.2 – continued from previous page

DSR activities	EANA meta-model	EANA library	Data derivation strategy	EANA process	Two methods for EANA
Design and develop-ment.	We use thematic analysis and SLR to deliver an EANA-MM as output.	We use thematic analysis and SLR to deliver a GQM-based EANA library.	We use concepts from network theory and DSM to deliver a matrix of EA derived viewpoints and a set of data derivation operators.	We based ourselves in processes initiatives from 3 papers found in the literature together with insights acquired during our empirical analysis to design the EANA process.	We use network theory and our insights to build the artifact.
Demons-tration.	We apply the framework classifying the related research literature. We also use the artifact during the design of our EA process.	We include the artifact in the design of our EA process.	We use the artifact to provide information to support EA analysis for three organizations.	We instantiate the artifact to guide the design of three other artifacts (EA data derivation strategy, cognitive-structural analysis and attribute check analysis).	We validate the artifact during two workshops with EA experts. We apply the method in two datasets of two different organizations (A and B).

Continued on next page

Table 3.2: Distribution of activities of DSRM for each proposed artifact

Table 3.2 – continued from previous page

DSR activities	EANA meta-model	EANA library	Data derivation strategy	EANA process	Two methods for EANA
Evaluation.	We evaluate the artifact with EA experts and instantia-tions.	We evaluate the artifact with EA experts.	We evaluate the artifact with EA experts and instantiations.	We evaluate the artifact by instantiations of it.	We evaluate the artifact with EA experts and instantiations.
Commu-nication.	To publish a paper with the artifact evaluation.	To publish a paper with the artifact evalua-tion.We also build a web portal to publicize the library.	Discussions with the Experts. Paper to be published.	Paper to be published.	Report sent to experts and practitioners involved. Paper to be published.

Table 3.2: Distribution of activities of DSRM for each proposed artifact

Last but not least, we need to make an important distinction about the use of DSRM and a common development of a software artifact, for example. According to HEVNER et al. (2004), to be considered as DSR, the research, must produce an artifact; the problem must be relevant; the design must be rigorously evaluated; there must be original research contributions; the methods for artifact construction and evaluation must be rigorous; there must be a search process to obtain the artifact; the research must be adequately communicated to both technical and managerial (decision-making) audiences.

3.3.3 Thematic analysis

3.3.3.1 Basics of the method

MAYRING (2004) defines thematic analysis as a selective analysis of subject matters or attributes of the text. Its aim is to identify themes and to analyze frequencies and contingencies of the content categories. Thematic analysis is an approach that is often used for identifying, analyzing, and reporting patterns (themes) within data in primary qualitative research (CRUZES; DYBA, 2011). A theme captures something important about the data in relation to the research question, and represents some level of patterned response or meaning within the data set (MAYRING, 2004).

Like most research methods, this process of data analysis can occur in two primary ways - inductively or deductively (BRAUN; CLARKE, 2006). In our case, the data coding follows an inductive approach (latent analysis), preventing researchers from erroneously ‘forcing’ a preconceived result (CRUZES; DYBA, 2011). Data are reviewed line by line from each paper in detail, and as a concept becomes apparent, a code is assigned. To ascertain whether a code is appropriately assigned, the analyst compares text segments to segments that have been previously assigned the same code and decides whether they reflect the same concept. Using this ‘constant comparison’ method, the researcher refines the dimensions of existing codes and identifies new codes (CRUZES; DYBA, 2011), over a series of stages: (1) an open coding stage, wherein data are chunked into smaller segments that are all given a descriptor, or code; (2) an axial coding stage (wherein codes are grouped into similar categories), (3) and a selective coding stage, wherein the theory is integrated and refined (ONWUEGBUZIE; LEECH; COLLINS, 2012).

3.3.3.2 Operationalization of the method

CRUZES; DYBA (2011) proposed a set of main steps and checklist for thematic synthesis described in Table 3.3:

Steps	Description	Checklist
Extract data	Extract data from the primary studies, including bibliographical information, aims, context, and results.	1. Have all papers been read carefully to be immersed with the data? 2. Have specific segments of text pertaining to the objectives of the synthesis been identified? 3. Have publication details, context descriptions, and findings been extracted from all papers? 4. Has another researcher checked the extraction?

Continued on next page

Table 3.3: Main steps and checklist for thematic synthesis proposed by CRUZES; DYBA (2011)

Table 3.3 – continued from previous page

Steps	Description	Checklist
Code data	Identify and code interesting concepts, categories, findings, and results in a systematic fashion across the entire data set.	5. Have important segments of text like concepts, categories, findings, and results been labeled and coded? 6. Has coding been done across the entire data set on a level that is appropriate for the research questions? 7. Has a list of initial codes with definitions and frequencies been created and checked by another researcher? 8. Have consistency checks or inter-rater reliability checks been performed to establish the credibility of the coding? 9. Are there clear, evident connections between the text and the codes?
Translate codes into themes	Translate codes into themes, subthemes, and higher order themes.	10. Have themes been created from a thorough, inclusive, and comprehensive review of the codes of all papers? 11. Has overlap between codes been reduced and the remaining codes been collated and translated into themes? 12. Have themes been checked against each other and back to the data of the original papers? 13. Are themes internally coherent, consistent, and distinctive?
Create a model of high order themes	Explore relationships between themes and create a model of higher-order themes.	14. Have themes been compared across studies, translated into each other, and interpreted into higher-order themes? 15. Have higher-order themes and relationships between themes been checked against the research questions of the synthesis? 16. Are there clear descriptions of the higher-order themes and the relationships between these themes? 17. Has a model been created to show the relationships between the higher-order themes?

Continued on next page

Table 3.3: Main steps and checklist for thematic synthesis proposed by CRUZES; DYBA (2011)

Table 3.3 – continued from previous page

Steps	Description	Checklist
Assess the trustworthiness of the synthesis	Assess the trustworthiness of the interpretations leading up to the thematic synthesis.	18. Have the assumptions about, and specific approach to, the thematic synthesis been clearly explicated? 19. Is there a good fit between what is claimed and what the evidence shows? 20. Are the language and concepts used in the synthesis consistent? 21. Are the research questions answered based on the evidence of the thematic synthesis?

Table 3.3: Main steps and checklist for thematic synthesis proposed by CRUZES; DYBA (2011)

3.3.3.3 Limitations of use

For BRAUN; CLARKE (2006), the flexibility of the method – which allows for a broad range of analytic options - means that the potential range of things that can be said about the data is broad. While this is an advantage, it can also be a disadvantage, since it can be potentially paralyzing to the researcher trying to decide what aspects of their data to focus on. Another issue to consider is that a thematic analysis has limited interpretative power beyond mere description if it is not used within an existing theoretical framework that anchors the analytic claims that are made. Many of the disadvantages depend more on poorly conducted analyses, which requires a strong method orientation focus; or inappropriate research question.

3.3.3.4 Applying the method

In this thesis, we follow the guidelines of CRUZES; DYBA (2011) to perform the thematic analysis. The method is used in conjunction with two other research methods, SLR and DSR. As one output of the applied thematic analysis, we categorize EA analysis approaches, models and analysis concerns according to the steps suggested by CRUZES; DYBA (2011) and described in Table 3.3. To increase the reliability of the process, the thematic analysis was carried out two times and tested as to whether the same findings result. Two researchers were involved in this method execution. In the end, the categories found were also reviewed by an external EA expert, to ensure their consistency. We also applied thematic analysis to evaluate the artifacts proposed with EA experts' opinion.

3.4 Data collection

Although we did not consider each organization as a case study, we collect EA data from three different organizations aiming to diversify the data sources and enhance our results. The first organization operates in the media industry and employs several thousand people at its headquarters in Germany and about twenty international sites; the second one is a large multi-industry player also headquartered in Germany and operating in a few other European countries. The third one is a German automotive company that has multibillion-dollar revenues and more than 40,000 employees around the world.

The EA management function of the first organization was established more than five years ago. The second organization's EA management function is about two years old. The third dataset comes from an organization without an officially established EA function, although, there is a high level of control about process documentation at the company. We refer to the resulting datasets from these three organizations as Dataset01, Dataset02, and Dataset03, respectively (due to anonymity reasons, we cannot be more specific and, e.g. provide network figures with names in this document). It is worthy to highlight that each organization had already a different set of documented EA data. In addition, the datasets provided for our research were selected by the experts, without any interference or participation from our side, which resulted in the three different dataset configurations described in Table 3.4.

Dataset01	**Available data**
Business Object	127
Business process	64
Application	201
Technology	183
Component´s atributes	Vendor, Availability, Type of development
Considered relations	BoxBp, AppxTec, AppxApp, AppxVendor
Dataset02	**Available data**
Application	228
Technology	169
Component´s attributes	Vendor, “TOP”, Business criticality.
Relations	AppxTec, AppxApp, VendorxVendor, TecxTec
Dataset03	**Available data**
Business unit	15
Business process	101
Business objects	247
Component´s attributes	Critical Bo; Critical Bp.
Relations	BuxBp, BuxBu, BpxBp

Table 3.4: Datasets description

Some components had their attributes considered in the analysis. For instance, "vendor" is an an attribute of a technology (e.g Microsoft is a vendor of MS SQLServer, a database technology considered as a component of the EA). Availability and type of (software) development were two attributes which belong to an application component of the Dataset01. For the Dataset02, there was also the "vendor" attribute and in addition, "TOP" (if the application is considered important in terms of strategy) and business criticality. Finally, for the Dataset03, the attributes "Critical Bo" (critical business object) and "Critical Bp" (critical business process) were also considered.

Two of the datasets were built manually from raw documents such PDFs and datasheets containing components and relations, complemented with experts ‘interviews. In one organization, the dataset was directly exported from an EA modeling tool. The data collection process was coordinated by the main researcher with at least, one EA expert from each company. After built the EA models, they were presented to the experts to verify inconsistencies.

3.5 Evaluation methods

Due to its complexity and the variety of application areas today, it is difficult to clearly delineate the term evaluation (CLEVEN; GUBLER; HÜNER, 2009). A potential definition is offered by SCRIVEN (1998) apud CLEVEN; GUBLER; HÜNER (2009):

> "Evaluation is the process of determining the worth, merit, or significance of entities; and evaluations are the outcome of that process. The evaluation may be external or internal, or a mix of these; and it may be quantitative or qualitative, or a mix of these. It is strongly although not always sharply distinct from explanation."

Foundational papers in the DSR literature stress the importance of evaluation (HEVNER et al., 2004). Nevertheless, despite the key role of artifact evaluation, this topic is under-explored in the IS DSR literature. Consequently, DSR researchers are often left to wonder what to evaluate (object and criteria of evaluation) and how to conduct the evaluation (PRAT; COMYN-WATTIAU; AKOKA, 2014). This very research gap is aimed by PRAT; COMYN-WATTIAU; AKOKA (2014), which provide a hierarchy of evaluation criteria for IS artifacts, a model providing a high-level abstraction of evaluation methods, and finally, a set of generic evaluation methods which are instantiations of this model.

In our thesis, we consider the works of PRAT; COMYN-WATTIAU; AKOKA (2014), CLEVEN; GUBLER; HÜNER (2009), as the theoretical ground for proposed artifacts´ evaluation.

3.5.1 Classification of the evaluation methods applied to the proposed artifacts

A taxonomy of methods evaluation is presented by CLEVEN; GUBLER; HÜNER (2009), which characterizes the evaluation approaches along the following twelve dimensions. Aiming to reduce the complexity of the classification of our methods, we selected only a subset of the dimensions from (CLEVEN; GUBLER; HÜNER, 2009), marked with a **“*”** in the left column of Table 3.5. Besides, we elaborate on that model using a second classification model proposed by PRAT; COMYN-WATTIAU; AKOKA (2014), adding the dimensions signed with

◊. We selected particular dimensions from both models, that in our view, could cover essential information about the evaluation methods and, at the same time, leave aside too much details about them.

Therefore, for each proposed artifact evaluation, we use the classification schema presented in Table 3.6.

- Approach(CLEVEN; GUBLER; HÜNER, 2009)

Qualitative. Characteristics of the evaluation object are not appraised on a numerical, but on a value basis, and "emphasizes the description and understanding of the situation behind the factors".

Quantitative. In the case of a quantitative evaluation characteristics of the evaluation object are, in contrast, assessed on a numerical basis.

- Artifact focus(CLEVEN; GUBLER; HÜNER, 2009)

Technical. Technical DSR artifacts are for example routing algorithms, communication protocols, image processing algorithms, hardware designs, or robots.

Organizational. Organizational DSR artifacts are for example process models and grammars, methods for organizational re- design, or accountability matrices.

Strategic. DSR artifacts on the strategic level are for example designs for decision support systems, roadmap development methods, or balanced scorecards.

- Artifact type (CLEVEN; GUBLER; HÜNER, 2009)

They can be an instantiation, a model, a design theory and so on. These concepts were explained in section 3.3.2.1.

Dimension	Value				
Approach*	Qualitative	Quantitative			
Artifact focus*	Technical	Organizational	Strategic		
Artifact type*	Construct	Model	Method	Instantiation	Theory
Object*	Artifact	Artifact construction			
Time*	Ex-ante	Ex post			
Method*◊	Survey-based	Analytical with logic reasoning			
Level of evaluation◊	Abstract	Instantiation			
Relativeness' of evaluation◊	Absolute	Relative to comparable artifacts	Relative to the absence of the artifact		
Secondary participant◊	Student	Researcher	Practitioneer		
Evaluation Criterion dimension◊	Goal	Environment	Structure		

Table 3.5: Classification schema for the evaluation methods of the artifacts

- Object (CLEVEN; GUBLER; HÜNER, 2009)

Artifact. A DSR artifact itself can be the object of evaluation.

Artifact construction. The process in which DSR artifacts are constructed may as well be evaluated (PEFFERS et al., 2007b).

- Time (CLEVEN; GUBLER; HÜNER, 2009)

Ex-ante. With ex-ante evaluation, candidate systems or technologies are evaluated before they are chosen and acquired or implemented.

Ex-post. With ex-post evaluation, a chosen system or technology is evaluated after it is acquired or implemented

- Method (CLEVEN; GUBLER; HÜNER, 2009; PRAT; COMYN-WATTIAU; AKOKA, 2014)

For the evaluation of IS artifacts different methods and technologies have been suggested such as formal proofs, action research, prototype, case study, survey, analysis and logical reasoning. In this thesis, we use **survey-based and analysis with logical reasoning as evaluation methods**. In Appendix B, we present the survey instrument for evaluate the EANA-MM and EANA library artifacts. The Evaluation itself of the artifacts is presented in chapter 9.

- Level of evaluation (PRAT; COMYN-WATTIAU; AKOKA, 2014)

The evaluation may be performed at two levels: the **abstract artifact** is either assessed directly or through one or several **instantiations**.

- Relativeness of evaluation(PRAT; COMYN-WATTIAU; AKOKA, 2014)

The evaluation of the artifact may be **absolute** (e.g., does the artifact achieve its goal?), **relative to comparable artifacts**, or **relative to the absence of artifact** (e.g. when no comparable artifacts can be found).

- Secondary participant(PRAT; COMYN-WATTIAU; AKOKA, 2014)

Secondary participants may take part in the evaluation of the IS artifact, e.g. by using a prototype and giving their feedback. They may be **students, practitioners, or researchers**.

- Evaluation criteria

Considering an IS artifact as a system, PRAT; COMYN-WATTIAU; AKOKA (2014) propose a model of evaluation criteria for IS artifacts organized according to the fundamental dimensions of a system(goal, environment, structure, activity, and evolution). In addition, the authors define a systematic organization of evaluation criteria, from which DSR researchers may choose some according to their own developed artifacts characteristics. The research of PRAT; COMYN-WATTIAU; AKOKA (2014) provides a holistic view of evaluation criteria and generic evaluation methods to assess them. The holistic view is achieved by applying general systems theory. We ground our evaluation process in their model composed of the five dimensions. Table 3.6 shows the selected criteria for our evaluation, according to the taxonomy of PRAT; COMYN-WATTIAU; AKOKA (2014):

Artifact´s dimension	Selected Criteria	Description
Goal	Efficacy/usefulness	It is the degree to which the artifact produces its desired effect (i.e. achieves its goal)
	Generality	Artifact generality is the goal generality (AIER; SCHÖNHERR, 2006): the broader the goal addressed by the artifact, the more general the artifact.
Environment	Utility	The criterion utility, which can be applied to people and/or to the organization, measures the quality of the artifact in practical use.
	Understandability	Understandability about its concepts and use (MARCH; SMITH, 1995),
Structure	Consistency	Consistency is “agreement or harmony of parts or features to one another or a whole”.
	Level of detail	It measures if the level of details of the conceptual elements is suitable to its goal.

Table 3.6: Selected criteria for the artifacts’ evaluation methods

3.6 The role of the researcher

As expected, I took part in the entire research process, starting from the explanation about the network analysis potential for the EA analysis context. Normally, an *in loco* slide presentation (about 40 minutes) , covering the main network metrics, the network levels of analysis and possible outputs, was done for experts of each analysis case. The goal was then to promote an conceptual alignment regarding network analysis and to think together about potential data sets that could be provided by the companies. I also took part in building the EA models when raw data was provided. During the artifacts´ design phase, I offered initial ideas of artifacts. The artifacts´potential results were initially discussed with the experts by emails or Skype conferences. Finally, after running all the analysis by myself, the empirical results were presented for the experts of each case, separately, in two rounds (in loco, first and by Skype, later). Usually each of those presentation sessions lasted about 50 minutes.

3.7 Chapter Summary

This chapter described our general research strategy, in terms of research questions, objectives, as well as the methods and procedures selected to answer each specific research question and finally, our proposed artifacts. In Table 3.8, it is described how each research phase was related with the employed research methods and targeted research questions. This sequence is not linear, and some activities were actually concurrently executed.

As one can see in Table 3.7, there is no specific phase for evaluation, since it is a cross-cutting activity performed for each artifact individually. The artifacts' evaluations are consolidated in chapter 9. In the next chapter, we start to present our research results.

Phase	Goal	Targeted research questions	Research method employed	Research outputs
SLR1	**Identify** which EA concerns, models and analysis approaches have been applied in EA research **with the purpose of** tracing an EA analysis state-of-art.	RQ1.1	Systematic Literature Review (mapping study); Thematic analysis.	State-of-art about the EA analysis. EA models, analysis concerns and approaches mapped. Reseach agenda is suggested.
SLR2	**Identify** which EA concerns, models and analysis approaches have been applied in EANA research **with the purpose of** tracing an EA analysis state-of-art.	RQ1.2; RQ2.1	SLR(conventional); Thematic analysis.	State-of-art about EANA. 74 NAIs are mapped.
DSR1	**Design and evaluate** an EANA MM **with the purpose of** establishing the fundamental concepts for the field.	RQ2.1; RQ3 DSRM;	Thematic analysis	EANA-MM
DSR2	**Design and evaluate** a GQM-based EANA library **with the purpose of** guiding expert in choosing network initiatives from a repository according to an EA concern.	RQ2.1; RQ4	DSRM; Thematic analysis.	GQM-based EANA library.
				Continued on next page

Table 3.7: Research phases, research questions, methods and outputs

Table 3.7 – continued from previous page

Phase	Goal	Targeted research questions	Research method employed	Research outputs
DSR3	**Design and evaluate** the EA data derivation strategy research **with the purpose of** generating additional EA viewpoints for analysis and evaluating the EANA process.	RQ2.2	DSRM; Network theory.	EA data derivation strategy.
DSR4	**Design and evaluate** the EANA process research **with the purpose of** guiding experts in performing EANA	RQ4	DSRM	EANA process.
DSR5-6	**Design and evaluate** the cognitive-structural analysis and attribute check analysis methods research **with the purpose of** combining structural and expert sources of knowledge in EANA and evaluating the EANA process.	RQ4	DSRM, Network theory; Thematic analysis.	Cognitive-structural analysis and attribute check analysis methods.

Table 3.7: Research phases, research questions, methods and outputs

4

EA ANALYSIS: WHAT DID WE ANALYZE SO FAR?

Although researchers have shown an increased interest in the EA analysis topic in the last two decades, there is not a shared and acknowledged comprehension about EA analysis. Additionally, the extension of EA analysis research, major accomplishments, difficulties, analysis techniques and the main concerns targeted are unknown. This chapter aims to answer the following research question: *What is known about enterprise architecture analysis?*

In order to achieve that answer, the following specific questions are proposed:

- What is considered EA analysis?
- What are the analysis approaches developed by researchers?
- What are the main EA concerns targeted?
- What are the main used models?
- What are the main difficulties in performing EA analysis?

In Section 4.1, the key information from our SLR research protocol (KITCHENHAM, 2007) are presented. Next, the answers to the previous questions are showed in Section 4.2. Finally, we make our final considerations in Section 4.3.

4.1 THE SLR RESEARCH PROTOCOL

In the sequence, we present the main components of our research protocol.

- Research query

From our research query, we derive the main terms for our query string described in Table 4.1:

Original terms	Synonyms
Enterprise Architecture	("Enterprise architecture" OR "business architecture" OR "process architecture" OR "information systems architecture" OR "IT architecture" OR "IT landscape" OR "information architecture" OR "data architecture" OR "application architecture" OR "application landscape" OR "integration architecture" OR "technology architecture" OR "infrastructure architecture")
Analysis, goals or concerns	Goals OR concerns OR methods OR procedures OR approaches OR analysis OR evaluate* OR assess* OR indicator OR method OR measur* OR metric

Table 4.1: Research query

In our final string we combined the terms related to EA and enterprise architectural subsets, derived from the work of SIMON; FISCHBACH (2013), and terms related to the analysis aspect, as goals, metrics and evaluation, as listed by (ANDERSEN; CARUGATI, 2014). Thus, our final string is described in Table 4.2:

Research query string adopted
("Enterprise architecture" OR "business architecture" OR "process architecture" OR "information systems architecture" OR "IT architecture" OR "IT landscape" OR "information architecture" OR "data architecture" OR "application architecture" OR "application landscape" OR "integration architecture" OR "technology architecture" OR "infrastructure architecture") AND (Goals OR concerns OR methods OR procedures OR approaches OR analysis OR evaluate* OR assess* OR indicator OR method OR measur* OR metric)

Table 4.2: Research query string adopted for the EA analysis SLR

- Engines

We selected the main engines/databases accessed in the information system community as our data sources for primary studies: Scopus, IEEE, ScienceDirect, ISI web of knowledge and AIS electronic library. Table 4.3 shows the results returned by each engine.

Engine	IEEE	ScienceDirect	Scopus	AISEL	ISI	TOTAL
Papers returned	1,762	832	3439	25	1,162	7,220

Table 4.3: Studies returned by each engine.

- Inclusion and exclusion criteria

Inclusion and exclusion criteria should be based on the research question. They should be piloted to ensure that they can be reliably interpreted and that they classify studies correctly (KITCHENHAM; CHARTERS, 2007). We describe our inclusion and exclusion criteria in Table 4.4 and Table 4.5, respectively.

Criteria	Description
CI-01	Papers containing techniques, methods or any kind of initiative to evaluate EA. E.g., Papers which uses EA as input for taking decision or Papers that analyze EA itself, its changes and evolution.

Table 4.4: Inclusion Criteria

Criteria	Description
CE-01	Papers in any language but English
CE-02	Repeated studies found in different sources (i.e. the same paper published in a journal and in conference proceedings) or reporting similar results. In this case, we adopt the most complete study found.
CE-03	Papers whose files are not recovered in full not available for the Brazilian universities. I those cases, we email the authors twice requesting their papers.
CE-04	Summary of keynotes, tutorials, white papers and incomplete papers (i.e. only abstract, work in progress reports), book chapters, dissertations
CE-05	Papers related to Product architecture analysis.
CE-06	Papers related to internal architecture of software.
CE-07	Paper containing only modeling approaches
CE-08	Papers about Virtual EA
CE-09	Papers do not approach EA analysis
CE-10	EA as a phenomenon. Papers which do not analyze EA itself but instead they relate the EA as an organizational function to another organizational variable like performance, communication etc.
CE-11	Literature reviews about EA (secondary studies).
CE-12	Meta-analysis. Papers dealing with discussion of analysis approaches but not performing any.

Table 4.5: Exclusion criteria.

- Screening phases

Considering the previous inclusion and exclusion criteria, our "conducting phase" (KITCHENHAM; CHARTERS, 2007) was divided into three subphases. Initially, we read 7220 abstracts and titles of primary studies returned by the engines. In the next steps, the readings focus on the introduction and conclusion sections. Finally, the remaining candidate papers were completely read. Table 4.6 describes the screening process and its respective outputs.

Phase	Papers outputted
0 - Papers returned from the engines	7220
1 - Title and abstract readings	803
2 - Introduction and conclusion readings	183
3 - Full paper readings	120

Table 4.6: Output of each research phase

The list of the 120 papers analyzed in this review can be found in Appendix 3.

- Data extraction

We extracted the information described in Table 4.7 from the final set of selected papers.

Field	Description
ID	An identification number for our internal control.
Document Title	Self-described.
Link	The link for the paper and its source on the web.
Year	The paper publication year.
Country	Country of each author.
Evaluation (0:Exclude 1:include)	Represents a decision flag for inclusion or exclusion of a paper.
Concern	The analysis concerned target by the paper. E.g. risk, cost, EA alignment etc.
Layer	The targeted layers: value, business, application, information, technology
Approach: M-Model or F-Functional	The kind of input for the analysis process. A model-based analysis has necessarily a model in any language for the EA under analysis. Functional analysis is made with indicators about the EA but without using any model for that.

Continued on next page

Table 4.7: Data fields extracted from the EANA primary studies

Table 4.7 – continued from previous page

Field	Description
Model	Type of modeling language used in the analysis. E.g. archimate based, UML, own designed etc.
Model Data source and Data Collection	It is related to the type of data collection methods employed to build the EA model. For example, some papers use automated data collection methods. Other authors use interviews with the expert team to map EA components and their relations.
Techniques/Methods	The analysis technique or method employed in the EA analysis.
Type of analysis: Theoretical, Simulated or Empirical	If the analysis technique uses empirical data, illustrative examples (toy example) or of the techniques is just theoretically explained without any application.
Qualitative or Quantitative	The nature of the analysis approach executed.
Use of tools	If tools are used or proposed in the study.

Table 4.7: Data fields extracted from the EANA primary studies

We categorize the extracted data in EA analysis concerns, models and approaches, using thematic analysis methods as presented in chapter 3. We also discuss questions such the nature of the analysis, types of datasources and the use of tools in the next Sections.

4.2 Results of the SLR

The query string was executed on the first week of December, 2015 for the IEEE, ScienceDirect and Scopus engines; and on the first week of January, 2016 for the AISEL and ISI web of knowledge.

4.2.1 Meta-data

As depicted in Figure 4.1, the trend of growth of EA publications presented by SIMON; FISCHBACH; SCHODER (2013) is also reflected in EA analysis publications, since they become more frequent after 2006. Since our document collection started in December 2015, the results do not include 2016 publications.

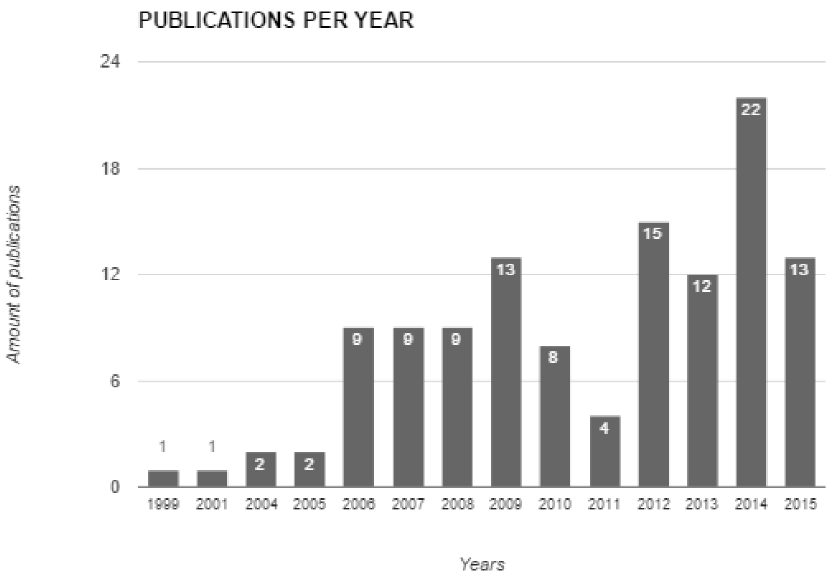

Figure 4.1: Distribution of EA analysis research papers per year

In this work, we compute the nationalities of the papers according to the authors' nationalities without duplicates (e.g., two Germans and one Brazilian author in the paper represent only one German and one Brazilian authorship in our calculation). Figure 4.2 depicts the country distribution for the field.

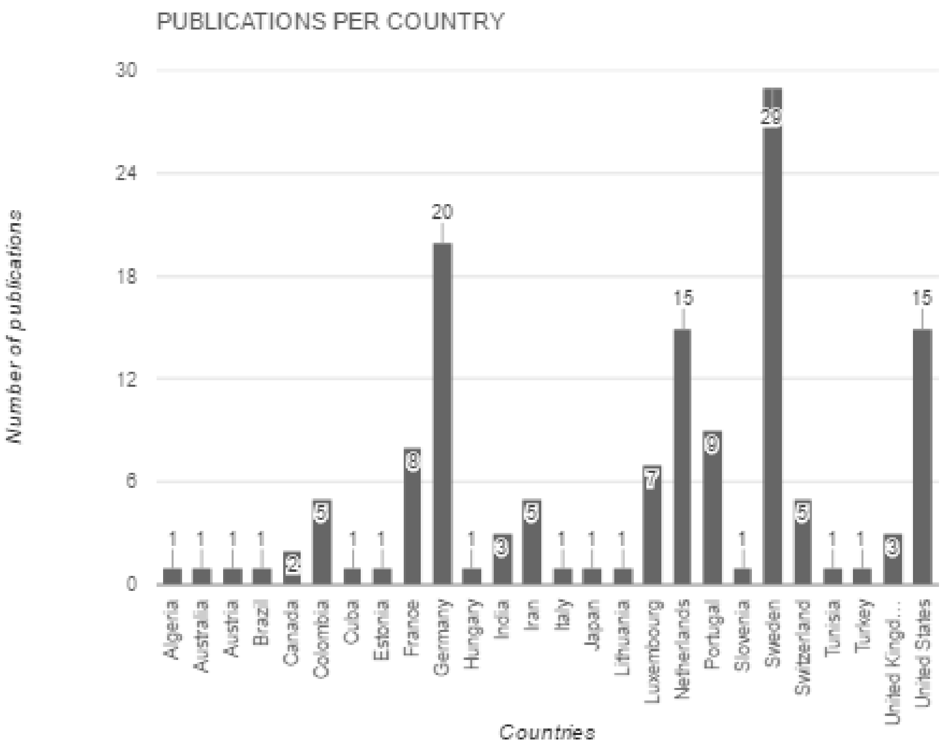

Figure 4.2: Distribution of EA analysis research papers per country

Considering the country distribution, Sweden and Germany have the highest number of publications, corresponding to more than 25% of all publications. We can point some reasons for this fact. In the papers selected, Swedish researchers (especially from the Royal Institute of Technology, KTH) demonstrated a strong interaction with enterprises of the electrical sector, as case studies.

The participation of Germany reflects a considerable role of universities researching EA, like the universities of Munich, Berlin and Stuttgart. The existence of consulting firms in the field is another factor that corroborates the diffusion of the concept in the country.

The United States has a key role in the research, with contributions especially from Harvard Business School, University of Hawaii and OHIO State University. The Netherlands also present an important industry-research interaction. For instance, Archimate and Archimate extensions were conceived by Dutch researchers.

4.2.2 EA concerns mapped

Our definition of concern agrees with the Oxford Dictionary of English definition, which is “A matter of interest or importance to someone”. A concern is the main analysis objective of an approach. Analyzing the primary studies, we identified 47 concerns in total, classified into 15 categories. Table 4.8 presents the definitions of each category of concerns.

ID	Category of analysis concern	Description
1	Actors aspects	This category covers papers dealing with impact of EA changes in human aspects such actor’s performance, actor’s relationship with resources and activities
2	Application portfolio analysis	This category covers research related with portfolio analysis of application architecture.
3	EA standards conformance	In this category, we classify the research, which aimed to establish EA patterns and analyze real world EAs comparing them with those EA patterns or standards (e.g. conformance with EA pattern catalog from BUCKL et al. (2009)).
4	Cost analysis	It covers cost analysis regarding different aspects of EA and EA as a whole (return of investment of the EA). It is usually related to the IT architecture, as most of the specific concerns comprise costs related with applications/components.

Continued on next page

Table 4.8: EA analysis concerns´s categorization

Table 4.8 – continued from previous page

ID	Analysis concern	Description
5	EA Alignment	Ideally, it should have a top-down alignment from strategy business process, which realize it, which in turn, are supported by applications and finally, are executed with a suitable infrastructure. Papers in this category are concerned in verifying issues like alignment or coverage or support between components of adjacent layers.
6	EA change	This category covers concerns related to modifications in the current state of the EA – e.g. its consequences of a particular change, scenarios' choices and what are the necessary steps to implement a specific change (gap analysis).
7	EA decisions	This category covers analysis approaches related especially with the decision-making process itself. The concerns here cover the rationale behind decisions, analysis of the stakeholders' influence on the decision process, types of methods to evaluate alternatives and so on.
8	EA governance	It considers EA as an whole organizational function, to analyze its overall quality. Papers in this category covers concerns like EA effectiveness, EA data quality, and EA documentation
9	Information dependence of an application	Papers in this category analyze the dependency of a given application A, to one or more business objects form the Information layer. In other words, this category is related to the in-and-out information flow analysis among applications.
10	Model Consistency	This category aims to evaluate the integrity of EA models and its consistency through time and organization' evolution.
11	Performance	It is mostly concerned with specific measures of performance, e.g. business component ´s performance, measures of system quality.

Continued on next page

Table 4.8: EA analysis concerns´s categorization

Table 4.8 – continued from previous page

ID	Analysis concern	Description
12	Risk	This category covers different risks aspects: risk of component´s failure (mainly in the application layer) and its consequences. information security aspects of the architecture as a whole; EA project risks (e.g. risks present in an integration initiative); EA implementation risk
13	Strategy Compliance	It analyses if EA decisions, EA projects, models and the structure itself (business, applications, infrastructure and so on) are compliant with the organization' strategy established (business goals, principles, directives, capabilities).
14	Structural aspects	It analyses how components are organized, the relations among them with their emergent complexity, possible ripple effects caused by changes, clustering issues, positional value in the structure and related concerns.
15	Traceability	It represents the need of querying or tracking components that are connected/linked to a particular component and/or have specific attributes values.

Table 4.8: EA analysis concerns´s categorization

In Table 4.9, we detail each category presented previously with their respective elements (concerns).

Concern category	Concern	ID Concern	Description
Actors aspects	Human aspects in EA	1	This category covers papers dealing with Human and relations with business process, goals etc.

Continued on next page

Table 4.9: Detailing the EA analysis concerns

Table 4.9 – continued from previous page

Concern category	Concern	ID Concern	Description
	EA actor's competence impact	2	This concern tries to investigate the impact of EA changes in the performance of actors of the EA. The definition describes the ability to combine in an efficient manner a number of non-material resources (knowledge know-how and social attitude) and material resources (instruments, machines, etc.) in order to respond to the need of an activity.
Application Portfolio Analysis	Application Portfolio Analysis	3	This category covers research related with portfolio analysis of application architecture.
EA standards conformance	EA standards conformance	4	In this category, we classify the research, which aimed to establish EA patterns and analyze real world EAs comparing them with the EA patterns or standards.
Cost analysis	Cost allocation	5	It is related to the analysis of the cost for each business unit considering the financial cost of EA components.
	EA ROI	6	This concern is related to the analysis of EA cost as a whole, such as return investment techniques.
	IT architecture cost	7	It approaches techniques to estimate or assess the cost of the current IT architecture or proposed scenarios.

Continued on next page

Table 4.9: Detailing the EA analysis concerns

Table 4.9 – continued from previous page

Concern category	Concern	ID Concern	Description
	IT portfolio analysis	8	An important element of portfolio management therefore is the valuation of IT projects and assets in terms of their costs, benefits, risks and contribution to strategic objectives.
	Cost of change of components	9	"An essential issue with today's software systems is that many of them are interconnected, thus a modification to one system may cause a ripple effect among other systems". Application costs are affected significantly by the number of the application's interdependencies (the higher the interdependencies, the higher its operations and maintenance costs).
EA alignment	Alignment between layers	10	It is related to strategies to identify, keep/promote alignment between two or more EA layers.
	EA redundancy	11	"Unintended or unidentified redundancies can affect the data quality, duplicating efforts which may lead to inconsistencies". "inside the information domain of an EA, analysis of actual information assets should be conducted to detect and / or eliminate unplanned redundancies [4]."

Continued on next page

Table 4.9: Detailing the EA analysis concerns

Table 4.9 – continued from previous page

Concern category	Concern	ID Concern	Description
	Support coverage	12	"...to show which IT systems support operations of a company, a three-dimensional map could be imagined which captures the mutual dependence of business functions and business products of that company on application components [10]".
	Application Redundancy	13	In this category, redundancy of EA components, such applications, is the main concern for this category.
	Overall alignment	14	This category of papers use EA models to measure the overall IT-Business alignment. This alignment differs from EA alignment category, since it does not look at the alignment between pairs of EA layers, but at alignment of the EA as a whole
EA change	EA change impact	15	It is related to the analysis of the effect/impact provoked by changes in the architecture. It also covers the suitability of the architecture considering an EA change candidate. "The goal of a change impact analysis is to see what would happen if a change occurs, before the change really takes place".
	EA gap analysis	16	This concern considers methods and techniques to analyze the gap between one AS-IS and other TO-BE states.

Continued on next page

Table 4.9: Detailing the EA analysis concerns

Table 4.9 – continued from previous page

Concern category	Concern	ID Concern	Description
	Scenario analysis	17	It considers two or more EA TO-BE states and analyze them according to some criteria. Then, it focus in a technique or method to decide which scenario would be the best alternative. The term "EA scenario" is used to denote an architecture, an architecture proposal, or a solution for an enterprise, which can be on any level of granularity. However, of course, the architecture scenario candidates to be compared should have the same level of granularity.
EA Decisions	Decision rationales tracking	18	This concern is about to map the relations between EA decisions taken. "EA Decision represents a decision that has been made or rejected in order to resolve an issue. An EA decision shows decisions that are captured in the context of an Enterprise Transformation".
	Decisions in EA projects´portfolio	19	This category brings the concept of arbitration, which consists in making a decision of choice between a collection of alternatives (e.g. requirements, projects or evolution scenarios) by ordering them based on one or several decision criteria.

Continued on next page

Table 4.9: Detailing the EA analysis concerns

Table 4.9 – continued from previous page

Concern category	Concern	ID Concern	Description
	Prioritization of EA project initiatives	20	It is related to the decision and prioritization of projects that change the EA, adding, removing or updating it. “We use the term "transformation" to describe substantial, business-critical changes in an application landscape that have significant impact on an organization’s business processes and on the people that work with the applications.” (ID2_580). It does not consider EA scenarios.
	Stakeholder analysis	21	It is related to the role, importance or influence of the stakeholders for EA domains, process, applications or other components, regarding EA decisions.
EA Governance	EA quality	22	By an EA quality attribute we mean a non-functional property that has meaning in all aspects of an enterprise (such as maintainability in application layer, business, technology and so on).
	EA data quality	23	This concern is related to EA overall analysis, considering the aspects of data quality. "Data quality is a multi-faceted concept. The most common dimensions of data quality are completeness, consistency, currency, relevance and accuracy (Redman, 1996)".

Continued on next page

Table 4.9: Detailing the EA analysis concerns

Table 4.9 – continued from previous page

Concern category	Concern	ID Concern	Description
	EA documentation	24	As for architectural documentation, it is common practice to partition and refine the overall EA into more detailed domain architectures that cover specific segments (e.g., marketing and sales) of the enterprise and allow the distribution of architectural work in large enterprises.
	EA effectiveness	25	The concern here is related to the ability to proper implement and use EA function. "This involves assessing the organization on its ability to reach the goals it strives for with enterprise architecture" (ID_105, p.1)
	Metrics monitoring	26	These concerns relates to the tracking and monitoring of EA metrics.
Information dependence of an application	Information dependence of an application	27	Papers in this category analyze the dependency of a given application A, to one or more business objects form the Information layer. In other words, this category is related to the in-and-out information flow analysis among applications.
Model consistency	Model checking during EA evolution	28	This concern is related to check if any previous defined constraint or rule is being obeyed for any TO-BE state.

Continued on next page

Table 4.9: Detailing the EA analysis concerns

Table 4.9 – continued from previous page

Concern category	Concern	ID Concern	Description
	EA model consistency	29	This concern is related to application of formal methods (e.g. ontologies) to verify the correctness of the EA modelled according to the formal rules.
Performance	Application reliability	30	Papers concerned with the ability of an application to offer correct and timely service.
	Business performance	31	This concern covers research that analyzed how other EA components (such business goals, applications, technologies and so on) can impact the business process performance.
	EA component performance	32	This type of analysis includes performance metrics, i.e. response time, utilization, workload.
	System quality	33	It covers papers that evaluate system quality attributes, including functionality, reliability, usability, efficiency maintainability and portability. "ISO 9126-1 defines software quality in terms of six quality characteristics. These are functionality, reliability, usability, efficiency maintainability and portability".
Risk	Risk analysis	34	In the scenario analysis context, risks can come from potential negative impacts of the planned change of an EA component on others. Therefore, modifications of positionally important applications could expose the organization to significant risks.

Continued on next page

Table 4.9: Detailing the EA analysis concerns

Table 4.9 – continued from previous page

Concern category	Concern	ID Concern	Description
	Risk failure	35	It contains papers concerning risk of failure in EA components with its impact.
	Information security risk	36	“Usually related to confidentiality, integrity and availability of EA components” or cyber security analysis".
	Availability	37	Availability analysis of EA components. By availability we mean the property of EA component being part of an industrial process (e.g. paper manufacturing) or a business process (e.g. trading on the stock market) where its downtime poses significant risks to human life, large costs or other consequences for business.
	EA implementation risk	38	This category is about the risks that arises in a EA implementation process
	Integration risk	39	"Due to the fact that applications depend on other applications, it is often difficult to modify or replace them, because unexpected side effects can occur which may result in failure of business critical applications".
Strategy Compliance	Compliance of EA decisions with requirements, principles, directives and capabilities	40	It is related with the compliance of EA decisions with EA goals, requirements, principles, directives and capabilities. The idea is to keep track on the historic of EA decisions and the rationale behind them.

Continued on next page

Table 4.9: Detailing the EA analysis concerns

Table 4.9 – continued from previous page

Concern category	Concern	ID Concern	Description
	Goal compliance	41	Concerns related to how EA components contributes to the achievement of EA goals. “It provides traceability between high-level business objectives and low-level enterprise architecture elements..”.
	Compliance of EA principles, by EA models	42	This category is about the alignment of EA models with EA design principles defined by the organization. "Understanding, an EA principle focus on how the design of an enterprise meet its essential requirements. They are declarative statements that can be made more precise using design instructions (by modelling and formalizing the design principles)”
Structural aspects	EA domain analysis	43	It checks the alignment of EA domains already implemented with the domains designed by the experts through functional decomposition.
	EA complexity	44	This category tries to analyze complexity related to the number of components and heterogeneity of components.
	Positional value	45	This category analyze the effect of occupying structural positions for individual EA components; structural group patterns and network modeling approaches for EA analysis.

Continued on next page

Table 4.9: Detailing the EA analysis concerns

Table 4.9 – continued from previous page

Concern category	Concern	ID Concern	Description
	Impact of Shadow IT systems	46	This concern conjugates structural analysis performed with network analysis together with information about IT shadow systems.
Traceability	Traceability	47	It represents the need of querying or tracking components that are connected/linked to a particular component and/or have specific attributes values.

Table 4.9: Detailing the EA analysis concerns

The classification of each work and the evidence which support this classification are available in Appendix D. Since an approach may suit more than one concern at a time, several papers are classified with more than one concern (e.g. SIMON; FISCHBACH; SCHODER (2013)) and VASCONCELOS et al. (2004)).

According to our research' results, the focus of EA analysis has been in five main categories: EA change, EA alignment, strategy compliance, performance and structural aspects, as shown in Figure 4.3.

Papers covering these concerns correspond to 77% of the whole final set. EA change and EA alignment represent the biggest research branch (43 studies from 120) and use a plethora of different analysis techniques. In contrast, human aspects are almost not considered in EA analysis so far - only two papers covered this concern. Considering EA as a complex sociotechnical system, this result is quite surprising.

4.2.3 EA models mapped

In our final set of primary studies, 88% of all analysis approaches are model-based. We mapped the models used by each analysis technique and classified them into eight categories, as shown in Table 4.10.

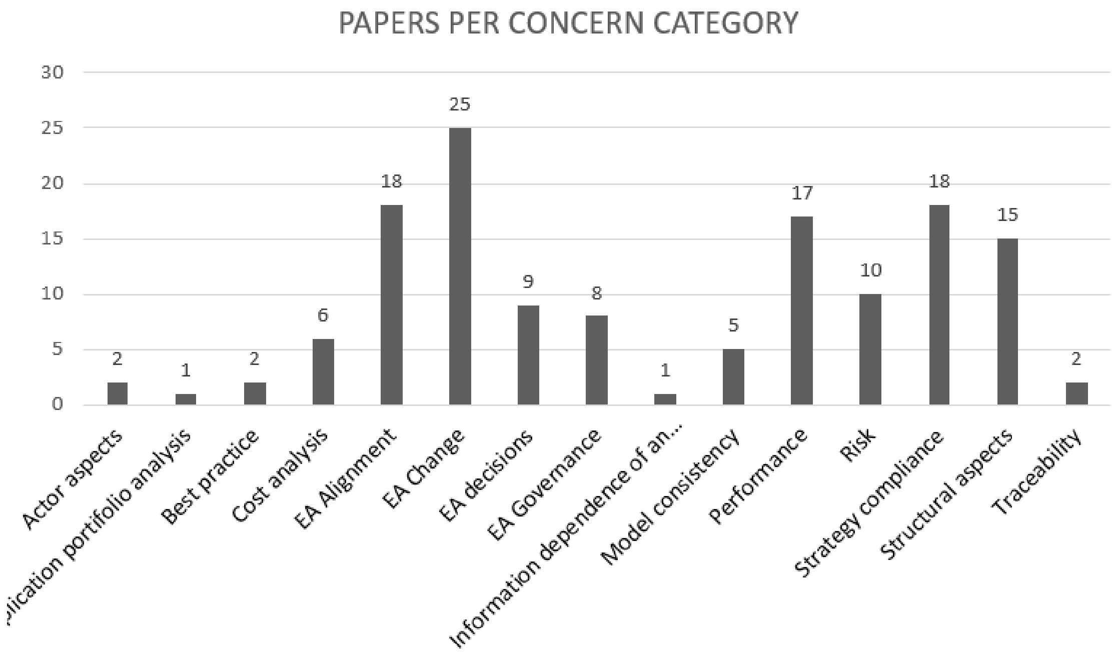

Figure 4.3: EA analysis research papers per concern category

ID	Category	Definition
1	Archimate-based	In this category, we classified all the papers that modeled their EA using the standard modeling from TOGAF. We have yet two subcategories: archimate-based (papers, which used the Archimate adapted or merged with other entities and attributes) and Archimate extensions (papers that used extensions for particular concerns such motivation, security issues and so on). Several concerns were targeted with these models. Example of primary study is BOER et al. (2005).
2	Combined models	This category comprises papers that use more than one model to perform their analysis. For example, SUNKLE et al. (2014a) uses Business Motivation Model (BMM) and Intentional Modelling together with Archimate to evaluate if and how business rules and goals are compliant with the organization's directives. Alternatively, a paper using only the BMM would fit in the specific category for intentional modeling related papers. Example of primary study is SUNKLE et al. (2014b).
		Continued on next page

Table 4.10: EA analysis models categories

Table 4.10 – continued from previous page

ID	Category	Definition
3	Formal and semi-formal specification based	This category is characterized by the attempt to describe EA models with the textual languages or mathematical specification such as set theory, ontology, XML and domain specific languages. Normally, the related papers build their models aiming to take advantage of reasoning techniques to support EA analysis. Example of primary study is CASTELLANOS; CORREAL; MURCIA (2011).
4	Graphs based	In this category, the EAs are modeled as graphs, with their components and relations being represented by nodes and edges, respectively. In addition, design structure matrix are included since they are structurally equivalent to graphs. Example of primary study is DREYFUS; IYER (2008).
5	Intentional Modeling	This category covers papers concerned with goals, modeled with I* framework and related models. Usually, these papers aim to analyze strategy related concerns. Example of primary study is YU; STROHMAIER; DENG (2006).
6	Own	In this category are included papers that present their own EA modeling framework and it does not fit in any of other specific categories. For example, ZACARIAS et al. (2007) presents a framework called CEO EA framework (extended) to evaluates Human aspects in EA.
7	Probabilistic networks based	It includes papers that used Probabilistic Relational Models, Extended Influence Diagrams, Influence Diagrams, Bayesian Networks and Fault Tree Analysis models. All those models work with uncertainty and probability principles in their modeling approaches. Performance metrics, from the infrastructure to the business process level are common concerns. Example of primary study is JOHNSON; NORDSTRÖM; LAGERSTRÖM (2007).

Continued on next page

Table 4.10: EA analysis models categories

Table 4.10 – continued from previous page

ID	Category	Definition
8	UML-based	This category covers papers that use the Unified Modelling Language (UML) or UML-based model to perform their analysis. UML is present also with combined models. Example of primary study is MORKEVIIUS; GUDAS; ILINGAS (2010).

Table 4.10: EA analysis models categories

In Table 4.11, we detail each of those previous categories of models, totalizing 19 types of models identified.

Model-category	ID	Modeling approach	Brief Description
Archimate-based	1	Archimate	Well defined in theory
	2	Own (Archimate-extended)	Models here present Archimate with a specific extension to perform the intended analysis. E.g., Manzur et al. (2015) presented the xArchimate, which consists in the Archimate with an extension to dynamic attributes that allow perform simulations and evaluate components performance.
	3	Own (Archimate-based)	The archimate-based category covers papers that explicitly used the Archimate adapted or merged with other entities and attributes. E.g., Plataniotis et al. (2015) uses Archimate' structure as a start to construct their EA Anamnesis metamodel.
	4	Bayesian Belief Networks	Well defined in theory

Continued on next page

Table 4.11: EA Analysis model categorization (* These are recognized models in EA field and also well defined in theory. Thus we did not include their descriptions.)

Table 4.11 – continued from previous page

Model-category	ID Approach	Modeling approach	Brief Description
Probabilistic networks based	5	EID	Well defined in theory
	6	PRM	Well defined in theory
	7	PRM based	Well defined in theory
	8	P2AMF	"The Predictive, Probabilistic Architecture Modeling Framework (P2AMF) is a probabilistic Object Constraint Language. The main feature of P2AMF is its ability to express uncertainties of objects, relations and attributes in UML-models and perform probabilistic analysis incorporating these uncertainties" ÖSTERLIND; LAGERSTRÖM; ROSELL (2012)
Combined models	9	Combined models	They use a combination of two or more model categories
Graphs based	10	DSM-based model	Design structure matrix enables the expert to model, visualize, and analyze the dependencies among the entities of any system and derive suggestions for the improvement or synthesis of a system.
	11	Graphs	Well defined in theory
Intentional modeling	12	GRL	Well defined in theory
	13	Intentional Modeling	Well defined in theory
	14	User Requirements Notation (URN)	Well defined in theory
Formal specification based	15	Description language based	Well defined in theory
	16	Ontology	Well defined in theory

Continued on next page

Table 4.11: EA Analysis model categorization (* These are recognized models in EA field and also well defined in theory. Thus we did not include their descriptions.)

Table 4.11 – continued from previous page

Model-category	ID Approach	Modeling approach	Brief Description
	17	Set theory	Well defined in theory
Own	18	Own	Analysits define a particular semantic for components and relationships
UML-based	19	UML-based	Well defined in theory*

Table 4.11: EA Analysis model categorization (* Well defined in theory models are well recognized in the field of EA and are well defined in theory, thus we did not include their descriptions.)

Even though there is a trend to using Archimate-based and graph-based models –the two categories correspond to 35,5%. 38% of the model based approaches used a self-developed model or a combined models to perform their analysis, as shown in Figure 4.4. This plurality of different modeling approaches corroborates the affirmation from JOHNSON; NORDSTROM; LAGERSTROM (2007) that "there is no clear understanding of what information a good enterprise architectural model should contain". It is natural the existence of several types of models, as organizations have different analysis concerns and goals for them. However, one downside of this is that, so far, the use of several modeling languages makes hard the comparison among research results of primary studies. Despite that, it is important to highlight the importance of Archimate models (representing 23% of the model choices) in the research as a standardization effort developed by the industry.

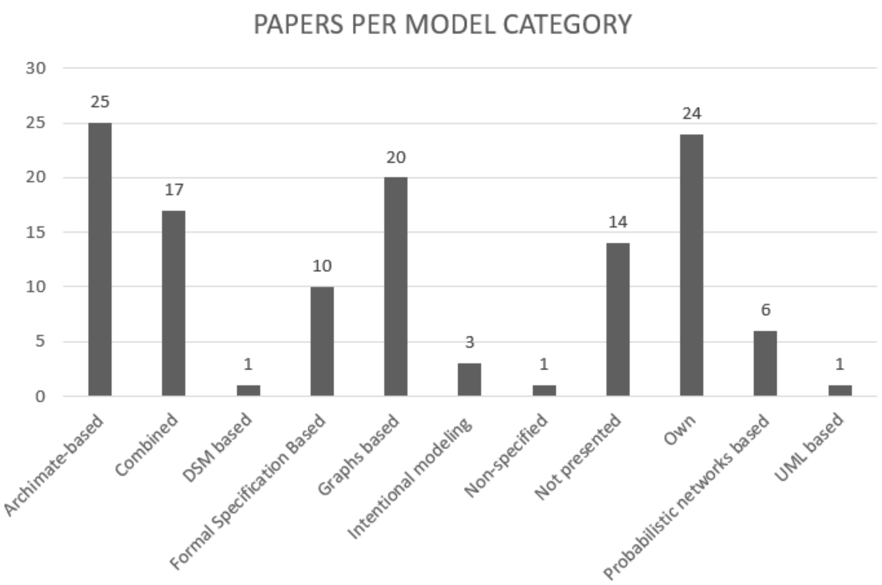

Figure 4.4: EA analysis research papers per model category

We list in Table 4.12 the six most applied categories of models and their analysis concerns, to compare the graph-based analysis with the other ones:

Category of model	Amount of papers	Analysis concerns targeted
Archimate-based	25	In this category, we classified all the papers that modeled their EA using the standard modeling TOGAF, archimate-based and archimate-adapted or merged with other entities and attributes, archimate extensions. From a total of 15 categories of concerns, Archimate related models were used to analyze all of them, especially for EA change(6 articles) and strategy compliance (5 articles). Although, no occurrences for structural concerns.
Own	24	Includes papers that present their own EA modeling frameworks and do not fit in any other specific categories. For example, ZACARIAS et al. (2007) present a framework (extended) to evaluate human aspects in EA. No tendency for specific concerns was identified.
Graph-based	20	In this category, the EAs are modeled as graphs, with their components and relations being represented by nodes and edges. This is also the approach we take in this thesis. Design structure matrices are also included. In 8 papers, graphs were used to analyze structural concerns, 3 for strategy compliance, 3 for risks, 2 for performance, 5 for EA decisions, 2 for EA alignment and 2 for cost analysis.

Continued on next page

Table 4.12: Most used EA models according to our EA analysis literature review

Table 4.12 – continued from previous page

Category of model	Amount of papers	Analysis concerns targeted
Combined models	17	This category comprises papers that use more than one model to perform their analysis. For example, SUNKLE et al. (2014a) use Business Motivation Model (BMM) and intentional modeling together with Archimate to evaluate if and how business rules and goals are compliant with the organization's directives. Models in this category together target several different concerns from cost analysis to strategy compliance without any predominant concern.
Formal and semi-formal specification based	10	This category is characterized by the attempt to describe EA models with the textual languages or mathematical specification such as set theory, ontology, XML and domain specific languages. Normally, the related papers build their models aiming to take advantage of reasoning techniques to support EA analysis.
Probabilistic network based	6	Includes papers that used probabilistic relational models, extended influence diagrams, influence diagrams, Bayesian networks and fault tree analysis models. The main concerns analyzed are related to EA performance metrics, from the infrastructure to the business process level.

Table 4.12: Most used EA models according to our EA analysis literature review

According to the Table 4.12, graph-based models represent the third most used category of models in EA analysis. This might indicate the familiarity of analysts with this type of modeling choice. Graphs are used not only to perform pure structural analysis but also in seven other concerns categories, highlighting the versatility of graphs for several analysis concerns. It is worth pointing out that analyzing an individual component while considering the overall structure is an exclusive capability of the structural paradigm. The challenge becomes, thus, to interpret the network metrics in the EA analysis context to identify its value. For instance, one may ask what does it mean for a component having a high betweenness value inside the application layer or what kinds of structural patterns can emerge in each layer or inter-layers and their meanings. Finally, when compared to other analysis paradigms such as Archimate-based, formal and semi-formal specification-based, efforts to collect information to be used by their methods have a lower level of complexity due to the relatively simple nature of the data (only the relations and components) comparable to domain specific models as depicted in Figure 4.5.

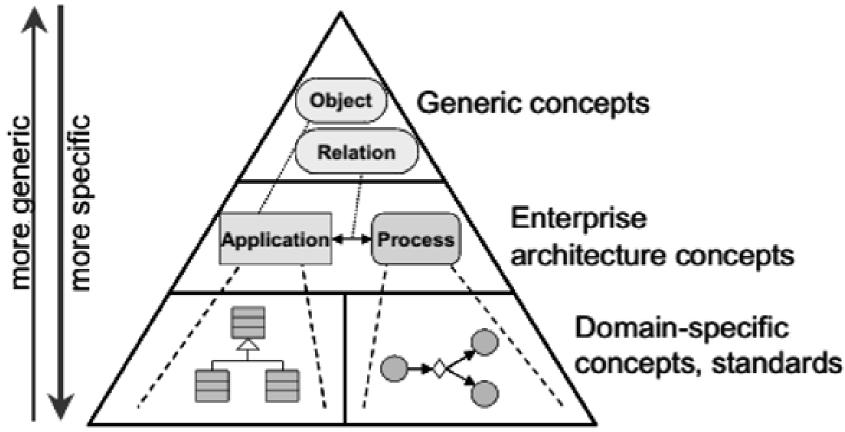

Figure 4.5: Levels of models' abstraction and EA contextualization among them (STEEN et al., 2004).

4.2.4 EA analysis approaches mapped

Analysis approaches cover techniques and methods used to perform EA Analysis. We identified a plurality of different approaches, classified in 23 categories according to their main characteristics (see Table 4.13). A large portion of the approaches was proprietary and many were poorly detailed, focusing on the results rather than the analysis process.

ID	Analysis approach	Description
1	Agent simulation based	Well defined in theory.*
2	Analytic Hierarchy Process (AHP)	Well defined in theory.
3	Architecture Theory Diagrams (ATDs) based	Well defined in theory.
4	Axiomatic Design	Well defined in theory.
5	Best practice conformance	In this category, we included the papers aimed to establish EA patterns and analyze real world EAs comparing them with the EA patterns or standards.
		Continued on next page

Table 4.13: Definition of the EA analysis approaches categories

Table 4.13 – continued from previous page

ID	Category	Description
6	Business intelligence based	Research in this category tried to introduce components´ operational data(such cost of a server, transactions per hour, workload etc.) together with EA models in order to produce ad-hoc analysis
7	BITAM	The Business and IT Alignment Method (BITAM) “provides a systematic, engineering-principled way of detecting and correcting misalignment from the strategic business model level down to the IT architecture level. The BITAM offers an information model with standard means of eliciting, collecting, prioritizing and organizing the information needed by the alignment/realignment process. The method invites its stakeholders to consider a range of realignment strategies— architectural and business—and provides a decision procedure for choosing among the alternatives” (Chen et al., 2005, p.2)
8	Compliance analysis	Approaches in this category aim to check if components and changes are compliant and aligned with the enterprise’s goals, principles or directives.
9	Cost benefit analysis	Well defined in theory.
10	Design Structured Matrix	Well defined in theory.
11	EA Anamnesis	“EA Anamnesis captures decision characteristics such as decision criteria and used decision-making strategy, and shows the relation between business-level and IT-level decisions.” (Plataniotis et al., 2015, p1)
12	EA executable models/Simulation	Papers in this category apply an experimental approach to simulate the behavior of EA, creating scenarios and indicators, and then for design and run experiments to back up decision-making processes on the enterprise.

Continued on next page

Table 4.13: Definition of the EA analysis approaches categories

Table 4.13 – continued from previous page

ID	Category	Description
13	Fuzzy Cognitive Maps (FCM) with Ordered Weighted Averaging (OWA) operators	"The proposal makes use of causal relations modeling using fuzzy cognitive maps. An OWA operator based on distance are used to rank the scenarios dependent upon the decision makers risk preferences." (Leyva-vázquez et al., 2014, p.1)
14	(Semi) Formalism based	It includes description languages, ontologies, set theory and other formalisms are part of this category. All those techniques try to take advantage of reasoning mechanisms to perform (semi) automated analysis of the EA, through queries, model consistency and restrictions checks, for example.
15	Functional-Business Object Matrix (FBOM) e ROCK algorithm	This method constructs a Functional-Business Object Matrix (FBOM) – where rows represent business objects and columns represent functionalities (of two types Use or Create) – and applies an algorithm called ROCK (originated in another work) to transform the matrix and generate blocks on it thus representing the potential candidates to become components in the enterprise.
16	Fuzzy based	Well defined in theory.
17	Mathematical functions	Included in this category are methods, which use mathematical functions to evaluate an aspect of EA, e.g. use of linear regression to estimate actor performance based on EA changes.
18	Metrics based	This category includes several punctual quantitative metrics to evaluate operational data from the components (e.g. performance, usage, workload) or from the overall EA (e.g. entropy). Papers in this category may introduce just one metric like "return of investment", or a bunch of metrics for EA alignment. Several concern analysis are aimed in this category since they could not be grouped in a specific one.

Continued on next page

Table 4.13: Definition of the EA analysis approaches categories

Table 4.13 – continued from previous page

ID	Category	Description
19	Multi-criteria analysis	Well defined in theory.
20	Prescriptive models	The method evaluates the AS-IS organization's maturity in each layer trough qualities aspects and based on that evaluation it prescribes actions to assist the organization to reach the intended TO-BE state.
21	Probabilistic-based	Cause and effect, uncertainty and probabilistic events are concepts present in all variations of methods belonging to this category. Typical techniques are Bayesian networks, probabilistic Bayesian networks, extended influence diagrams and fault-tree analysis. Those, in turn, are frequently used to run EA components' performance analysis
22	Structural analysis	In this category, structural aspects of the overall EA or specific layers are analyzed. Methods and techniques based on network science are employed to identify critical points, clusters or overall indexes for the EA structure.
23	Visual analysis	This category covers several techniques that use the power of visualization intrinsic to the models to extract valuable information for the experts. Typical concerns analyzed are alignment between layers, the impact of changes or failures in the overall structure.

Table 4.13: Definition of the EA analysis approaches categories

In Table 4.14, we relate each analysis approach together with their respective concerns analyzed.

Approach category	Analysis approach	Concern analyzed
Agent simulation based	Agent simulation based	Change impact
Analytic Hierarchy Process (AHP)	AHP	Scenario analysis

Continued on next page

Table 4.14: EA Analysis approaches categorization

Table 4.14 – continued from previous page

Approach category	Analysis approach	Concern analyzed
Architecture Theory Diagrams (ATDs) based	ATDs for modifiability	Impact analysis
Axiomatic Design	Axiomatic Design	Business-IT Alignment
Best practice conformance	EA analysis pattern catalog	Best practices analysis
BI	Total cost of server	EA Alignment
		EA Change
		Cost allocation
BITAM	BITAM	Business-IT alignment
Compliance analysis	Analysis of the contribution of EA components to goal achievement;	Alignment of changes with EA capabilities
	Goal evaluation algorithm	Goal compliance
	Expert reasoning about influence links and compliance to the principles	Compliance of EA decisions with Goals, requirements, principles, directives and capabilities
Cost benefit analysis	Cost benefit analysis	Prioritization of projects for EA change
Design Structure Matrix	Design Structure Matrix	Human aspects
		Business-IT Alignment
		EA Complexity
EA Anamnesis	EA Anamnesis with ex-ante and ex-post analysis	Decision rationales tracking
EA executable models/Simulation	EA executable models/Simulation	EA component performance

Continued on next page

Table 4.14: EA Analysis approaches categorization

Table 4.14 – continued from previous page

Approach category	Analysis approach	Concern analyzed
FCM (Functional Cognitive Maps) with OWA (Ordered Weighted Averaging) operators	FCM with OWA operators	Scenario analysis
(Semi) Formalism based	Event-Condition Action (ECA) rules	Model Correctness checking during EA evolution
	AgreementMaker – algorithm of ontology comparison and match	Data redundancies
	Indicator based on domain specific language	Business performance
	Queries based on Model-based Expression Language (DSL) including time series analysis	EA complexity
	Queries with formal language	EA alignment
	Ontology matching for consistency analysis	Traceability
	Queries with OWL-DL	EA change impact
	Model checking with Alloy language and rules	EA gap analysis
	Formal Reasoning with OWL	Decision rationales tracking
	Constraints rules for Decision Design Graphs	EA alignment
	Model checking	Model checking
	Propagation rules for change in ontology models	Change impact
	Description Logic Checking	EA model consistency
		Traceability
		Change Impact

Continued on next page

Table 4.14: EA Analysis approaches categorization

Table 4.14 – continued from previous page

Approach category	Analysis approach	Concern analyzed
	Description Logic Checking	Compliance of EA decisions with Goals, requirements, principles, directives and capabilities
		Model checking during EA evolution
	Traceability based on set theory	EA alignment
Functional-Business Object Matrix (FBOM) e ROCK algorithm	Functional-Business Object Matrix (FBOM) e ROCK algorithm	System quality
Fuzzy based	Data envelopment Group analysis based on fuzzy credibility constrained programming and p-robustness	Scenario analysis
	Failure mode and Effect analysis (FMEA) and VIKOR	EA implementation Risks
Mathematical functions	Linear regression to estimate actor performance based on EA changes	EA change; EA actor's competence impact
Metrics based	A set of metrics for measure IT fitness (alignment)	EA alignment
	Production function	Availability
	IT importance, IT cost and IT effectiveness analysis	EA cost; Application portfolio valuation, EA gap analysis
	Return of investment metric	Return of Investment

Continued on next page

Table 4.14: EA Analysis approaches categorization

Table 4.14 – continued from previous page

Approach category	Analysis approach	Concern analyzed
	the workload, the processing time, the utilization	EA component performance
	Entropy measure	EA Complexity
	Several mathematical formulas for EA alignment between layers	EA alignment between layers
	Several mathematical formulas for EA project prioritization	Prioritization of EA project initiatives
	Complexity metrics	Complexity
	Time, Resource use analysis	EA Component performance
	EA Viewpoint Alignment Measure	Overall alignment
	Business information object (BIO) criticality	Information dependence of an application
	Effectiveness measure	EA effectiveness
	Metric rules with OCL through models composition	EA Alignment
	Enterprise quality data indicator	EA Quality
	Overall Risk Score:	Integration risk (risk)
	Total cost of server (analysis based on BI paradigm)	Cost allocation
Multi-criteria analysis	Multiple attribute decision analysis (MADA)	scenario analysis
	Cost analysis based on technology layer attributes	Cost analysis
	Propagation framework	Goal conformance

Continued on next page

Table 4.14: EA Analysis approaches categorization

Table 4.14 – continued from previous page

Approach category	Analysis approach	Concern analyzed
Multi-criteria analysis	Propagation framework	Traceability
Prescriptive models	Prescriptive model for EA Alignment evolution	Overall alignment
Probabilistic-based	Bayesian Networks	Data quality (accuracy)
		EA quality (maintainability)
		System quality
	EID + GQMEID rules	Availability
		Risk analysis
		Goal compliance
		Business performance
		System quality
		EA change impact
		Scenario analysis
		Information security risks
	PRM	Modifiability analysis (Change cost)
		Availability
		Data accuracy
		EA Component Performance
		Application Portfolio Analysis
		Cost of change of components
		Information security risk
	Fault Tree Analysis	Dependency Analysis
	The Predictive, Probabilistic Architecture Modeling Framework (P2AMF)	Structural analysis – Modifiability analysis (Change cost)
	leaky Noisy-OR model	Availability
	(Bayesian Network based)	

Continued on next page

Table 4.14: EA Analysis approaches categorization

Table 4.14 – continued from previous page

Approach category	Analysis approach	Concern analyzed
	Bayesian Belief Networks	Risk failure
	Probabilistic rules for impact analysis	Change Impact;
		Risk failure
Structural analysis	Stakeholder Crosswalk	Stakeholder analysis
	Clustering analysis	Business-IT alignment
	Structural complexity metric	EA complexity
	Network analysis(degree in and out)	EA decisions about EA portfolio projects
	Quality of clusters maps (such decoupling, cohesion, number of clusters)	EA alignment
	Clustering algorithms + Maturity and strategic weight	Impact analysis
	Clustering algorithms + Visual analysis	Change impact
	Social Network analysis	EA domain analysis
	Network analysis (eigenvector and degree centrality)	EA scenario analysis
	Network analysis (degree centrality)	Risk analysis
	Network analysis (out-degree and eigenvector)	EA cost of change
	Network analysis (closeness and clustering, average degree centrality)	Risk failure
	Clustering techniques (Girvan/Neumann algorithm)	

Continued on next page

Table 4.14: EA Analysis approaches categorization

Table 4.14 – continued from previous page

Approach category	Analysis approach	Concern analyzed
	Markov-Chain-MonteCarlo Gibbs	Resource allocation for EA documentation
Visual analysis	Visual analysis + interviews	EA goal compliance
	Clustering algorithms + Visual analysis	Application reliability, redundancy and risk failure
	Visual analysis + ordering algorithms	Risk
	Visual analysis + Business Process analysis patterns	Alignment between layers
	Visual analysis + qualitative questions	Strategy compliance
	Graph Visual dependency analysis	Impact analysis
	Traceability analysis of goals changes and its impact on low-level components (process, applications etc.)	EA Complexity
	Visual Semantic analysis	Structural aspects
		Overall alignment
		Positional value
		EA change impact
		Stakeholder analysis
		goal compliance

Table 4.14: EA Analysis approaches categorization

The five main approaches - probabilistic-based, (semi) formalism-based, metrics-based, visual analysis and structural analysis - correspond to 57% of all work. Figure 4.6 depicts the number of papers per analysis approach category. In section 4.5, we will discuss the occurrence of this approaches and their relation to the concerns mapped by them.

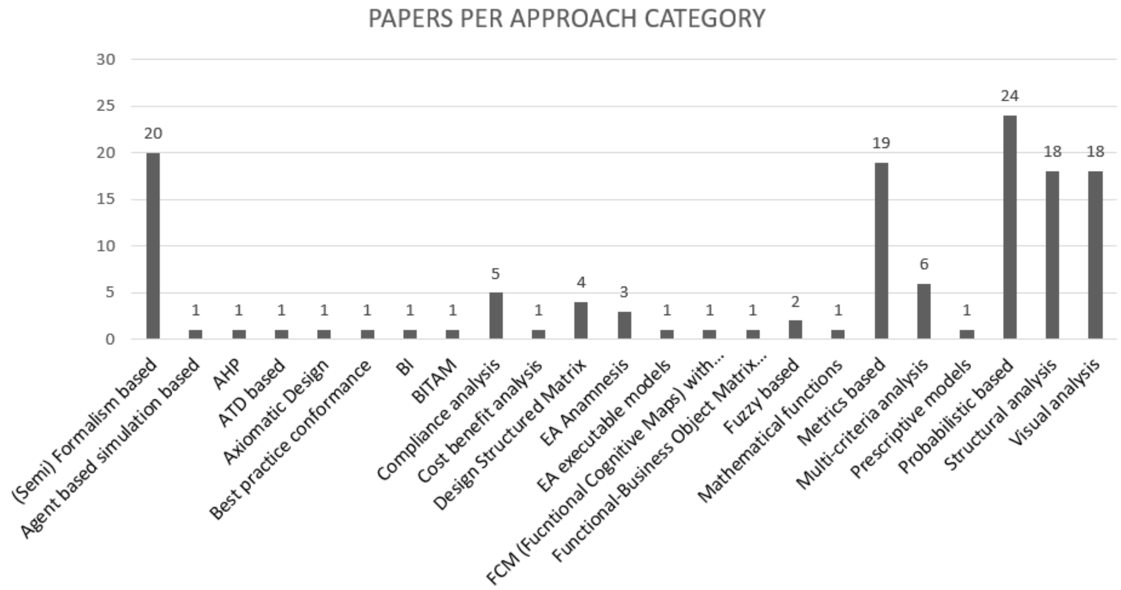

Figure 4.6: EA analysis research papers per category of approach

4.2.5 Crossing concerns, analysis approaches, models

4.2.5.1 Concerns x Analysis approaches

We perform a data crossing to evaluate the use of approaches regarding each concern in Tables 4.15. The amount of different approaches evaluating the same concern (e.g. strategy compliance is evaluated by 18 papers, using 7 (distinct) kinds of techniques) reflects the diversity of ways to perform EA analysis for the same concern and the absence of standard techniques for a specific concern. On the other hand, this offers the possibility of managers to select an approach that best fits the organization's reality.

Concern category	Analysis approach	Number of uses
Actor aspects	Design Structure Matrix	1
	Mathematical functions	1
Application portfolio analysis	Probabilistic-based	1
EA standards Conformance	EA standards conformance	1
	Visual analysis	1
Cost analysis	Metrics based	3
	Multi-criteria analysis	1
	Probabilistic based	1
	Structural analysis	2
	(Semi) Formalism based	5
	Axiomatic Design	1

Continued on next page

Table 4.15: EA Analysis approaches (categories) used for each concern

Table 4.15 – – continued from previous page

ID	Category	Number of uses
EA Alignment	BITAM	1
	Design Structure Matrix	1
	Metrics based	3
	Prescriptive models	1
	Structural analysis	2
	Visual analysis	5
EA Change	(Semi) Formalism based	5
	Agent-based simulation based	1
	AHP	1
	ATD based	1
	BI	1
	EA executable models	1
	FCM (Functional Cognitive Maps) with OWA (Ordered Weighted Averaging) operators	1
	Fuzzy based	1
	Metrics based	2
	Multi-criteria analysis	2
	Probabilistic-based	4
	Structural analysis	3
	Visual analysis	2
EA decisions	(Semi) Formalism based	1
	Cost benefit analysis	1
	EA Anamnesis	2
	Metrics based	1
	Structural analysis	4
	Visual analysis	1
EA Governance	AHP	1
	Metrics based	2
	Multi-criteria analysis	1
	Probabilistic based	3
	Structural analysis	1
Information dependence of an application	Metrics based	1

Continued on next page

Table 4.15: EA Analysis approaches (categories) used for each concern

Table 4.15 – – continued from previous page

ID	Category	Number of uses
Model consistency	(Semi) Formalism based	3
	Metrics based	2
Performance	(Semi) Formalism based	1
	Functional-Business Object Matrix (FBOM) e ROCK algorithm	1
	Metrics based	3
	Probabilistic-based	10
	Visual analysis	2
Risk	Fuzzy based	1
	Metrics based	1
	Probabilistic-based	4
	Structural analysis	2
	Visual analysis	2
Strategy compliance	(Semi) Formalism based	4
	BI	1
	Compliance analysis	5
	EA Anamnesis	1
	Multi-criteria analysis	2
	Probabilistic-based	1
	Visual analysis	4
Structural aspects	(Semi) Formalism based	1
	Design Structure Matrix	2
	Metrics based	3
	Probabilistic-based	1
	Structural analysis	5
	Visual analysis	3
Traceability	(Semi) Formalism based	2

Table 4.15: EA Analysis approaches (categories) used for each concern

In Table 4.16, we show how concerns are analyzed considering the diversity of analysis approaches.

Concern category	Distinct analysis approaches used	Total of analysis approaches used
Actor aspects	02	02
Application portfolio analysis	01	01
EA Standards Conformance	02	02
Cost analysis	05	07
EA analysis	08	19
EA change	13	25
EA Decisions	06	10
EA Governance	06	08
Information dependence of an application	01	01
Model consistency	02	05
Performance	05	17
Risk	05	10
Strategy compliance	07	18
Structural aspects	06	15
Traceability	01	02

Table 4.16: Number of distinct EA analysis approaches in relation to the total

EA change presents the widest variety of analysis approaches, with 13 (distinct) analysis approaches categories, from a total of 25 that are used to analyze this concern. This reflects the strength of this concern since one of the main objectives of EA analysis is to assist in transition situations. It is important to emphasize that works may present the same approach to more than one concern (e. g., BOUCHER et al. (2011)) or use different approaches to the same concern (e.g. TARENSKEEN; BAKKER; JOOSTEN (2015)). In both cases, each technique is counted once in each concern.

4.2.5.2 Concerns x Models

By crossing information about the models used to model each concern we can examine if there is a pattern - i.e., which modeling approaches are more applied to a specific concern. Table 4.17 shows the different modeling approaches categories used to model each concern category.

Concern category	Model category	Number of uses
Actor aspects	Not presented	1
	Own	1
Application portifolio analysis	Own	1
EA standards conformance	Archimate-based	2
Cost analysis	Archimate-based	2
	Graphs based	2
	Not presented	2
	Probabilistic networks based	1
EA Alignment	Archimate-based	2
	Combined	2
	Formal Specification Based	4
	Graphs based	2
	Non-specified	1
	Own	7
EA Change	Archimate-based	6
	Combined	1
	Formal Specification Based	2
	Graphs based	6
	Intentional modeling	1
	Not presented	3
	Own	3
	Probabilistic networks based	3
EA decisions	Archimate-based	2
	Graphs based	5
	Not presented	1
	Own	1
EA Governance	Archimate-based	2
	Combined	3
	Graphs based	1
	Not presented	1
	Own	1
Information dependence of an application	Own	1
	Archimate-based	2

Continued on next page

Table 4.17: Models (categories) used for each concern

Table 4.17 – continued from previous page

Concern category	Model category	Number of uses
Model consistency	Combined	1
	Own	2
Performance	Archimate-based	2
	Combined	5
	Formal Specification Based	1
	Graphs based	2
	Not presented	2
	Own	4
	UML based	1
Risk	Archimate-based	2
	Combined	2
	Graphs based	3
	Not presented	1
	Own	1
	Probabilistic networks based	2
Strategy compliance	Archimate-based	5
	Combined	2
	Formal Specification Based	2
	Graphs based	3
	Intentional modeling	3
	Own	3
	Probabilistic networks based	1
Structural aspects	Combined	2
	DSM based	1
	Formal Specification Based	1
	Graphs based	8
	Not presented	3
Traceability	Formal Specification Based	2

Table 4.17: Models (categories) used for each concern

EA change is the most approached concern category and it is targeted by eight different models. Archimate-based approaches are used in nine of the fifteen concern categories and graph-based in nine. This states the popularity and flexibility of both methods.

4.2.5.3 Analysis approaches x models

We organize the analysis approaches regarding the models used. The results are shown in Figure 4.7 and Table 4.18. It is possible to observe that the same technique can be used in combination with more one or more kinds of models to perform EA analysis.

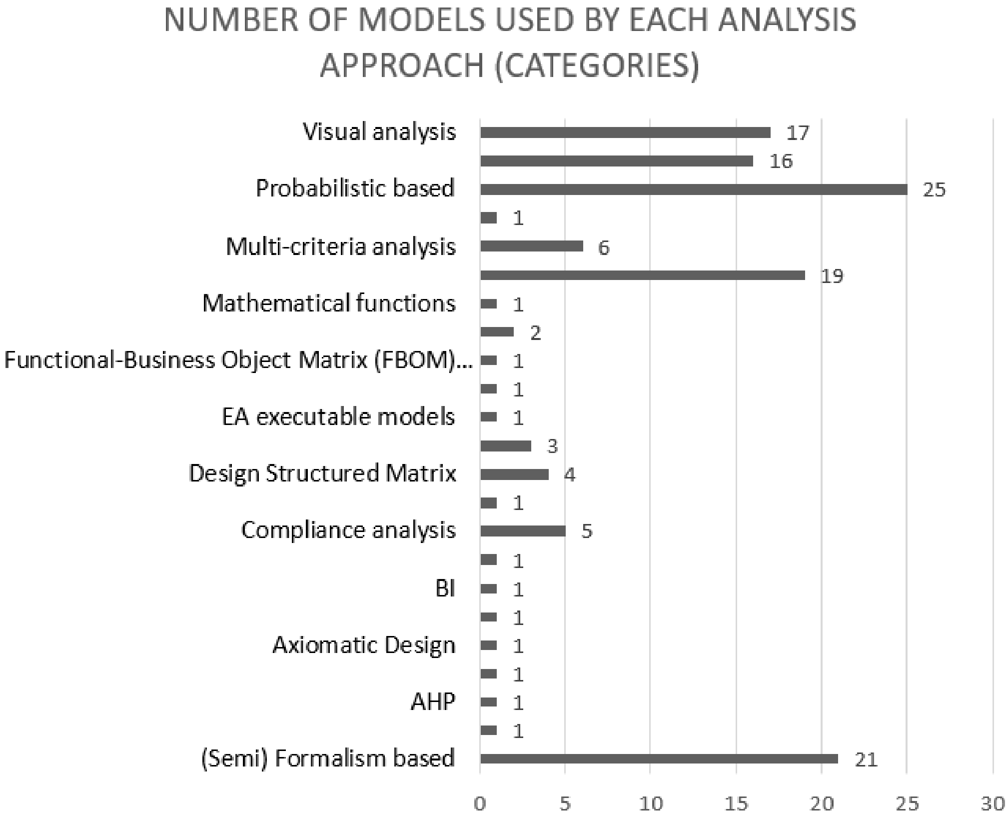

Figure 4.7: Number of models used by each EA analysis approach

Analysis approach category	Model category
(Semi) Formalism based	Archimate-based Combined Formal Specification Based Graphs based Own
Agent based simulation based	Own
AHP	Combined
ATD based	Not presented
Axiomatic Design	Own
Best practice conformance	Archimate-based
BI	Intentional modeling
	Continued on next page

Table 4.18: Models (categories) used by each analysis approach category

Table 4.18 – continued from previous page

Analysis approach category	Model category
BITAM	Own
Compliance analysis	Archimate-based Graphs based Intentional modeling Own
Cost benefit analysis	Not presented
Design Structured Matrix	Graphs based Own
EA Anamnesis	Archimate-based
EA executable models	Archimate-based
FCM (Functional Cognitive Maps) with OWA (Ordered Weighted Averaging) operators	Archimate-based
Functional-Business Object Matrix (FBOM) e ROCK algorithm	Formal Specification Based
Fuzzy based	Not presented
Mathematical functions	Not presented
Metrics based	Archimate-based Combined Graphs based Non-specified Own UML based
Multi-criteria analysis	Archimate-based Graphs based Own
Prescriptive models	Own
Probabilistic based	Archimate-based Combined Own Probabilistic networks based
Structural analysis	DSM based Graphs based Own

Continued on next page

Table 4.18: Models (categories) used by each analysis approach category

Table 4.18 – continued from previous page

Analysis approach category	Model category
Visual analysis	Archimate-based Combined Graphs based Own

Table 4.18: Models (categories) used by each analysis approach category

4.2.6 Considerations about theoretical, simulated and empirical EA analysis

We extracted infromation from the publications regarding the nature of the data they used. The "theoretical" category refers to works that describe the analysis approach but do not present any case studies to demonstrate the approach. For instance, BEETZ; KOLBE (2011) presents a set of metrics to measure complexity aspects of EA but does not apply them to a real-world scenario. The "simulated" category covers works that presented an example of what would be the application of the approach in the real world, as is the case of many papers that use The Open Group's Archisurance example - e.g., S.; S.; V. (2013) and GUNGOR; OGUZTUZUN (2014). The latter points that despite "Archisurance is a well-known and widely used case study; it is still a fictitious scenario" and the best way to assess the validity of their method would be a real-world case study. Finally, papers in "empirical" category bring a real case study presenting empirical results of the evaluation approach, as GAMMELGÅRD; EKSTEDT; NÄRMAN (2007) presents. Figure 4.8 displays the distribution of the analysis approaches according to this classification.

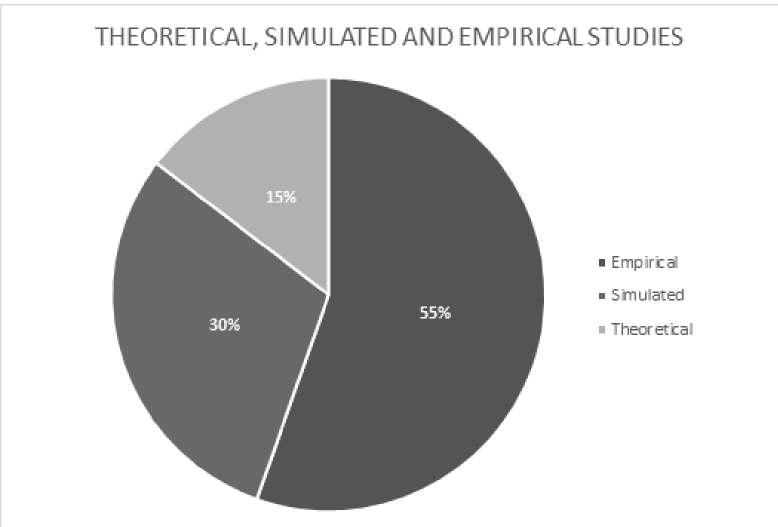

Figure 4.8: Distribution of theoretical, simulated and empirical studies

Almost half of the works proposed analysis approaches only at the theoretical level, or present illustrations to demonstrate their methods. Although the majority of the publications in our research presented empirical approaches, there is still a need for more detailed empirical demonstrations. Several works did not present enough information about how the study was conducted and the benefits obtained from the analysis approach.

The absence of details also raises an issue regarding the data required to perform a specific analysis approach. Usually, they require a large number of data specifications (regarding business and IT components) based on the assumption that all those specifications are available. Most of them are not adaptable to the possibility of incomplete data, the reality of many organizations (GMATI; RYCHKOVA; NURCAN, 2010). Therefore, although many theoretical and simulated approaches may present useful techniques, they might not apply to real scenarios where a lack of data is an issue.

Table 4.19 shows the distribution of empirical, simulated and theoretical approaches per concern. EA alignment, EA change, and strategy compliance (the three most studied concerns) present most of the research work empirical. This reflects the maturity of this concerns, as they are directly related to the main objectives of EAM -maintain the enterprise alignment, support the decision-making and scenario selection, and assure the compliance with organization's strategic objectives.

Concerns	Empirical	Simulated	Theoretical
Actor aspects	1	1	0
Application portifolio analysis	0	1	0
Best practice	2	0	0
Cost analysis	7	0	0
EA alignment	11	6	2
EA change	15	9	1
EA decisions	6	4	0
EA governance	4	4	0
Information dependence of an application	1	0	0
Model consistency	2	3	0
Performance	2	6	9
Risk	8	1	1
Strategy compliance	11	3	4
Structural aspects	6	6	3
Traceability	1	1	0

Table 4.19: Number of empirical, simulated and theoretical studies per concern

On the other side, performance related analysis is a clear example of the need for empirical validation, as most of the analysis approaches presented only a simulated or theoretical approach. Instead, impact analysis, represented by EA change and structural aspects, have most of the work related presenting empirical research as in LANGERMEIER; SAAD; BAUER (2014). The absence of empirical studies also might reflect the lack of tools to support the analysis process, since perform some of those methods manually would require much more time and effort. This issue will be more discussed in the next section.

4.2.7 Tool support for EA analysis

The lack of tools to support analysis techniques is an issue reported by many publications. We grouped the evidences about EA analysis tools and related issues in the categories described in Figure 4.9.

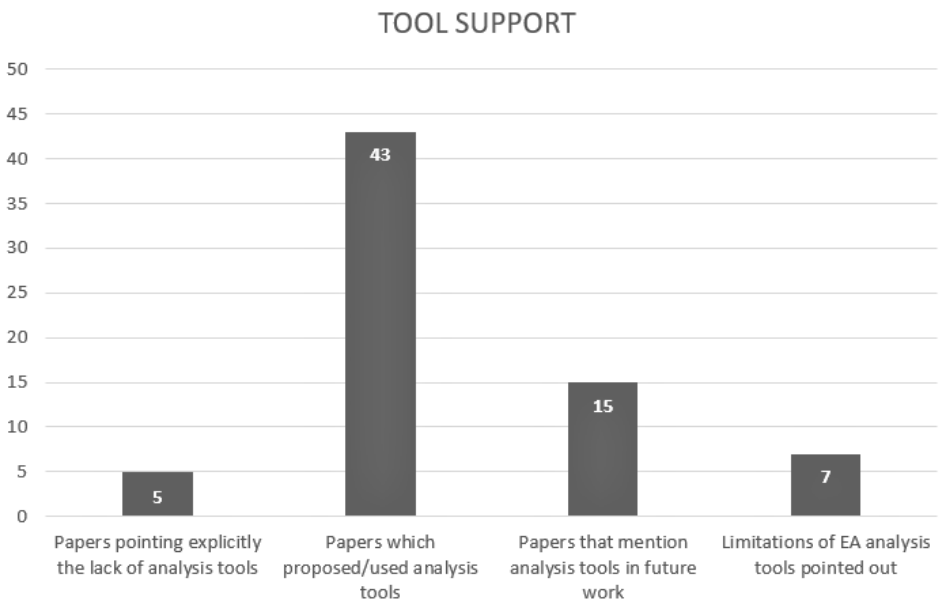

Figure 4.9: EA analysis tool support considerations extracted from primary studies

Paper pointing out explicitly the lack of tools. Authors argue about the need for tools that support a specific kind of analysis (BUCKL et al., 2009; CASTELLANOS; CORREAL; MURCIA, 2012; DAM; LÊ; GHOSE, 2015) or point the necessity of tools in the general field (PLATANIOTIS et al., 2015; MANZUR et al., 2015) in 5 papers.

Papers pointing the limitations of the existent tools. Another limitation pointed is the inflexibility of existing tools, as they do not adapt to the reality of real organizations. Several tools rely on a static meta-model structure, as organizations tend to adapt languages and models to their reality (RICO, 2006; CASTELLANOS; CORREAL; MURCIA, 2012; LANGERMEIER; SAAD; BAUER, 2014). Furthermore, according to LAGERSTRÖM; JOHNSON; EKSTEDT (2010), existing tools do not offer support to multiple kinds of analysis or support only a phase of the evaluation process, obligating organizations to use multiple tools to apply a technique (ÖSTERLIND; LAGERSTRÖM; ROSELL, 2012).

Papers mentioning the use tools or propose enhancements of existentones. From

120 primary studies, 43 papers used an existing tool or proposed enhancements for an existent one.

Papers mentioning the development analysis tools in future work: 15 (12,5%) papers mention the development of an analysis tool to support the proposed technique in the future.

4.2.8 Qualitative/Quantitative aspects in EA analysis

Analysis approaches may have a qualitative or quantitative character. The qualitative analysis focuses on the assessment of EA qualities, not focusing on numerical data. This type of methods provide information based on questions such as “is the AS-IS state of EA aligned?”, ”is this scenario compliant with the strategy goals?”, “which would be the impact of changes in this element?”. In contrast, quantitative methods cover the quantitative aspect of relationships between different enterprise architecture elements and layers, using usually metrics and functions to assess different aspects of an EA (ŠAŠA; KRISPER, 2011). According to our SLR, almost half of studies present qualitative approaches for EA analysis (46%), while 42% were quantitative, as shown in Figure 4.10.

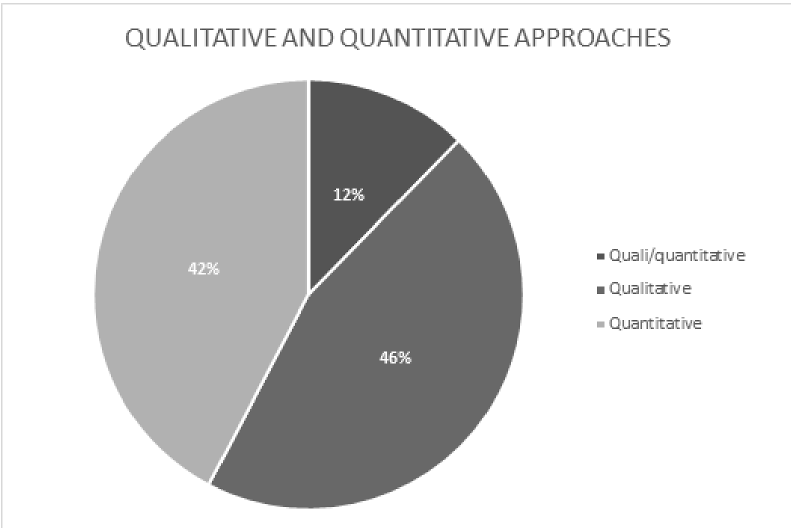

Figure 4.10: Qualitative and quantitative approaches of EA analysis

Most of the qualitative studies were represented by analysis approaches of visual analysis, formalism based techniques and other punctual techniques such AHP, multicriteria, EA standards conformance and so on. Concerns related to strategy compliance and EA Alignment have the largest quantity of qualitative studies (see Figure 4.11).

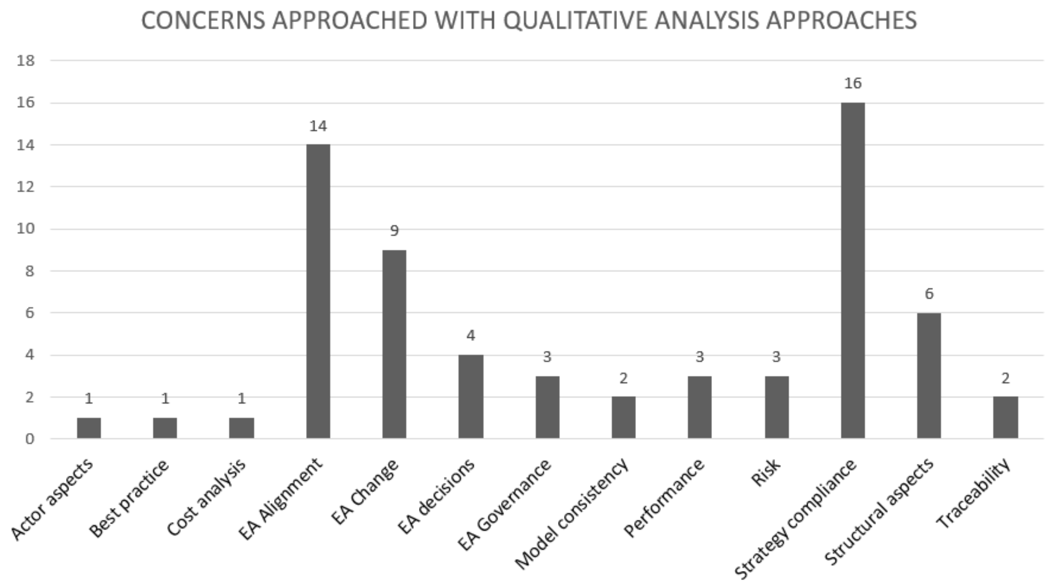

Figure 4.11: Qualitative approaches by concern category

The imminent studies in the quantitative analysis were probabilistic based techniques such PRMs, metric-based and structural analysis (mainly represented by network analysis). EA change and performance have the higher incidence of quantitative studies (see Figure 4.12). Performance, alone, represents 22% of the quantitative studies and most of those initiatives (61,5%) use probabilistic-based methods, a trend in this concern analysis.

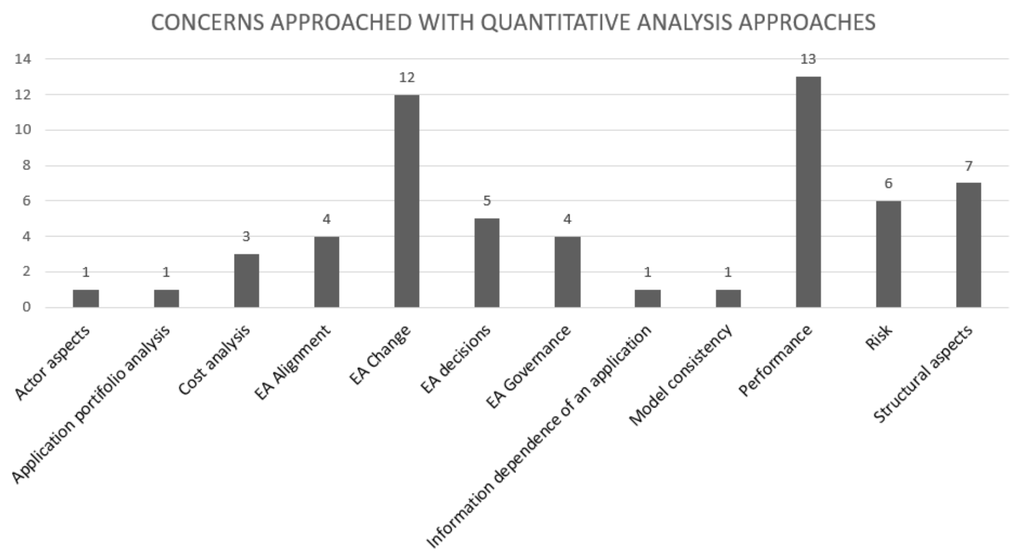

Figure 4.12: Quantitative approaches of EA analysis by concern category

There was also mixed-approaches (qualitative and quantitative approaches) representing 12% of found works. For instance, for FASANGHARI et al. (2015), particularly the EA scenario analysis should be considered as a multi-criteria decision management approach that encompasses both quantitative and qualitative criteria to interconnect both IT and business-related issues to achieve the desired output comprehensively.

4.2.9 Data sources and data collection methods in EA analysis

In our SLR, we also investigate an important aspect of model-based analysis approaches: the data source. GMATI; RYCHKOVA; NURCAN (2010) state that some techniques require an extensive set of data as input that organizations often cannot provide. According to those authors, this makes the implementation of these approaches difficult (if at all possible). This is also a reason why many approaches do not present empirical results (as discussed in section 4.4.1).

EA analysis techniques must adapt to incomplete data since the complete specification of a system may be complex and require a time and effort the organization does not have. Thus, to create a "complete" EA model is not a trivial (if it is possible) task as "information on complex application landscapes is both incomplete and beyond comprehensibility" (HOFER, 2013).

NÄRMAN et al. (2011) present strengths and weaknesses of different data collection methods, as depicted in Table 4.20. They also discuss the cost of data collection approaches, arguing they may variate regarding "on the ease of retrieving these artifacts and the amount of data which need to be processed".

Data collection method	Strengths	Weakness	Cost
Questionnaires	Can evaluate large samples.	Difficult to reach in-depth data. Difficult to describe complex situations.	They are an efficient tool.
Interviews	Focuses on the topic. Provides perceived causal inferences.	Possible bias.	The cost of interviews depends on the format of them.
EA artifacts (documents)	Stable, Unobtrusive, Exact, Quantitative.	Difficult to retrieve information	Depends on the ease of retrieving these artifacts and the amount of data which need to be processed.

Table 4.20: Strengthens and weakness of different data collection methods (YIN, 2003) contextualized in EA (adapted from NÄRMAN et al. (2011).

We classified the data source as documents (artifacts as models and descriptions), interviews or automated, as depicted in Figure 4.13. 43 of 120 studies did not mention the data collection approaches to build their EAs. Among the remaining 77, the data sources were distributed in documents (54%), interviews (45%) and tool supported (5%).

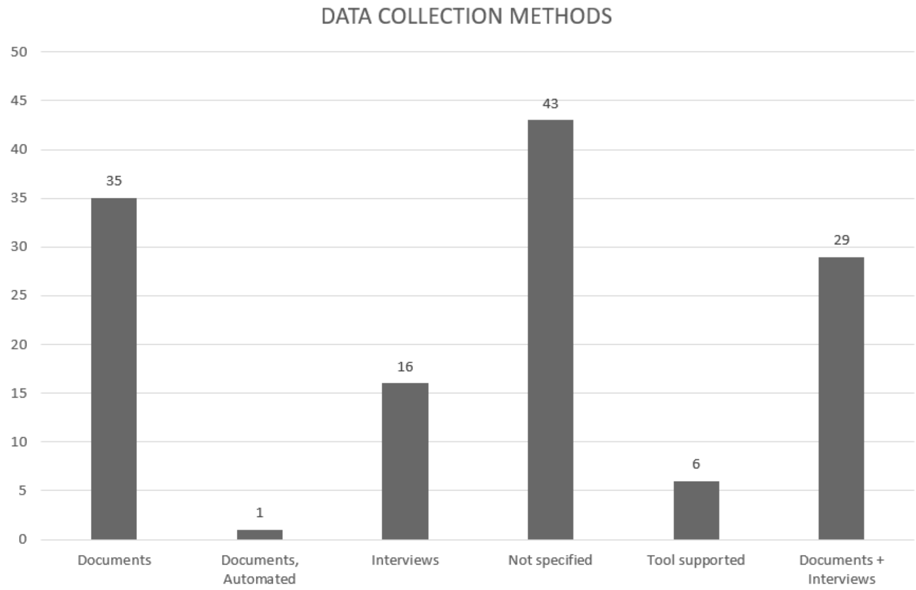

Figure 4.13: Distribution of the data collection methods in EA analysis papers

Most of the data collection approaches are performed manually, which is not only time consuming but difficult when it comes to model updates (CHEN; CHIANG; POOR, 2013). There is no common methodology or notation for EA documentation reported in the studies. Another important issue during the data collection process is to have high-quality input from experienced process performers. Otherwise, results credibility would be affected (NÄRMAN et al., 2011). Manual data collection requires effort and resources not always available in organizations.

Automated data collection would reduce the effort and provides an efficient mean to reuse and interconnect existing data sets, and assist in the "related data management processes to keep the information within an EA up-to-date" (CHEN; CHIANG; POOR, 2013). However, automated data collection is quite challenging in terms of defining and recognizing components such business process and applications, since gathering the correct data distributed across the organization is not a trivial task.

4.3 CONCLUSIONS

Motivated by the absence of a state-of-the-art in EA analysis, we presented a systematic literature review, in order to answer the question "what are the main analysis concerns, approaches and models and employed in EA analysis field?"

With a final set of 120 papers, we performed a thorough EA analysis SLR. An initial finding is the lack of a common definition of the term EA analysis. Most papers do not present a definition. Instead, they refer to existing ones or only bring justifications regarding why they are performing it. Based on the main existing definitions of EA analysis found in the literature, we proposed our own definition: EA analysis is the assessment of any EA´s property, based on models or other EA related data, in other to inform or bring rationality to decision support of stakeholders. We believe this description covers the main aspects of the field.

We studied the main analyzed concerns, classifying them in general categories. Most of the existing analysis approaches are related with five categories: EA change, EA alignment, strategy compliance, performance and structural aspects. The first three of them are present in half of the studies while all of them take part in 77, 5% of the primary studies. In total, we defined 43 concerns and 15 main categories.

We identified methods and techniques used to perform EA analysis and classified them in 23 different categories. Probabilistic-based techniques were present in 20% of the studies, representing the main analysis approach category. (Semi) formalism based, metrics-based, structural analysis and visual analysis, are the other predominant categories. The five together represent 82% of all studies. Most of the studied analysis approaches were model based (88%). Thus we analyzed the state of art of models used in EA analysis. We identified eight main classes of models, covering 19 kinds of modeling approaches. An interesting data was the amount of self-developed models, which corresponded to 20% of the identified models. Archimate-based models were the main used models, while self-developed and graph-based models , were the second and third main model choices, respectively.

By crossing data about concerns, models and analysis techniques, we found that some concerns, e.g. EA change, are targeted by a combination of several different approaches and models. Despite the diversity of models and analysis approaches, there is a general call for more empirical validation in the field, as almost half of the primary studies presented theoretical or simulated research.

We did not identify any common guidelines to perform EA analysis. The vast majority of the works did not indicate analysis activities performed, giving little information about their EA analysis process workflow.

4.3.1 An attempt of research agenda for EA analysis.

The results of our SLR highlight several research avenues in EA analysis, as we describe in the following.

Explore further some concerns. Surprisingly, despite EA being a socio-technical discipline, human aspects were almost not taken into consideration - with only two approaches on the topic identified. Concerns related to application portfolio analysis, best practices, information dependence of an application also receive little attention.

Perform empirical evaluations with analysis approaches. Another issue on EA analysis field is the lack of empirical works to validate some analysis approaches. Concerns related to EA performance demonstrate this: though it is the fifth most analyzed category, only two of the 17 publications covering this topic present empirical validation, with almost half of the works presenting theoretical or simulated approaches, what clearly calls for more empirical validation.

Development of EA automated data collection methods. There is a lack of tools for supporting the EA analysis process, from its very initial data collection phase. The absence of EA documentation in the organizations is a factor that influences in the scarcity of papers performing empirical EA analysis. The field may benefit considerably from tools that helps to automate or make the data collection process easier. The existing automated approaches are still incipient and need further development.

Development of support tools for EA analysis. The analysis process itself is usually complex and it might require time and resources the organization does not have. As discussed in Section 4.2.8, only 35,8% of the primary studies presented or used a tool to support their analysis and many stated the lack of this asset.

Define an EA analysis process. Most of the primary studies did not detail their data collection process, though many pointed the shortage of tools to support this process. We did not find any generic process or guidelines to perform EA analysis. However, initial ideas are discussed in FUERSTENAU; ROTHE (2014); ADDICKS; APPELRATH (2010); VENEBERG et al. (2014).

4.3.2 Limitations

Despite there is not a consensus in the literature about EA and EA analysis, we have adopted broad definitions for both concepts to derive our research search, covering several related terms and definitions adopted in the literature.

In our SLR, we did not perform a qualitative assessment of primary studies. We accepted intentionally all the works that aimed to perform EA analysis, without a very strict quality criteria, to be able to have a broad understanding of the field and the authors' purpose.We consider the 120 papers in the final set give a good perspective on state of the art in the field.

4.3.3 Future works for EA analysis

In order to obtain a full perspective of the EA analysis field, we aim to improve this work investigating the practice of EA analysis, surveying practitioners in order to complement the presented academic-based perspective. Understanding both perspectives is essential to develop aligned methods and techniques suitable for the practice in real organizations. We also aim to validate the categories defined for concerns, approaches and models with other EA practitioners through the application of a survey. This would help to build standardization of these categories in the field, prioritize analysis approaches and concerns, and promote the development of tools.

4.4 Chapter Summary

In this chapter, we aimed to build a knowledge base about EA analysis and portrait the current status and existent gaps. The contributions of this SLR for our overall research are twofold: first, it allowed us to get familiarized with definitions of EA analysis, EA models, EA analysis techniques and main empirical related issues. Secondly, we were able to position the EA structural analysis paradigm in the overall EA analysis scenario,a s discussed in Section 2.2.7.

Gathering different analysis approaches, we believe that the presented EA analysis consolidation may enable organizations to select techniques and methods that best suit their current architecture and analysis needs. Finally, it may also allow researchers to recognize opportunities in the field, focusing on concerns and techniques that did not receive attention or enhancing existing methods, adapting them to different organization's architectures.

5

TOWARDS THE EANA META-MODEL AND LIBRARY

In the previous chapter, we focused in the EANA's main constructs such as concerns, models, and analysis approaches to have a general comprehension of the EA analysis process. We now perform a more focused SLR in the EANA context. Following our research baseline, we investigate the state of art of EANA in Section 5.1. Grounding on its findings, we design the first two proposed DSR artifacts, the EANA-MM (section 5.2) and the EANA library (section 5.3). Our considerations about these artifacts are made in section 5.4.

5.1 THE EANA SLR

In chapter 1, we elicited our reasons to investigate further the EA with the structural paradigms. One of the reasons was that EA, as an intertwined network of components and relations, is intrinsically suitable for the network (graph) modeling approach. Later, we saw some foundational limitations to uncover the network analysis potential for EA context such as 1- no common language shared by researchers; 2- no clarity about what concerns can be analyzed with network analysis initiatives and 3- techniques and methods' implementation are not clear in the papers.

In this chapter, we aim to contribute to minimize those gaps, at first, fostering a set of EANA metrics and methods, classifying and organizing them in a repository of NAIs.

Since a complex network has many interconnected components (SCHNEIDER et al., 2015), in our first exploratory and ad hoc review, we identified research which also has been analyzing EA components and interrelations under the label "EA complexity." SCHNEIDER et al. (2015) proposed three categories to classify complexity metrics in EA. Their "topology metrics" category contains studies developed under the EA complexity perspective which are aligned with the EA network model discussed in our present work. Their "topology metrics" identified essentially represent network methods.

With this in mind, we want to make a distinction between our work and SCHNEIDER et al. (2015). The present work differs in aspects like the coverage of our query string and types of research objectives. We try to characterize the state of the art network measures applied over EA components and relationships, not necessarily only within the EA complexity literature.

We also aim to focus on structural aspects of intra- and inter-relations of EA components as in SIMON; FISCHBACH (2013), while paying less attention to the properties of individual components and heterogeneity, which were central aspects focused in SCHNEIDER et al. (2015).

Thus, in order to identify the network analysis measures that were applied in EA, it is necessary to look at the existing body of knowledge of EA complexity (see Figure 5.1, Areas 2 and 3) due to overlaps with concepts discussed. Furthermore, we identified works which dealt specifically with the network analysis of EA components, its relation to all EA subdomains, and adding these respective terms in our query string (see Figure 5.1, Area 1). In Figure 5.1, Area 1 – our focus –represents potential network analysis works not mapped in EA complexity literature. With regard to Area 2, it depicts the overlap between EA complexity and EA network analysis. Area 3 in Figure 5.1 represents possible works related to EA network analysis that might not explicitly contain network-related terms, due to no standardized nomenclature in the fields.

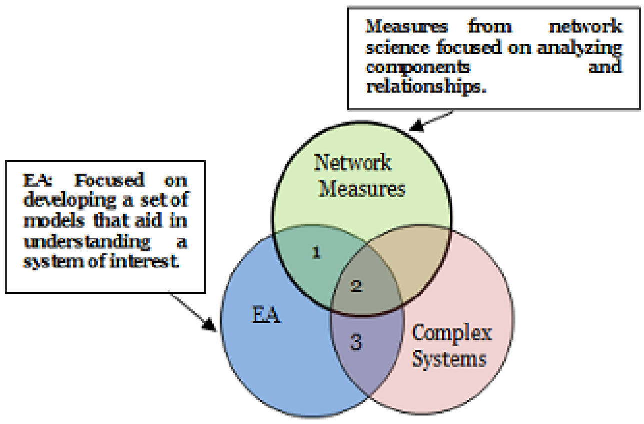

Figure 5.1: EANA research focus and overlapping areas

Therefore, our goal is to map techniques and methods implemented and found in the literature, analyzing their EA concerns targeted, which will be consolidated in order to answer the following research question (RQ1.2):

What is known about the application of network measures in order to analyze components and relationships in the EA context?

We break down the previous question as follows:

- What are the analysis concerns of these studies?
- What are the main methods of data collection used?
- What are the perceived effects/benefits of the application of NAIs in the EA context?
- What are the research gaps?

As discussed in chapter 3, we use the guidelines of KITCHENHAM; CHARTERS (2007) to perform a broad literature review and create a state of the art description of EANA, its applied measures and major accomplishments. We also provide a research agenda for the field. These contributions will be presented along this chapter. In the next section, we start describing the SLR design.

5.1.1 The research protocol

The main parts of our research protocol are described in the following.

- Research query

We derive our search terms from the definition of EA presented by SCHÜTZ; WIDJAJA; GREGORY (2013). In addition, we used elementary terms related to network science like "network measures," "centrality," and "network analysis," since these terms could represent research related to relations analysis. Three research experts in the network analysis field have validated these related terms. As SIMON; FISCHBACH (2013) did, we believe that querying all EA subsystems terms together with "enterprise architecture" might produce a broader coverage of research. According to SCHÜTZ; WIDJAJA; GREGORY (2013), there are possibly various analysis approaches from other fields which might be transferable to EA. For example, research from System of Systems Theory (SoS), system thinking theory, system engineering and others which did not specifically mention the expression "enterprise architecture" did cover one of the EA subsystems – application architecture – and presented some network measures/metrics. In those cases, we investigated if those measures had the potential to be listed as a contribution to EA analysis research. Our query string is presented in Table 5.1.

Research query string adopted
("enterprise architecture" OR "business architecture" OR "process architecture" OR "information systems architecture" OR "IT architecture" OR "IT landscape" OR "information architecture" OR "data architecture" OR "application architecture" OR "application landscape" OR "integration architecture" OR "technology architecture" OR "infrastructure architecture") AND ("complexity" OR "centrality" OR "network analysis" OR "network measures" OR "betweenness" OR "eigenvector" OR "degree" OR "network metrics" OR density OR "modularity" OR "clustering" OR "cluster")

Table 5.1: Research query string adopted for the EANA SLR

- Selected engines

The query was executed in March 2015 for Scopus, IEEE, and Engineering Village; and executed again in August 2015 for EBSCO and AISeL engines. We did not limit the initial year of the publications.

- Inclusion and exclusion criteria

We describe our inclusion and exclusion criteria in Tables 5.2 and 5.3, respectively.

Criteria	Description
CI-01	Papers analyzed had to contain techniques, methods or any network-related initiative to evaluate the infrastructure, application, information, or business process architectures (any EA subsystem) or the EA in a holistic way (EA as a whole system). Different authors allude to EA complexity but are actually referring to different facets or types of EA complexity (MOCKER, 2009). In our research, we focused on the complexity related to interdependence among EA components, which can be of two types: intra-architecture (e.g. AA) and inter-architectures (e.g. AA and BA).

Table 5.2: Inclusion criteria for papers of the EANA SLR.

Criteria	Description
CE-01	Papers containing other types of complexity like diversity or heterogeneity and deviation from standards (MOCKER, 2009) without considering any network measures.
CE-02	Papers related to product architecture or software code analysis.
CE-03	Papers in any language but English.
CE-04	Papers with files that are not recoverable in full.
CE-05	Summaries of keynotes, tutorials, white papers, book chapters, theses or dissertations.
CE-06	Incomplete papers.
CE-07	Repeat studies found in different sources or reporting similar results (e.g. DREYFUS; IYER (2008) replaces a similar work in DREYFUS; IYER (2006)).
CE-08	Papers containing measures not related to a component's interdependency like "human resource workload at the business process level," an analysis function identified by RAMOS et al. (2014).

Table 5.3: Inclusion criteria for papers of the EANA SLR

- Screening phases

The papers were selected through the following screening steps described in Table 5.4

Phase	Papers outputted
0-Papers returned from the engines	5739
1-Title and abstract readings	262
2-Introduction and conclusions readings	66
3-Full paper readings	24

Table 5.4: Screening phases of EANA SLR

We also performed forward and backward searches which resulted in additional five papers to our dataset after applying the I/E criteria. In the last phase, we extracted the evidence from 29 (24+5) papers.

- Data extraction

We extracted the information described in Table 5.5 from the final set of selected papers.

Information group	Information	Description
Paper	Sequential	Information for internal control.
	Paper_ID	Information for internal control.
	Document title	The publication´s title.
Analysis strategy	EA layer	Targeted EA layers: Value, Business, Application, Technology.
	EA components	EA components used in the analysis. For instance: a goal mission, strategy for the value layer; Stakeholder,Business process, business functions,product in the business layer; Information objects, business information objects, entity for the information layer; Information systems, application, software component for the application layer;application server, database server for the technology layer.
	EA component´s attribute	Attributes used in the analysis. E.g.technology vendor, , the strategic value of an application etc.
	Level of analysis	It can assume the component, group, network levels.
	ID	Information for internal control
	Name(label)	Name of the NAI

Continued on next page

Table 5.5: Data extracted from the selected primary studies for EANA SLR

Table 5.5 – continued from previous page

Information group	Information	Description
NAI identification	Calculation	A brief resume of how the NAI is calculated or some external source on which it is explained.
Analysis concern	ID_Concern	Information for internal control
	EA_Concern	EA concern under analysis
	ID_Question	Information for internal control
	EA_Question	The operational question that is used to guide the NAI selection. It explains what kind of analysis is performed.
Data strategy	EA Data generation strategy	It can assume the following values: (Primary data; Derived data; Simulated data; No data)
	Unimodal or Multi-modal?	It represents the type of network under analysis.
	Data source	It is related to the data sources or data collection techniques used such as documents, diagrams, interviews, automated tools and so on.
Modeling decisions	Relation meaning	This is the semantics of the relationship among the modeled components. E.g. "Application is connected with application.
	Directed or undirected data	It qualifies the direction of the relation between two components.
	Transitive relations	It determines if the transitivity property for the relation is available.
	Weighted?	It indicates the presence of weighted relations.
Output	Output type	It describes the type of produced outputs such as graphs, rankings, lists, diagrams, clusters etc.

Table 5.5: Data extracted from the selected primary studies for EANA SLR

The outputs of the EANA SLR process are described in the next section.

5.1.2 Results of the SLR

5.1.2.1 SLR meta-data

As depicted in Figure 5.2, the trend of growth of EA publications presented by Simon, Fischbach e Schroeder (2013b) and also identified with our EA analysis SLR in chapter 4, again is reflected in EANA publications, since they become more frequent after 2006.

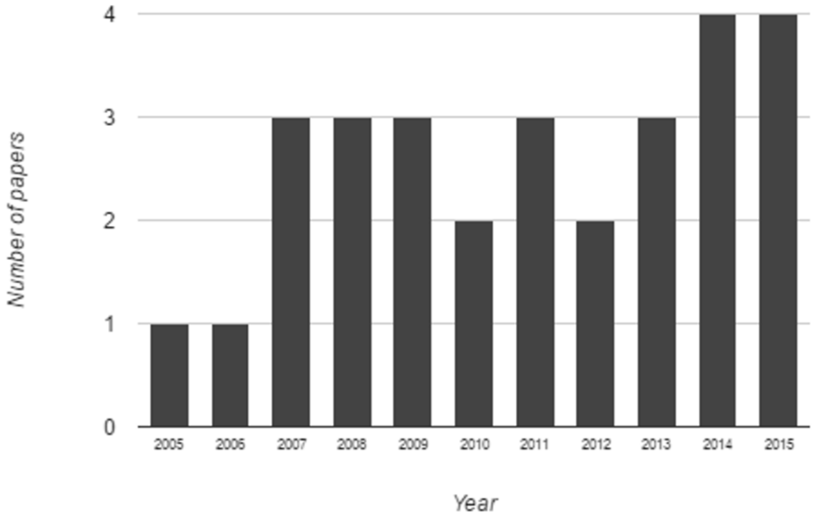

Figure 5.2: Distribution of papers per year for the EANA SLR

According to the Figure 5.3, Germany and USA are the predominant countries in terms of contributions to the field. All in all, European countries concentrate the biggest part of research efforts with the participation of Germany, Finland, Switzerland, The Netherlands and Sweden. The University of Los Andes in Colombia gives a significant contribution from South America as well.

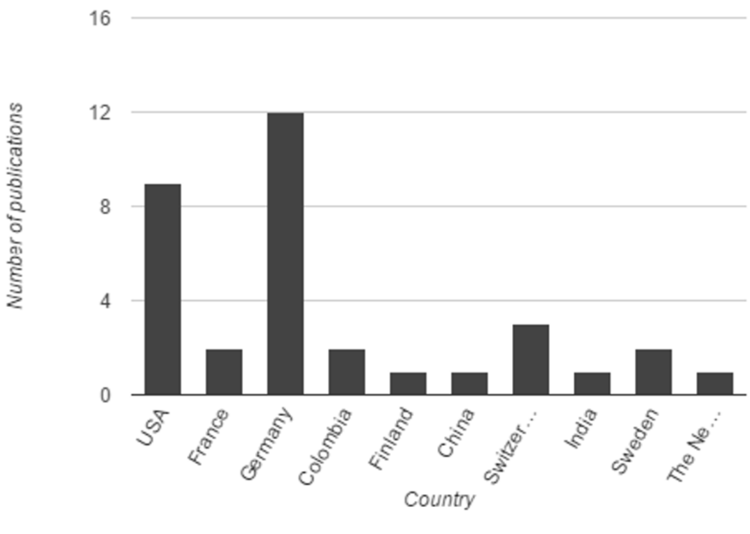

Figure 5.3: Distribution of papers per country for the EANA SLR

5.1.2.2 Categories of EANA concerns mapped

We identified fourteen categories of EANA concerns which are described in Table 5.6. The right column in Table 5.6 presents examples of primary studies which investigated the analysis concern described in each line.

Cat_ID	Concern Category	Description
C_01	Information dependence among applications	It is related to application dependency in terms of information (business objects) that one application may have to others. For instance, one can compute the number of valuable BOS produced by third-parties which are used by a specific application (WENDT; BRIGL; WINTER, 2005). In Addicks (2009), in order to compute the business information criticality of an application "a" based on their BOS, some refinements are considered: on which BOS the application "a" has the data sovereignty; and which BOS produced by third-parties and accessed by application "a" manually and automatically.
C_02	Key structural components	NAIs in this category focus in pure structural aspects, applying centrality metrics such as degree centrality, eigenvector, betweenness and others. The goal is to identify the architectural components which occupy significant structural positions. Examples of such studies are HOMMES (2008), FÜRSTENAU (2015) and SIMON; FISCHBACH (2013).
C_03	EA modularity	The goal of these metrics and methods is to evaluate the modularity of the EA layers individually or the EA layer as a whole. Modularization is used to manage the EA complexity in Hommes(2008), SIMON; FISCHBACH (2013); CHIRIAC et al. (2011).
C_04	EA change	It studies changes occurred in specific components and their impact on the remaining components (Hommes et al., 2008; (SINGH; SINGH, 2010)
C_04.1	EA cost of a change	In this subcategory, researchers tried to estimate the cost of changes considering the impact in the overall structure (SIMON; FISCHBACH, 2013; DREYFUS; WYNER, 2011; LAGERSTRÖM et al., 2014).
		Continued on next page

Table 5.6: Categories of EANA concerns mapped from the EANA SLR

Table 5.6 – continued from previous page

Cat_ID	Concern Category	Description
C_05	Stakeholder management	This concern comprises works which deal with stakeholder management and their involvement in the EA changes, regarding impacts or responsibilities. BARTOLOMEI et al. (2012) investigated the role and influence of stakeholders in order to involve the right people in taking the right decision about the architecture. WOOD et al. (2013) tried to capture the active stakeholders along different business process phases.
C_06	EA complexity	This category comprises indicators to measure complexity in the whole EA or its layers (KREIMEYER; GÜRTLER; LINDEMANN, 2008; BEETZ; KOLBE, 2011).
C_06.1	EA Application complexity	In this subcategory, researchers investigate application layer complexity (SCHNEIDER et al., 2015; NARANJO; SÁNCHEZ; VILLALOBOS, 2014).
C_06.2	EA business complexity	In this subcategory, researchers investigate business layer complexity (SCHNEIDER et al., 2015).
C_06.3	EA technology complexity	In this subcategory, researchers investigate technology or infrastructure layer complexity (SCHNEIDER et al., 2015)
C_07	EA domain analysis	The goal here is, based on clustering algorithms, to identify groups or clusters of components (LI et al., 2007). Another possible analysis is to compare cluster formed by algorithms with the ones produced manually by organization decomposition (BARTOLOMEI et al., 2012).
C_08	EA improvement	This analysis aims to identify key structural EA components which can represent important improvement candidates to be prioritized in case of architectural redesign LEE et al. (2014).
C_08.1	Business process improvement.	This category is about to identify business process improvement candidates with the support of network metrics. (LEVINA; HILLMANN, 2012)

Continued on next page

Table 5.6: Categories of EANA concerns mapped from the EANA SLR

Table 5.6 – continued from previous page

Cat_ID	Concern Category	Description
C_09	Application usage	The goal of metrics in this category is to identify how much specific applications are accessed by others (SIMON; FISCHBACH, 2013).
C_09.1	Cross-functionality	In this category, metrics and methods investigate which applications are accessed by multiple EA functional domains (e.g. marketing, billing, sales etc). A proxy for that was used in SIMON; FISCHBACH (2013) was the betweenness centrality.
C_10	Risk failure management	The metrics in this category evaluate the impact of a component failure in the remaining architecture, helping to perform risk management (SIMON; FISCHBACH, 2013; HOMMES, 2008).
C_11	Local dependency	This is related to the neighborhood analysis for a given component, focusing on the applications most dependent and the ones which generate the biggest dependencies (PRIMROSE: A GRAPH-BASED APPROACH FOR ENTERPRISE ARCHITECTURE ANALYSIS, 2015; SINGH; SINGH, 2010).
C_12	EA alignment	Works in this category investigated the alignment among components from different layers (BOUCHER et al., 2011)
C_12.1	Business-application alignment	It deals with the specific alignment between components from the application and business layers. A typical example is to verify if the applications are supporting a business process. In case not, it is an indication of applications not aligned with business functions directly and therefore, should be checked and justified (BOUCHER et al., 2011).
C_12.2	Application-information alignment	We also can verify the alignment between application and information layers as performed in SIMONIN; Le Traon; JÉZÉQUEL (2007) .
C_13	EA evolution	We grouped in this category works which aimed to analyze changes occurred in the EA over time. It represents a time series perspective for EANA (DREYFUS; IYER, 2008; BOUCHER et al., 2011).

Continued on next page

Table 5.6: Categories of EANA concerns mapped from the EANA SLR

Table 5.6 – continued from previous page

Cat_ID	Concern Category	Description
C_14	EA gap analysis	The goal of this analysis is to verify the transformation or change steps needed to go from the AS-IS state to the TO-BE state of an EA(POSTINA; SECHYN; STEFFENS, 2009).

Table 5.6: Categories of EANA concerns mapped from the EANA SLR

To sum up, from the 29 primary studies, we identified 67 network- related metrics, methods, models used to investigate 25 different concerns, which in turn, were classified in the 14 concern categories shown in Table 5.6. It is worth clarifying that a given metric, method, model (e.g. degree centrality), can be used to analyze different concerns (e.g. application dependence or business information criticality) depending on the targeted EA layers and the adopted semantics for components and relationships. At the opposite side, a concern can be analyzed using one or more network-based metrics, methods, models. We had those 67 metrics/methods/models applied in 74 analysis cases (i.e. metrics were reused in different NAIs). An excerpt of the mapping between concerns and their respective metrics is shown in next section. The full data of the identified NAIs can be found in our web repository at www.eanaresearch.org.

5.1.2.3 NAIs mapped

In Table 5.7, we present an excerpt with seven columns of our data extraction form shown in section 5.1.1. Each line in the Table 5.7 represents a NAI. However, due to readability reasons, we only show 13 NAIs from a total of 74 mapped. In the first two columns, we describe the name of the NAI and its analysis concern category. Next, we classify each NAI in three categories: applied in empirical cases (primary data), simulated ones (used illustration examples) or theoretically studied (no data). We use the concepts of EA layers from TOGAF (TOGAF) to classify NAIs according to the EA layers on which they were applied: business architecture (BA), data/information architecture (DA), application architecture (AA), and technology/infrastructure architecture (IA). The motivation extension of Archimate was also considered as the value layer. The level of network analysis and data source also are defined in Table 5.7 for each NAI. Finally, the last column of Table 5.7 indicates the publication source from where the NAI was extracted.

Name of NAI	Analysis concern category	EA layer(s) targeted	NAI-ID	Level of network analysis	Data source	Primary study
Degree of informational dependency	Information dependence among applications	*Application, Information*	M_01	*Component level*	Primary data	WENDT; BRIGL; WINTER (2005)
Degree of functional dependence	Information dependence among applications	*Application, Business*	M_02	*Component level*	Primary data	WENDT; BRIGL; WINTER (2005)
Degree of heterogeneity of the communication paths in the application layer	Information dependence among applications	*Application, Information*	M_03	*Component level*	Primary data	WENDT; BRIGL; WINTER (2005)
Business Information criticality of an application	Information dependence among applications	*Application, Information*	M_04	*Component level*	No data	ADDICKS (2009)
Static alignment measure	EA application-information alignment	*Application, Information*	M_05	*Ccomponent level*	Primary data	SIMONIN; Le Traon; JÉZÉQUEL (2007)
Key components with Centrality degree, Closeness, Betweenness	Key structural components	*Application*	M_06	*Component level*	Primary data	HOMMES (2008)

Continued on next page

Table 5.7: An excerpt of the NAIs identified during the EANA SLR

Table 5.7 – continued from previous page

Name of NAI	Analysis concern category	EA layer(s) targeted	NAI-ID	Level of network analysis	Data source	Primary study
Modularity based on Whitney Index	EA Modularity	*Application*	M_07	*Network level*	Primary data	HOMMES (2008)
Change Cost	EA Change	*Application*	M_08	*Network level*	Primary data	HOMMES (2008)
Stakeholder Crosswalk	Stakeholder management	*Business*	M_09	*Network/ Modules/ Component Level*	Primary data	WOOD et al. (2013)
Degree of connectedness	EA Complexity	*Application*	M_10	*Network level*	No data	MAZHELIS et al. (2006)
Time and intensity between connections	EA complexity	*Application*	M_11	*Component level*	No data	MAZHELIS et al. (2006)
List of Complex Adaptive Systems features	EA complexity	*Application*	M_12	*Network level*	No data	MAZHELIS et al. (2006)
Coupling (operationalized by closeness centrality)	EA cost of change	*Application*	M_13	*Component level*	Primary data	DREYFUS; WYNER (2011)

Table 5.7: An excerpt of the NAIs identified during the EANA SLR

We identified that 63 of 74 NAIs have AA as the input for their analysis process. This partially corroborates Schneider et al.(2015) when they say that "EAM research focuses on application landscapes only, which does not satisfy the premise of a holistic scope of EA."

Nevertheless, we did find analysis of application domains in association with others domains (e.g. WENDT; BRIGL; WINTER (2005); DREYFUS; WYNER (2011); AIER; SCHÖNHERR (2006); ADDICKS (2009)). The majority of the measures found have static lens to analyze EA since few measures were proposed to evaluate the dynamic aspects of EA. According to BUCKL et al. (2008), such dynamic analysis is a widely-known and accepted fact in many management disciplines, but regarding EANA, many considerations have not yet been undertaken. We also corroborate this position here.

Several measures identified in this study were not validated with empirical data: 10 of 29 papers (around 34%) contain an illustrated application of their measures; 5 of 29 articles (around 17%) proposed their measures only theoretically. We also classified each measure according to the structural level of analysis on which it was applied: component level, group and network level.

Some measures (e.g. closeness centrality) appeared more than once, in different NAIs, because they were applied in different contexts analyzing different constructs as suggested by SIMON; FISCHBACH (2013). Most of the measures were designed for the component level (37 of 74 NAIs). It is clear that major attention was dedicated to the EA application domain while social aspects of EA and infrastructure domains still have been focused on less by researchers. Also, we must comment on some modeling issues we identified:

Granularity level of measure analysis: In general, most measures were applied at a single granularity level of analysis without comparing the results of measures at different levels of the system. The proper level of granularity used in the modeling is an important issue. CHIRIAC et al. (2011) find that the degree of modularity can vary for the same system when the system is represented at two different levels of granularity. CHIRIAC et al. (2011) also affirm that it still is not clear which level of granularity is correct in terms of making decisions about which architecture would better achieve the benefits of modularity. Although, experiences from the industry confirm that EA models need to remain on an aggregated level instead of modeling very detailed structures SAAT; AIER; GLEICHAUF (2009). These questions about the "right" granularity level on EA modeling still need further understanding.

Indirect and direct dependencies:LAGERSTRÖM et al. (2014) highlight the importance of indirect dependencies which were important in estimations of change cost propagation. Although, indirect dependencies do generate an additional modeling effort.

Symmetric and asymmetric relationships (or directed and undirected): Some studies did not distinguish between symmetric and asymmetric relationships. Studies should clarify these questions in their modeling process

The full list of NAIs can be found on our website www.eanaresearch.org.

5.1.2.4 Data collection methods

In the primary studies described in Table 5.8, the data were collected commonly through interviews with IT staff and other stakeholders. In those cases, surveys or conversations could be set in order to identify the EA components and their relations. A second approach was the document analysis. In those cases, researchers had access to files containing UML or business process diagrams, for instance, describing components and relationships, which sometimes, could need to be converted for a specific standard or modeling language before analyzed. Interviews and document analysis are often time-consuming methods during the collection process (it was, at least in our experience working with empirical data) and need careful revision during the model-building phase.

In a few cases, researchers have benefited from the support of an automated data collection (e.g a tool which could identify components and automatically extract their relationships) or a EA modeling tool (e.g when companies already use EA modeling tools such ADO-IT). Likely as a consequence, many studies applied illustrative data, as showed in Table 5.8.

Method found	Primary studies
Interviews with IT staff	FÜRSTENAU (2015); LANGERMEIER; SAAD; BAUER (2014); DREYFUS; WYNER (2011); SCHNEIDER et al. (2015); MOCKER (2009); WOOD et al. (2013); POSTINA; SECHYN; STEFFENS (2009); ADDICKS (2009) .
Repository/document analysis	SCOTT (1992); LAGERSTRÖM et al. (2014); FÜRSTENAU (2015); LAGERSTRÖM et al. (2014); DREYFUS; IYER (2008); DREYFUS; WYNER (2011); SIMON; FISCHBACH (2013); SCHNEIDER et al. (2015); MOCKER (2009); WOOD et al. (2013); BARTOLOMEI et al. (2012); POSTINA; SECHYN; STEFFENS (2009); KREIMEYER; GÜRTLER; LINDEMANN (2008); LI et al. (2007); BOUCHER et al. (2011).
Automated collection	POSTINA; SECHYN; STEFFENS (2009); RAMOS et al. (2014)
Simulation/iIllustrative data	FÜRSTENAU (2015); DREYFUS; IYER (2008); MAZHELIS et al. (2006); AIER; SCHÖNHERR (2006); SINGH; SINGH (2010); LEE et al. (2014); CHIRIAC et al. (2011); SIMONIN; Le Traon; JÉZÉQUEL (2007); WENDT; BRIGL; WINTER (2005); NARANJO; SÁNCHEZ; VILLALOBOS (2014); LEHNERT et al. (2015); ADDICKS (2009)
No-data	SINGH; SINGH (2010); BEETZ; KOLBE (2011)

Table 5.8: Data collection methods used by primary studies of the EANA SLR

5.1.2.5 Benefits of the EANA

For each primary study, we searched for explicit mentions made by the authors regarding benefits of EANA. The evidences are organized in Table 5.9.

Effects/Benefits category	Primary studies	Evidences
Improve stakeholder management/governance support	(WOOD et al., 2013)	“allow them to identify which stakeholders are important for any particular decision at any time throughout the program’s development".(WOOD et al., 2013, p 256) "a map that identifies paths of influence for each particular stakeholder by graphically portraying which other stakeholders influence the target stakeholder" (WOOD et al., 2013, p 256)
Identification of architectural improvements to be made (e.g. improve reuse of components, reduce dependencies, refactoring)	(SIMONIN; Le Traon; JÉZÉQUEL, 2007; LEE et al., 2014; LI et al., 2007; RAMOS et al., 2014; LAGERSTRÖM et al., 2014; BARTOLOMEI et al., 2012; HOMMES, 2008))	“Therefore, these case studies further proved that the centrality measures were good ways to identify areas for architecture improvements” (HOMMES, 2008) “Then the manager can easily know which IT applications are not aligned with the business, improving the Business IT-alignment”. (RAMOS et al., 2014, p 446) “The wing and tail connectors and the batteries subsystems were identified as opportunities for flexibility in the system” (BARTOLOMEI et al., 2012, p 58)

Continued on next page

Table 5.9: Benefits pointed out by authors of primary studies of the EANA SLR

Table 5.9 – continued from previous page

Effects/Benefits category	Primary studies	Evidence
Improve communication among stakeholders	(POSTINA; SECHYN; STEFFENS, 2009; DREYFUS; IYER, 2008)	“According to the chief architect at FinServ, the conceptual model, network visualization, and associated simulation is “a powerful communication tool to gain the support of people to do the right thing.” (DREYFUS; IYER, 2008, p 125) “IT architects were able to gain a better idea of possible next steps within these development projects and also used our approach to communicating these steps to their contractors” (POSTINA; SECHYN; STEFFENS, 2009, p 279).
Identification of critical components in terms of structure and cost	(LAGERSTRÖM et al., 2014; MOCKER, 2009; BARTOLOMEI et al., 2012; SIMON; FISCHBACH, 2013; KREIMEYER; GÜRTLER; LINDEMANN, 2008)	"We find that the classification of applications in the architecture (as being in the core or the periphery) is significantly correlated with architectural flexibility (of the costs of architectural change)." (MOCKER, 2009, p 4) “(. . .) systematically point an experienced user to possible weak spots that might otherwise be overlooked. (KREIMEYER; GÜRTLER; LINDEMANN, 2008, p 439)

Continued on next page

Table 5.9: Benefits pointed out by authors of primary studies of the EANA SLR

Table 5.9 – continued from previous page

Effects/Benefits category	**Primary studies**	**Evidence**
Manage evolution of EA	(HOMMES, 2008; SIMON; FISCHBACH, 2013; DREYFUS; IYER, 2008; AIER; SCHÖNHERR, 2006; POSTINA; SECHYN; STEFFENS, 2009)	“It was possible to set architecturally informed rules that guide EA emergence” (DREYFUS; IYER, 2008, p 9) “Nevertheless, the closeness values can be used to compare the modularity of the same system before and after an architecture improvement, and hence can be a useful indicator for system architects” (HOMMES, 2008, 8).
Identification of implicit domains in EA	(AIER; SCHÖNHERR, 2006)	"(...) encapsulated service domains(...) may be desirable for a variety of reasons—technology, business requirements, and also politics. Eventually, encapsulation leads to a better manageable complexity of enterprise architecture" (AIER; SCHÖNHERR, 2006, p 199).
Support EA documentation process	(SIMON; FISCHBACH, 2013)	"As for documentation, they can support the check for EA involvement in domain initiatives, and also the review of documentation in terms of its quality” (SIMON; FISCHBACH, 2013, p 31). “Network metrics thus also help resolve questions of data quality and aid in reviewing architectural documentation in terms of its correctness and completeness”. (SIMON; FISCHBACH, 2013, p 30)

Continued on next page

Table 5.9: Benefits pointed out by authors of primary studies of the EANA SLR

Table 5.9 – continued from previous page

Effects/Benefits category	Primary studies	Evidence
Help in the implementation phase of EA lifecycle	(SIMON; FISCHBACH, 2013)	"The support of architectural relevance checks also makes network metrics relevant in the implementation phase" (SIMON; FISCHBACH, 2013, p 31)

Table 5.9: Benefits pointed out by authors of primary studies of the EANA SLR

5.1.2.6 An attempt of research agenda for EA analysis

In this section, we provide some recommendations for the research community, listing suggestions for future works extracted from the primary studies. We organize them in the following ten topics together with their sources.

Abstraction techniques for modeling components and relations. This topic focuses on the definition of techniques in order to reduce the complexity of the EA analysis process such as better strategies for data abstraction (e.g. how much data must be collected and at which level of granularity?) and the development of criteria for when and how systems should be modeled are worthwhile subjects for research CHIRIAC et al. (2011); BARTOLOMEI et al. (2012).

EA analysis at different levels of granularity. Also related to the previous topic, the inclusion of various mixed levels of granularity of the system architecture representation to analyze the sensitivity of network metrics to those choices would also be valuable future work (SIMON; FISCHBACH, 2013; MOCKER, 2009; CHIRIAC et al., 2011).

Relationship types, weights and symmetry. Differentiation between business and infrastructure applications and its representation in separate networks is considered reasonable by SIMON; FISCHBACH (2013). Analogously, the application and technology architectures analyzed separately could also benefit from this approach. Centrality measures, for example, applied to bimodal networks present different values when compared to unimodal representations. In addition, one could quantify the edges, for example, in terms of frequency of data exchange among applications SIMON; FISCHBACH (2013) and explore how the analysis methods apply to symmetric and asymmetric dependency EA networks (MOCKER, 2009; FÜRSTENAU, 2015).

Qualitative aspects in EA analysis measures. According to POSTINA; SECHYN; STEFFENS (2009), research should be focused on finding and classifying more qualitative measures regarding nonfunctional aspects of application landscapes in general. Strategic value can be added to a component as an attribute, for example. Validation of network measures in more cases and industries. Empirical validation of the measures is claimed explicitly in several works such as LAGERSTRöM et al. (2014); DREYFUS; WYNER (2011); MAZHELIS et al. (2006); MOCKER (2009); AIER; SCHÖNHERR (2006); LEE et al. (2014); NARANJO; SÁNCHEZ; VILLALOBOS (2014); LEHNERT et al. (2015).

Correlate network measures with performance measures like the cost of changes, business agility, etc. This topic was also approached by LANKHORST (2013); LAGERSTRÖM et al. (2014); POSTINA; SECHYN; STEFFENS (2009).

Correlation among network measures. For SINGH; SINGH (2010), another interesting area could be to explore if and how various metrics are related to other metrics. A special case of correlation is suggested in MAZHELIS et al. (2006) to investigate if a network measure Y1 can be expressed as a function of component measures X1, X2, . . . , Xn.

Analysis of EA evolution. Not only to perform gap analysis between AS-IS and TO-BE states of EA, for POSTINA; SECHYN; STEFFENS (2009) long-term studies show great promise for using distance measuring for cost prediction and complexity management (SCHNEIDER et al., 2015; FÜRSTENAU, 2015) on application landscape development. Another avenue in this direction is to examine whether an IS architecture is in a phase of expansion or consolidation (FÜRSTENAU, 2015) and Boucher et al. (2010).

More topological/structural analysis. Some approaches might take full advantage of the several topological properties of enterprise models seen as networks, such as discovery of paths, clusters, or graph metrics (NARANJO; SÁNCHEZ; VILLALOBOS, 2014).

Tool development support for EA analysis. The need is pointed out by LAGERSTRÖM et al. (2014); RAMOS et al. (2014); WENDT; BRIGL; WINTER (2005). According to LAGERSTRöM et al. (2014), future work needs to be directed towards data collection support in the enterprise architecture domain. In LEVINA; HILLMANN (2012), authors suggest experiments on the perception of EA components visualization as a network using the cognitive fit theory. Regarding EA analysis supported by tools, in RAMOS et al. (2014) they have an extendable tool to perform analysis over EA models. The EA field may also benefit from recent development of tools in areas such process mining (AALST, 2016) and data science (POWER, 2016), which are making data from information system executions available for analysis. However, these areas were out of our research scope.

5.1.2.7 Conclusions

We conducted an extensive literature review and identified 67 network-based metrics and methods from 29 analyzed articles with network measures applied in EA analysis. As an emergent field, we saw several analysis concerns (classified in 14 categories) aimed in the papers operationalized by different measures with most of them still requiring validation. Although, we consider the present literature review as an important step for the field since it contributes towards systematizing what is known about the application of network measures in EA analysis.

Finally, we hope that the identified measures list can foster a discussion among practitioners and consequently make the analysis process of EA analysis easier. We shed light on several research gaps and hope to attract the attention of the EA research community.

In the next section, we elaborate on the findings to design two important artifacts for EANA.

5.2 The EANA meta-model

While we were performing the SLR about EANA, a problem that we have identified is that NAIs often lack an explicit specification of the information and structures that they require to work. Thus, it is not easy to realize if an existent implemented EA model is suited or not to support a particular analysis function (RAMOS et al., 2014) or from where to start to collect EA data to perform a specific kind of analysis. We also realized that a common foundation of concepts - a sort of taxonomy- for the field was lacking. For instance, it was very common not find the semantics of relationships between EA components, which in turn, make it difficult for researchers to compare their results since it is hard to detect the similarities among the methods employed.

To address this issue, we designed a meta-model to characterize NAIs. The meta-model was based on the 74 NAIs found in our previous review (Santana A.; Fischbach K.; Moura H., 2016a). Thus, we classify the initiatives according to their analysis concern and other information requirements, designing thus, their fundamental constructs. Our approach is inspired by the work of RAMOS et al. (2014), which defined for each analysis initiative the information requirements need to perform it. In our case, we needed to include specific information requirements related to structural analysis issues: level of network analysis, uni/multi-modal character, types of outputs, derived data, and structural properties of relations. In doing so, we provide answers to the following research questions:

- What are the information requirements to perform EANA (RQ2)?
- How can NAIs be classified (RQ3)?

5.2.1 Overview of the artifact's design

As discussed in chapter 3, this piece of research can be classified as an exploratory, descriptive and basic research (WOHLIN; AURUM, 2015). As for research methods, we have adopted the design science research method (HEVNER; MARCH; RAM, 2008). The six steps of the DSRM are described in Table 5.10:

For SIMON (1996), one of the goals of DSR artifacts is to motivate the activity or the use of the artifact in the targeted environment, which in turn, may generate new research goals and artifacts. We hope to contribute with the proposed EANA-MM to the academy and industry

DSR phase	Research contextualization
Identify the problem	For the researchers and practitioners audience, the NAIs (e.g. metrics or methods) often lack an explicit specification of the information and structures that they require to work. It is difficult for researchers to compare their results since it is hard to detect the similarities among the methods employed. It is not easy to realize if existent implemented EA models in the organizations are suited or not to support a particular analysis function in practice.
Define solution objectives	Therefore, we aim to help EA researchers and practitioners, defining for each analysis initiative the information requirements needed to perform it, as taxonomy for the NAIs.
Design and development	We learned the design commonalities shared by the 74 NAIs found in our SLR, combined with the constructs of network science to design an EANA meta-model.
Demonstration	We demonstrate how the EANA-MM covers all the NAIs found in the literature.
Evaluation	We evaluate our artifact with the artifact instantiation with the 74 NAIs and surveying EA experts about the EANA-MM.
Communication	We make the research exposition itself in this thesis (to the academic audience) and later in publications outlets to be produced.

Table 5.10: DSR phases and contextualization for the EANA-MM (HEVNER; MARCH; RAM, 2008).

as we establish a common conceptual ground with fundamental elements that may serve as the first step to designing and apply new network-based analysis initiatives.

5.2.2 The EANA-MM artifact

The EANA meta-model, our first DSR artifact, is depicted in Figure 5.4. Its main constructs are discussed in the following.

- Network analysis Initiative (NAI)

As already discussed, we consider as an **NAI** each attempt to analyze EA modeled as a network. This can be a simple, well-known network metric, more sophisticated techniques or methods or yet analysis inspired by network models applied to analyze a specific concern. For example, in SIMON; FISCHBACH (2013), the degree centrality (a network metric) was used to analyze the application usage (the analysis concern) in the application layer. A more sophisticated approach, the "hidden structure method", is proposed by LAGERSTRÖM et al. (2014), as a method (not a metric) which aims to identify central structures and peripheral ones in an application layer. Both are examples of an NAI which differ clearly in terms of computational complexity.

One NAI aims to answer an **analysis question**, which in turn is associated with an **analysis concern**, which might be particularly important for one or more **stakeholders**.

To execute a NAI, we need to define a **data generation strategy** to get or generate (e.g. data sources and data collection methods) to search for (or to build) the proper EA model.

In order to build this EA model, it is necessary to define the **semantic of layers, components their respective relations**, which may change over **time**. Before to start analyzing the EA model, we also need to define on which **network level** this analysis will take place (e.g. at component level or modular level). Finally, each NAI produces at least one type of **output** (e.g. graph visualizations, rankings of components). All those conceptual elements are discussed as follows:

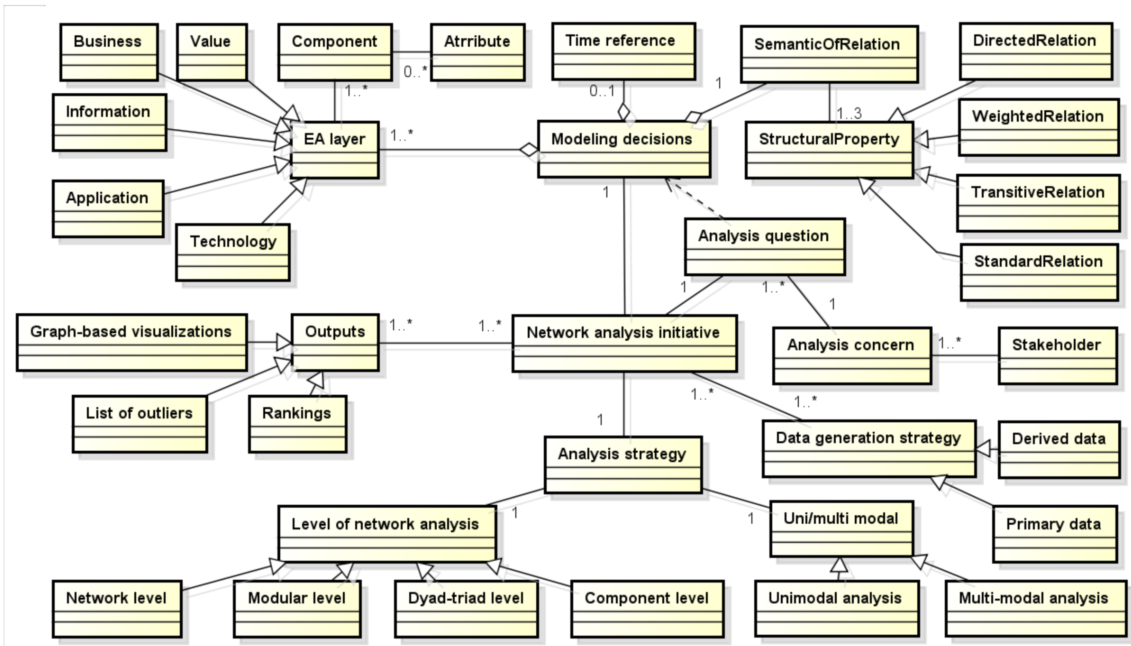

Figure 5.4: Proposed EANA-MM

- Analysis concern

Also previously discussed, it can be defined as an aspect of the EA components or of the EA as whole, that is investigated/measured by the experts in order to help EA evolution, to communicate with stakeholders and/or to ensure business-IT alignment. Each NAI has an associated concern such as to identify key structural components, to measure the EA layer modularity, analysis of the cost of a change and so on (see Table 5.6, section 5.1.2.2 for the concern categories mapped).

- Analysis question

For each NAI, we elicited the analysis aspect proposed by the researchers in their primary studies and describe it as an operational question. Such questions are based on our interpretation of the NAI and its outputted results. Examples of questions are: “what are the key structural components in the EA?" , How modularized is the application or business layer?", "what are the cost of changes if we decided to change a particular component?". One question can be answered by one or more correlated NAIs which aim to answer it.

- Data generation strategy

In this construct, we document the process of data collection or generation which are inputs for the NAIs. The first important aspect is the type of used data source: primary, derived or simulated. **Primary data** is related to approaches used to generate EA models from scratch, using documents modeled with specific notations such as UML or Archimate; interviews with architects and stakeholders or automated collection of EA data. For instance, in SIMONIN; Le Traon; JÉZÉQUEL (2007), UML diagrams are collected to analyze the static alignment between the application and information layers. In WOOD et al. (2013), textual documents, diagrams, and interviews are conjugated to build the dataset. **Derived data** uses primary data and some derivation mechanisms to create artificial edges between components and build derived networks of EA data as discussed in chapter 2. The third strategy is to use simulated data to demonstrate the NAIs as done in LEE et al. (2014). There are still some works which presented their NAIs and, however, did not realiz any simulation or empirical validation with them as in BEETZ; KOLBE (2011) and Boucher(2010).

- Modeling decisions

Another important aspect related to the EANA is the modeling choices for the EA model. In this sense, it is important for the analysis to define which **EA layers** it will approach, the **semantics** for the components themselves, the components´ relationships and some **structural properties** of those relations such as if these relations are directed, transitive and/or weighted as discussed in 2.2.8. In addition, **attributes** of EA components (e.g. cost, strategic value) can also be taken into consideration. Finally, the modeling may also involve **time-related issues** (e.g. evolution of the EA).

- Analysis strategy

With this construct, we define macro-level analysis decisions such as at which network level the analysis is performed. The options are at component level, dyad-triad level, group level and network level as discussed in section 2.2.3. Another characterization made regards to the EA layers used: one or more EA layers can be used by each NAI. For instance, in RAMOS et al. (2014), the business and application layers are used to measure their alignment to each other. Finally, we must define for each of those network levels of analysis, if unimodal, bimodal or multi-modal networks will need to be built .

- Outputs generated

This element describes what kinds of outputs were used to represent the results of the NAIs. Despite not being an exhaustive list, we identified papers producing outliers´rankings as in WENDT; BRIGL; WINTER (2005), list of numerical data as in MAZHELIS et al. (2006), graphs WOOD et al. (2013) and clusters BARTOLOMEI et al. (2012). The outputs' types were presented in section 2.2.5.

- Stakeholder

A stakeholder is an individual, team, or organization (or classes thereof) with interest in the EA analysis and which might benefit from a NAI´s output (e.g. software architect, process owner, business manager, CIO and so on).

Finally, in Table 5.11, we present a brief consolidation of the EANA-MM constructs presented in this thesis.

EANA-MM element	Description
NAI	Each attempt to analyze EA modeled as a network. This can be a simple, well-known network measure, more sophisticated techniques or methods or analysis inspired by network models
Analysis concern	An aspect of the EA components or the EA as whole that is investigated/measured by the experts in order to help EA evolution, to communicate with stakeholders and/or to ensure business-IT alignment.
Modeling decisions	Decisions such as EA layers and components approached, edges' semantics, edges' weight, direction or transitivity.
Analysis strategy	It is defined by the network level of analysis and the unimodal/bi/multi-modal character.
Uni/bi/multi-modal analysis	Unimodal analysis uses a network composed of just of type of component (e.g. application). Bi-modal analysis uses a network composed of two types of components (e.g. application and business process). Multi-modal analysis uses a network composed of more than two types of components (e.g. technology, application and business process).
Network level of analysis	The chosen structural level to analyze the network: component, dyad-triad, module or group and network level (overall network).

Continued on next page

Table 5.11: Summary of the EANA conceptual elements

Table 5.11 – continued from previous page

EANA-MM element	Description
Data generation strategy – Primary data	It is related to approaches used to generate EA models from scratch, using documents, interviews or automated collection of EA data
Data generation strategy – Derived data	This approach uses primary data and some derivation mechanisms to create artificial edges between components and build derived networks of EA data.
Outputs	Types of outputs used to represent the results of the NAIs such as graphs, rankings or raw numerical values.
Stakeholder	The user or stakeholder that might benefit from the results generated by the network analysis initiatives.

Table 5.11: Summary of the EANA conceptual elements

CHIRIAC et al. (2011) apud ANABY-TAVOR et al. (2010) affirm that the conceptual modeling is not an end in itself. Conceptual models are created as an intermediary goal, to serve to additional final goals. In this sense, we prospect some applications for the proposed meta-model:

- Help in the systematization of the knowledge produced so far regarding EANA.
- Help researchers to easier discuss their findings and trace comparisons among studies with similar analysis strategies or data generation strategies, for example.
- Establish a common language for researchers, tool designers and experts in EA.

Having the EANA-MM as a reference, researchers and practitioners know in advance the exact information requirements to perform EANA for specific analysis concerns. In this direction, the EANA-MM may help in the development of new NAIs.

In the next section, we advance our understanding of the application of NAIs in practice, with the design of the EANA library.

5.3 The GQM-based EANA library

Creating a taxonomy to describe the analysis initiative was our first step to help experts to perform EA structural analysis or even inspire them to create their own ones. In our second step, we organized the analysis initiatives in a decision structure to help experts to select the right measures for an intended specific case. To do so, we built a library based on the goal-question-metrics concept (SOLINGEN; BERGHOUT, 1999). Our hypothesis is that the expert can use the GQM based library to identify which analysis initiatives (and respective information requirements) provide the intended outputs. In doing so, we provide some answers to the research question "How can experts be guided to perform EANA?" (RQ4). Our second DSR artifact is designated for organizations and EA experts involved in the EA analysis.

In the following, we contextualize the EANA library design in the DSRM phases.

5.3.1 *Overview of the artifact's design*

The design process of the EANA library can be classified as an exploratory, descriptive and basic research. We again adopted the design science research method (HEVNER; MARCH; RAM, 2008). The six steps of the methods are described in Table 5.12:

DSR phase	Research contextualization
Identify the problem	There is no consolidation about NAIs produced so far. Also, there is a lack of a mechanism to help EA experts choose which NAI to select, according to a specific analysis concern.
Define solution objectives	To guide EA experts to perform structural analysis using the 74 NAIs identified in our EANA SLR
Design and development	We use the 74 NAIs identified in our SLR, together with the ideas of GQM models (SOLINGEN; BERGHOUT, 1999) to provide a decision mechanism to help EA experts to choose an NAI.
Demonstration	We build the mappings between each analysis concern to its respective network-based metrics, both pointed out in the primary studies. We organize this knowledge in a web portal.
Evaluation	We evaluate our artifact reasoning about its application and asking EA experts about its utility.
Communication	We make the research exposition itself in this thesis (to the academic audience). We also publish the EANA library in the web portal www.eanaresearch.org.

Table 5.12: DSR phases and contextualization for the EANA library (HEVNER; MARCH; RAM, 2008).

From the 29 primary studies, we mapped the 74 NAIs with the 14 categories of concerns. This process is depicted in Figure 5.5.

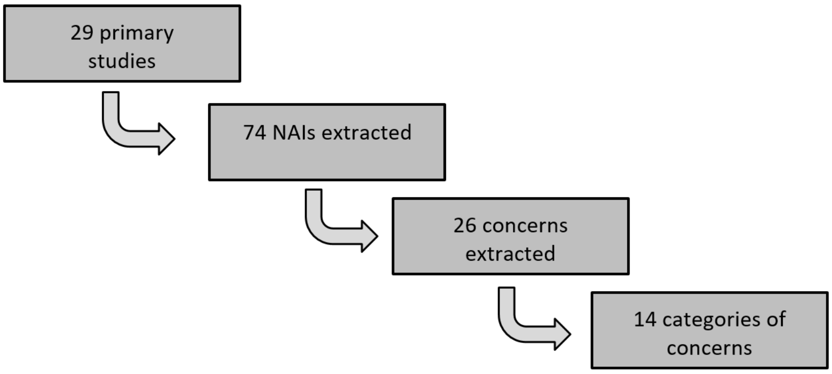

Figure 5.5: NAI library building process

We linked each NAI and its concern using principles of the goal-question-metric model (SOLINGEN; BERGHOUT, 1999). The idea is to help practitioners to operationalize the utilization of those NAIs to help them to perform EA structural analysis. The goal concept from SOLINGEN; BERGHOUT (1999) is translated to our context to "analysis concern". Similarly, a "question" as defined by SOLINGEN; BERGHOUT (1999) is equivalent in our context to an "analysis question". Finally, the "metric" concept from the GQM model is mapped to an NAI in our EANA library. In Figure 5.6, we present the implicit decision support mechanism for each NAI from the GQM-based EANA library.

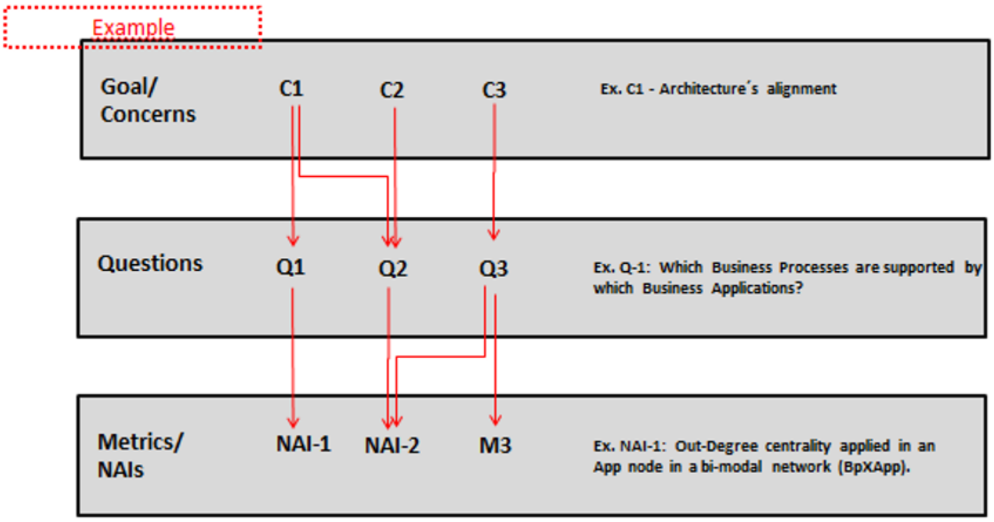

Figure 5.6: The decision mechanism of the GQM-based EANA library

As indicated in Figure 5.6, if the EA expert wants to check the architecture alignment (the analysis concern), he will find the questions "Q1" and "Q2" in the library which are implemented by the NAI-1 or NAI-2. We believe that such structure, together with the EANA meta-model, help researchers and practitioners in selecting an NAI according to a concern of interest. Once the NAI is selected, the data stored in the EANA library, in conformance with the EANA-MM, specifies which analysis strategy and data strategy to apply; identify stakeholders which can benefit and the modeling decisions needed to build the model and so on (these are the conceptual elements from the EANA meta-model).

In the next section, we described all the mappings between analysis concerns and NAIs.

5.3.2 *The GQM-based EANA-library artifact*

Table 5.13 describes the concerns' categories, concerns, operational questions and the NAIs to compute them. It has to be noted that one metric or method (the right column in Table 5.13) can be utilized by one or more NAIs. For instance, the metric "M_41" is used by the NAI which targeted the "EA application Complexity" (concern C_12) and also by the NAI which targeted the concern "Business-Application Alignment" (C_23). That is the reason why "M_41" appears twice in Table 5.13.

Category ID	**Concern ID**	**Concern**	**Question ID**	**Question**	**NAIs identified**
1	C_01	Information dependence of an application	Q_01	What are the most dependent applications in terms of information?	M_01
			Q_02	What are the applications which depend on information from different enterprise domains?	M_02
	C_02	Reuse of an application	Q_03	What is the degree of reuse of the application?	M_03
	C_03	Business criticality of an application in terms of information	Q_04	What is the level of business criticality of an app in terms of data usage?	M_04

Continued on next page

Table 5.13: The GQM-based EANA library

Table 5.13 – continued from previous page

Category ID	**Concern ID**	**Concern**	**Question ID**	**Question**	**NAIs identified**
2	C_04	Key structural components	Q_06	What are the key structural components of the EA?	M_06; M_18; M_19; M_56; M_54; M_55
3	C_05	EA Modularity	Q_07	How well the application/system architecture is modularized?	M_07; M_20; M_45; M_46
4	C_06	EA Change	Q_08	How to measure the propagation effect of a change in EA?	M_08; M_21; M_25; M_44
	C_07	EA cost of a change	Q_13	How to measure the flexibility of the application layer based on the cost of change of its components?	M_13
			Q_49	How can we identify the applications with the highest cost of change?	M_51
			Q_62	What is the propagation cost of a change in EA?	M_66

Continued on next page

Table 5.13: The GQM-based EANA library

Table 5.13 – continued from previous page

Category ID	Concern ID	Concern	Question ID	Question	NAIs identified
4	C_08	EA Change management decision-making	Q_55	How to generate an order of prioritization for EA changes implementation based on process clustering?	M_59
5	C_09	Stakeholder management	Q_09	How to identify the stakeholder groups according to EA Domains or along product architecture phases?	M_09
			Q_58	How is stakeholder behavior along the time according to betweenness centrality?	M_62
6	C_10	EA Complexity	Q_10	How to evaluate the network behavior of the EA regarding change over time?	M_10
			Q_11	How intensive the application interactions are?	M_11
			Q_12	How evaluate EA complexity?	M_12; M_26; M_29; M_30;

Continued on next page

Table 5.13: The GQM-based EANA library

Table 5.13 – continued from previous page

Category ID	Concern ID	Concern	Question ID	Question	NAIs identified
			Q_15	What are the components with high complexity regarding execution time?	M_15
			Q_59	What is the average path of the EA?	M_63
			Q_61	How is the cohesion evolution of the EA network?	M_65
			Q_63	What is the architecture flow? (meaning that more than % of the applications are either in, depend on, or are dependent on the Core)	M_66
6	C_11	EA Application Complexity	Q_38	How evaluate EA application complexity?	M_29; M_39; M_41; M_49
			Q_42	How evaluate EA application complexity based on the number of redundant applications?	M_42

Continued on next page

Table 5.13: The GQM-based EANA library

Table 5.13 – continued from previous page

Category ID	**Concern ID**	**Concern**	**Question ID**	**Question**	**NAIs identified**
			Q_43	How to evaluated EA domain complexity based on the number of internal and external connections between applications?	M_43
			Q_44	How to classify the application architecture regarding structure (Core, control, shared, periphery, propagation cost)?	M_44
			Q_47	How evaluate EA application complexity based on the number of in and out connections?	M_49, M_50
6	C_12	EA Business Complexity	Q_39	How evaluate EA business complexity?	M_38, M_49
	C_13	EA Technology Complexity	Q_40	How evaluate EA technology complexity?	M_40
7			Q_14	How to cluster applications based on the amount and frequency of transferred data?	M_14

Continued on next page

Table 5.13: The GQM-based EANA library

Table 5.13 – continued from previous page

Category ID	Concern ID	Concern	Question ID	Question	NAIs identified
7	C_14	EA Domain Analysis	Q_31	what is the degree of purity of domain for the applications' support?	M_31
			Q_43	how to evaluate EA domain complexity based on the number of internal and external connections between application?	M_43
			Q_60	Which domains or modules can be identified in EA?	M_64; M_67
8	C_15	EA improvement	Q_15	What are the components with high complexity regarding execution time?	M_15
	C_16	Business process improvement	Q_33	How can the business process be classified (core, automatable, information intensive, distributed, flexible) regarding information flux in order to facilitate Business Process Management?	M_33; M_34; M_35; M_36; M_37

Continued on next page

Table 5.13: The GQM-based EANA library

Table 5.13 – continued from previous page

Category ID	Concern ID	Concern	Question ID	Question	NAIs identified
8	C_16	Business process improvement	Q_50	How to establish a prioritization order for process improvement based on structural analysis?	M_52
9	C_17	Application usage	Q_16	What is the level of connectivity of the application?	M_16
	C_18	Cross-Functionality	Q_17	Which application has potential cross functionality?	M_17
10	C_19	Risk failure management	Q_18	Which applications can impact the IT landscape in case of failure?	M_18; M_19; M_06
11	C_20	Local dependency	Q_22	What the importance of an application in terms of the number of applications which depend on it?	M_22; M_27
			Q_23	What is the absolute dependence of an application in terms of the number of other applications this application depends on?	M_23; M_28

Continued on next page

Table 5.13: The GQM-based EANA library

Table 5.13 – continued from previous page

Category ID	Concern ID	Concern	Question ID	Question	NAIs identified
11	C_20	Local dependency	Q_24	What is the criticality of an application in terms of in-degree* out-degree product?	M_24
			Q_48	What is the local neighborhood of dependence of an EA component?	M_50
12	C_21	EA alignment	Q_53	How are aligned stakeholder, application, information and process dependencies?	M_57
	C_22	EA Business-IT alignment	Q_32	What is the potential importance of a current AL application for the business?	M_32
			Q_54	Are the applications supporting business process?	M_58
			Q_41	How to evaluate the application complexity based on the number of a business process supported by an application?	M_41

Continued on next page

Table 5.13: The GQM-based EANA library

Table 5.13 – continued from previous page

Category ID	Concern ID	Concern	Question ID	Question	NAIs identified
12	C_23	EA application-information alignment	Q_05	What is the static alignment between the information layer (Business Object) and the application layer?	M_05
13	C_24	EA evolution	Q_45	How to find architectural control points and measure the preservation of their influence during the EA evolution?	M_47
			Q_46	How to measure the preservation of the application network topology over the time?	M_48
			Q_51	How can we monitor the attachment behavior of new applications to existent applications?	M_53
			Q_56	How EA domain structural indicators are evolving over time?	M_60

Continued on next page

Table 5.13: The GQM-based EANA library

Table 5.13 – continued from previous page

Category ID	**Concern** ID	**Concern**	**Question** ID	**Question**	**NAIs identified**
14	C_25	EA Gap Analysis	Q_31	What is the degree of purity of domain for the applications' support?	(AS_IS - TO_BE) M_31
			Q_32	What is the potential importance of a current application layer for the business considered the TO-BE state?	M_32

Table 5.13: The GQM-based EANA library

We expect that the consolidation of the structural analysis expertise found in literature organized in a GQM-based fashion may facilitate the application of NAIs by EA experts in practice and also, in combination with the EANA-MM, may foster the design of new NAIs that can be further added to the library. We recognize this would require a publicization of these results in combination with a tool support that could allow the EA analysis community to have access to this knowledge base in a more dynamic way. We discuss this solution in the next section.

5.4 EANA LIBRARY WEB PLATFORM

We developed a web portal to publicize the EANA library from where the research community can have access to an initial set of NAIs. The web portal also allows other researchers to add new NAIs to the library and contribute to its improvement as well. In the next sections, we present the web portal's composition briefly.

5.4.1 Modules and user profiles

There are three types of users which are detailed in the following.

- Visitor. Any person who does not need to be registered in the system. A visitor has access to public functionalities such as the list of the published NAIs and a query interface in order to generate a catalog of NAIs according to parameters such as concern, targeted EA layer, stakeholder and so on.
- Researcher/Practitioner. In addition to the previous functionalities, this user can add, update and remove NAIs. The information required to add a new NAI are aligned with the EANA-MM elements described in section 5.2. All operations performed by this user are tracked and registered with their credentials.
- Administrator. This user has access to all system functionalities. The registration and management of users and other system configurations are exclusive to this user.

Figure 5.7 depicts the profiles and their respective use cases.

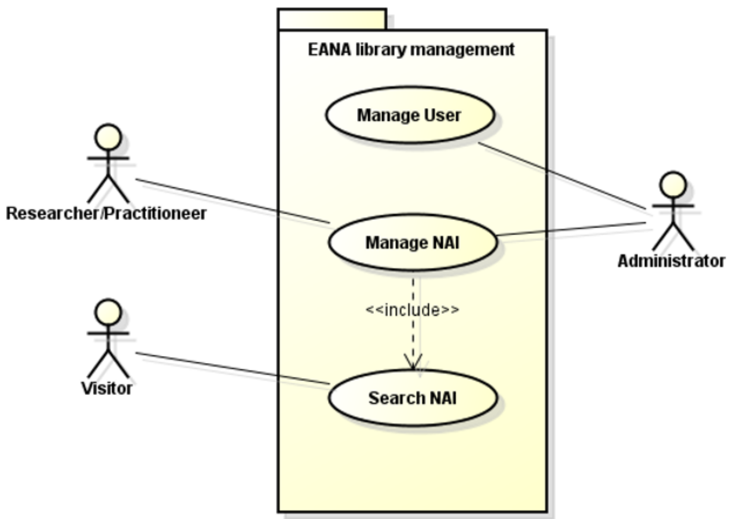

Figure 5.7: Use case diagram for the EANA library web portal

We chose one use case to describe in the next section, while other use cases can be found in Appendix H.

5.4.2 NAI search - use case

A visitor or any other system user that wants to find NAIs can enter keywords in the relative fields and search an NAI´s name as shown in the interface depicted in Figure 5.8:

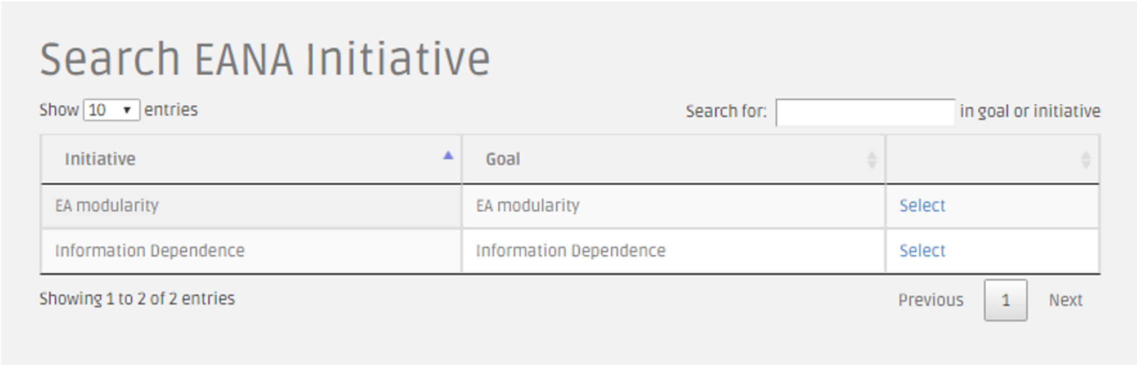

Figure 5.8: NAI search's interface

As a result, the system shows the matched NAIs in each line, which in turn can be selected (in the right column in the result data grid) to be detailed in a second interface depicted in Figure 5.9.

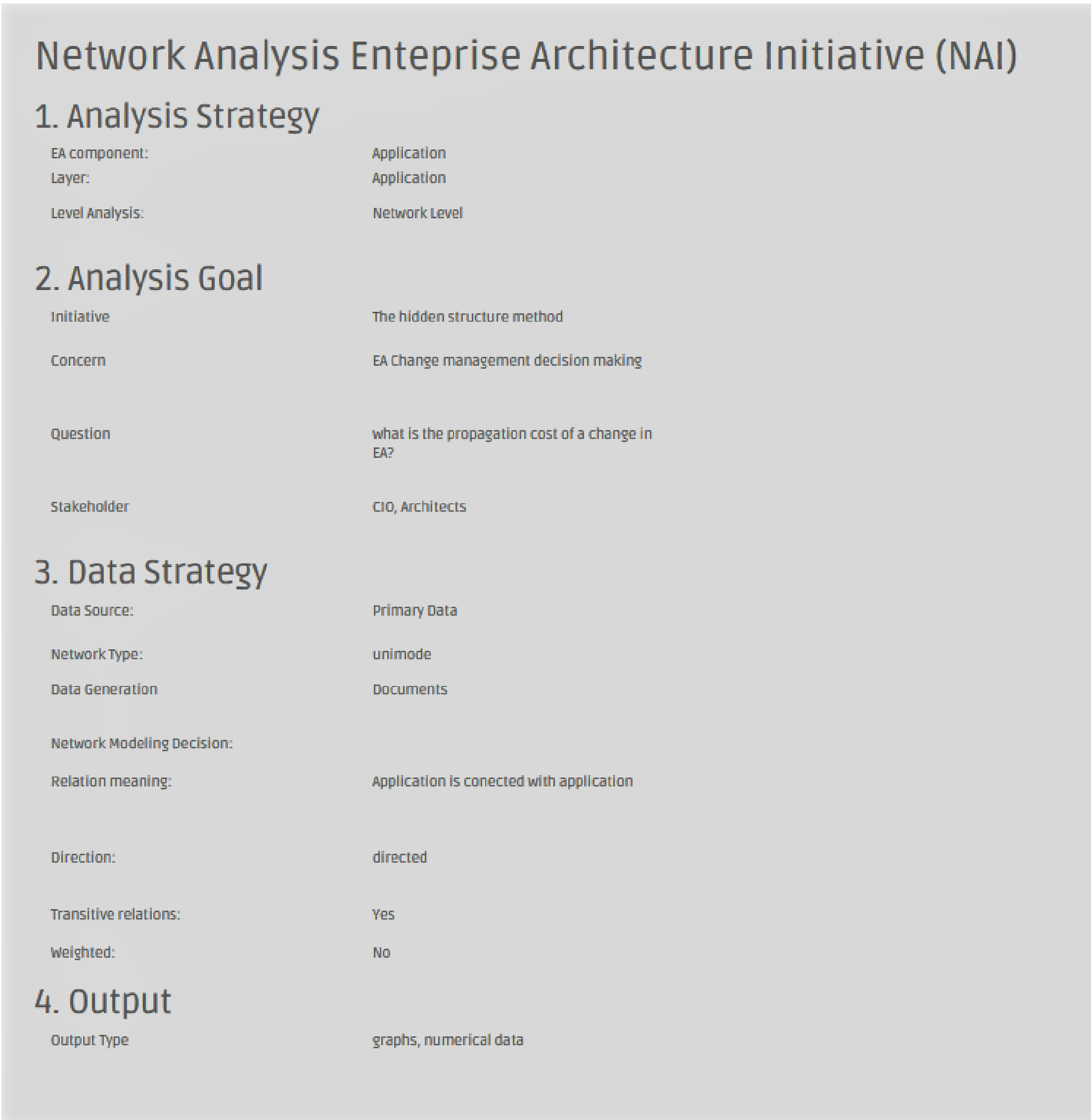

Figure 5.9: Interface showing the details of an NAI's

The EANA library is intended to be a central repository for NAIs accessed by any stakeholders. It can be found at www.eanaresearch.org and will be continuously maintained for research purposes. Further details about the web portal´s architecture can be found in the Appendix H.

5.5 Chapter Summary

In this chapter, we conducted an extensive literature review and identified 74 NAIs from 29 analyzed articles. As an emergent field, we saw several analysis concerns aimed at the papers operationalized by different measures with most of them still requiring validation. Although, we consider the present literature review as an important step for the field since it contributes towards systematizing what is known about the application of network measures in EA analysis. We shed light on several research gaps and hope to attract the attention of the EA research community. These results are also published in Santana A.; Fischbach K.; Moura H. (2016a). Grounded on these previous findings, we extracted from the NAIs the main concepts regarding EANA to build our EANA-MM.

For SIMON (1996), one of the goals of DSR artifacts is to motivate the activity or the use of the artifact in the targeted environment, which in turns, may generate new research goals and artifacts. We hope to contribute with the proposed EANA-MM to the academy and industry as we establish a common conceptual ground with fundamental elements that may serve as the first step to designing and apply new network-based analysis initiatives and to foster the improvement of existing EA analysis tools as in SCHMIDT (2013) and SCHÜTZ; WIDJAJA; GREGORY (2013).

Finally, we structured the NAIs and respective analysis concerns with the GQM model aiming to facilitate the EANA process, reusing existent metrics and methods proposed in the primary studies.

Some limitations can be pointed out. Our document collection does not include publications from 2016. Thus, this might be a minor downside of our review. Second, we did not perform a quality assessment of the primary studies (KITCHENHAM, 2007) since, based on our previous readings, we consider the field as a relatively recent and emergent area with many studies still attempting to validate their initial conclusions. We wanted to analyze papers with different hierarchies of evidence KITCHENHAM (2007) and different methodological designs. Therefore we relaxed our quality criteria to include the papers.

Despite being important steps towards it, the EANA-MM and EANA library are not sufficient regarding the EANA process guidance. Experts performing EANA from scratch (e.g. designing new NAIs) can benefit from the EANA-MM, regarding the design decisions to be made. From a top-down approach, a NAI selected from the EANA library does not specify the order of the activities needed to implement it. In the next chapter, we describe **how the EANA body of knowledge constructed so far could be used by researchers and practitioners** to advance its knowledge base and practice, respectively (RQ4).

6

TOWARDS A PROCESS VIEW FOR EANA

In the previous chapter, important contributions to clarify the conceptual elements involved in the EANA were given. In addition, a set of NAIs extracted from the literature was also presented. In this chapter, we advance in the operationalization of EANA library while we propose a set of EANA acivities.

We start the chapter discussing the lack of guidance in performing EANA, and how the EANA-MM and the EANA library may collaborate in that direction (section 6.1). Next, the design of the artifact is described in section 6.2. A set of ten EANA activities is presented in section 6.3. The two methods which compound the proposed EANA process, their main activities, inputs, outputs and involved actors are presented in section 6.4. Finally, the implications of the proposed artifact are discussed.

6.1 CONTEXT AND MOTIVATION

According to MATHEWS; KAESLIN; RYTZ (2014), the harnessing of the EA data accumulated over the years, documenting different architectural layers including business function, data, application, and technology, often remain unused for business benefit and widely performed manually using expert knowledge. In line with this, a survey amongst enterprise architecture practitioners shows that these practitioners often forego the use of a structured template/approach when rationalizing an architecture, relying instead on ad-hoc information capturing on tools such as Microsoft Office (PLATANIOTIS; De Kinderen; PROPER, 2013). Besides, as discussed in the EA analysis state of art presented in chapter 4, only a small number of papers presented details about how they performed the EA analysis in their studies. Even worse, almost half of the selected studies did not perform any empirical analysis at all.

In this scenario, exceptions are the works of ADDICKS; APPELRATH (2010), LANKHORST (2013) and VENEBERG et al. (2014). In the former, the authors present a generic method (for any organization and any kind of analysis in the application layer) together with its main parameters, for instance, evaluation criteria, metric, application data and key figures (tabular or graphical representation). Related analysis activities such as “configure evaluation” and “data gathering” are also defined. The main contributions of the authors are the generic information meta-model and analysis method, although no specific evaluation criteria, metric or empirical analysis is presented. In LANKHORST (2013), the analysis scope goes beyond the application layer. The efforts are put in developing guidelines to build effective EA models, aligned with the stakeholders´ concerns. Their guidelines are focused on how to perform the modeling activities and visual analysis but without presenting the step-by-step aspects as in ADDICKS; APPELRATH (2010). The work of VENEBERG et al. (2014), advanced on those two previous approaches, enriching EA models with operational data about EA components. The authors also describe the EA Intelligence Life Cycle (EAIL), a method, its phases, which combine business intelligence principles applied to the EA models. In the specific context of EANA, little has been written about how to perform EANA. The only exception is the work of FUERSTENAU; ROTHE (2014), which proposes a procedural method for a specific type of analysis: the identification of IT shadow systems.

Aware of the previous limitations, operationalizing the structural analysis in EA context is the main concern in this current research, since its initial design. It is reflected in the research questions, in the proposed GQM-based library of NAIs and also in the desire of transferring this knowledge to the EA analysis practice. However, one question that still can be posed here is "how can we execute it in practice"? To answer that question, the EANA meta-model concept plays an important role, providing the decision points necessary to design or run any NAI. The EANA-MM and the EANA Library, both might give us the first steps in that direction.

We already made some progress while we propose the GQM-based EANA library which can be used in prescriptive way by the EA experts. Nevertheless, despite being an intuitive top-down approach, the GQM does not provide enough guidance (i.e. a processual view) for the application of NAIs.

Therefore, in this chapter, we aim to fill this gap and propose a process to apply existent NAIs and help to design new ones. We reuse some concepts from ADDICKS; APPELRATH (2010), LANKHORST (2013), VENEBERG et al. (2014) and FUERSTENAU; ROTHE (2014), adding contextual issues of EANA discussed in the EANA-MM, to look closer at the EANA execution, detailing its phases, activities, outputs, involved actors, to compose the base of a process view for EANA. The purpose of this artifact is to help EA researchers and practitioners and their working groups follow an "objective" and systematic process for collecting and analyzing EA data. In the next section, the artifact's methodological aspects are explained.

6.2 EANA PROCESS RESEARCH DESIGN

As discussed in chapter 3, we have adopted the design science research method (HEVNER; MARCH; RAM, 2008). The six steps of the methods are described in Table 6.1.

The empirical instantiation of the artifact is presented in chapter 8. In the following, we present the main activities, inputs, outputs and tools that can support it.

DSR phase	Research contextualization
Identify the problem	In our context or environment, (HEVNER; MARCH; RAM, 2008), organizations have complex IT landscapes and there is a lack of guidelines to perform EANA on them.
Define solution objectives	To help EA researchers and practitioners to perform analysis of EA modeled as a complex network, proposing a set of activities presented in a step-by-step fashion.
Design and development	We use the initial ideas from 4 previous works combined with the EANA-MM and EANA library to design an EANA process. Those are our knowledge base (HEVNER; MARCH; RAM, 2008).
Demonstration	In chapter 8, a proof-of-concept demonstration of a subset of the proposed artifact is described.
Evaluation	We evaluate our artifact with the artifact instantiation.
Communication	We make the research exposition itself in this thesis (to the academic audience) which will be a base for producing a report for the experts (management audience).

Table 6.1: DSR phases and contextualization (HEVNER; MARCH; RAM, 2008).

6.3 The EANA process artifact

The EANA process view consists in ten main activities that guide experts in performing EANA, which are described as follows:

6.3.1 Analysis setup

As proposed in FUERSTENAU; ROTHE (2014), in the first step, a project team is appointed. The staffing should account for two kinds of complementary skills: business analysis competencies are essential throughout the data collection phase to align IT and business perspective; the business analyst acts as a boundary spanner. Team members should also be experienced with IS architectural solution patterns to appraise the data at hand. The project's breadth and depth should be limited, (e.g. line of business, organizational function, EA layer).

In this step, it is essential to ensure that team members have the skills to perform EANA as well and the access to the data sources or existent EA models. One must keep in mind the needed resources to collect missing data. If necessary, a data confidentiality agreement should be signed at this point.

6.3.2 Define an analysis approach

Team members must agree on an analysis concern (see section 5.3.2) to be investigated. The members can decide either to consult the EANA library for concerns options or even create a new concern, according to their current needs. In both cases, a heuristic applied to choose a concern is to define which level of network analysis is intended: component, group/clustering or network level, since they represent very distinct types of analysis as discussed in section 2.2.3.

Once the concern is defined (e.g. cost-analysis estimation, key structural components, application support coverage), members can reuse one of the NAIs from the EANA library specified for it. In this case, the remaining analysis activities would follow the NAI´s requirements specified in the EANA library (analysis strategy, data strategy, etc). If the needed data is not available a priori, team efforts will concentrate on data collection activities in later phases. We label this analysis scenario as “Scenario 1”(i.e. it uses a pre-defined NAI and colects the data). In the “Scenario 2”, experts propose to create a new NAI for a particular analysis concern or even create a new specific analysis concern. In both cases, the experts would have to design a NAI from scratch, proposing modeling semantics, network analysis metrics, an operational analysis question, types of outputs, etc. (in conformance with the EANA-MM) following a **concern-oriented EANA approach**. These two scenarios are depicted in the top part of Figure 6.1

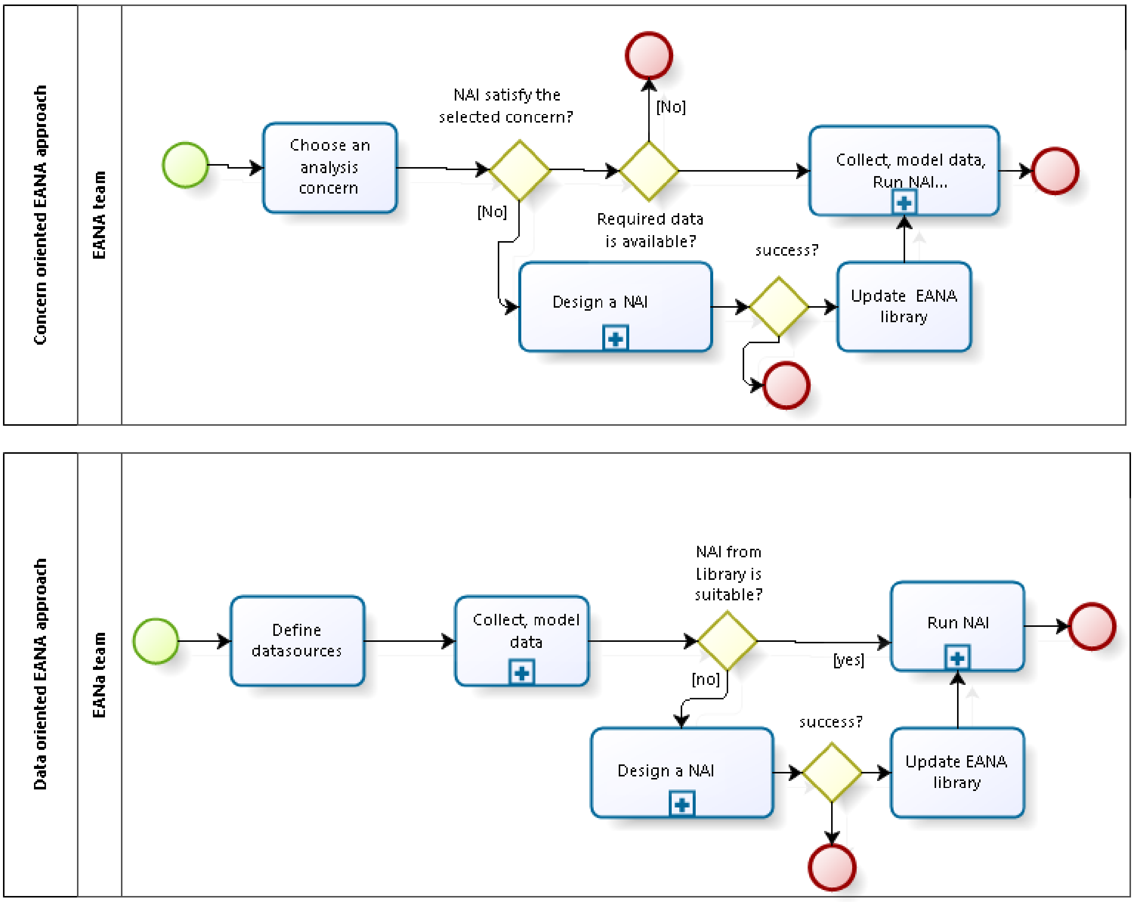

Figure 6.1: The concern and data oriented EANA approaches

In another realistic scenario, experts do not exactly know what analysis value can be extracted from the data they have, which is yet a delimited dataset (due to data collection restrictions; or due to the scope´s delimitation). In this case, experts can decide to look first at the available data to check if any NAI from the library fits with the data restrictions. This leads us to “Scenario 3”, which is logically equivalent to “Scenario 1” (i.e. uses a pre-defined NAI with the needed data). The difference here is that, first, we work in a limited dataset and after some data screening, we stick with concerns that are possible to analyze. Alternatively, experts can also decide to explore the analysis possibilities with the available data, with no pre-defined concern in mind ("Scenario 4"). The scenarios three and four implement an "mining approach", on which there is no analysis concern defined a priori, and the data will guide the concern´s choice. We call this the **data-oriented EANA approach**.

Figure 6.1 depicts the two EANA approaches. In this thesis, we design new NAIs using the data-oriented approach. In chapter 8, we present the results of this design and its empirical analysis. The activities "collect data, model the data", "Update EANA library" and "Design an NAI" appear agglutinated in Figure 6.1 to facilitate the EANA approaches 'understanding at this point. They are detailed in the following sections.

At the end of this activity, the team must be capable to select one of the two approaches. This choice will be determinant for the next activities.

6.3.3 Define actors

In this step, the internal actors involved in the analysis are specified according to the selected EANA approach. It has to be considered that, in later phases such "Build model" and "Run NAI", the internal actors need to bring in skills in advanced data analysis, network analysis and possibly statistical modeling. It is also important to ensure the presence in the team of members who can grant the access to the EA data sources. By internal actors, we mean any expert, from business or technical units, that can help in the design or application of the NAI. The role of those actors must be clear at the end of this activity.

6.3.4 Define stakeholders

Regarding the stakeholders, that is, the targeted users interested in the analysis results, STEEN et al. (2004) classify them in the categories described in Table 6.2:

Goal of the EA analysis	Possible stakeholders' roles
Informing	Customer, employee, others
Designing	Architect, software developer, business process designer
Deciding	Product manager, CIO, CEO

Table 6.2: The concern and data oriented EANA approaches. Adapted from STEEN et al. (2004)

In line with our EANA-MM, for each NAI, the team has to define its stakeholders, according to the categories described in Table 6.2, as an output of this activity.

6.3.5 Define data sources and data collection techniques

In this step, relevant EA data sources are identified and accessed. Thus, actors responsible for EA data are demanded to organize all the data sources available and consolidate a central

EA network data repository. Ideally, this task should be supported by tools, but in practice, it is common to see companies persist this data in ordinary datasheets. There are two categories of EA network data, according to the EANA-MM:

- Primary data

It may be an advantage if EAM solutions have already been introduced into the organization since a lot of required information of artifacts and their relations to each other might have been gathered and are maintained through cyclic processes (ADDICKS; APPELRATH, 2010). In these few cases where the enterprise function is mature, EA models, even based on Archimate language, are already available. For instance, in one of our case studies, the company used the ADOIT:CE tool that allowed us to export the EA data directly to common separated value (CSV) files (a commonly used input for network analysis tools).

If the data is not available, data responsible actors must **define collection procedures such interviews and document analysis**. FUERSTENAU; ROTHE (2014) suggest a template approach for EA data collection. In this case, the team prepares the template together with an experienced business expert. By doing so, the team ensures formulating the template in a comprehensible vocabulary and grounds it in the organization. After a pre-test, a fair fraction of subsidiaries/units should be sampled purposefully. An example of template proposed by FUERSTENAU; ROTHE (2014) contains source, target, transferred business object, and type of interface (e.g. online, manual, semi -manual) to map the systems' interfaces.

One practitioner from one of our case studies mentioned, for instance, that a web-based information system to identify each application and its interfaces, in a collaborative way, could improve the EA data management considerably. Unfortunately, such software solution was not available in his company. In other case study, we had access to EA data in the form of PDF files containing diagrams and textual information. Another important fact about data collection is that, depending on the organizational structure, the CIO can exert more or less influence on the organization's divisions to convince their directors to gather the required data (ADDICKS; APPELRATH, 2010). Thus, their importance in the data gathering process should be noted, too.

As discussed in section 4.2.10, according to our EA analysis SLR, 43 of 120 studies did not mention the data collection approaches for building their EAs. Among the remaining seventy-seven (77), we have papers using documents (54%), interviews (45%) and tool supported (8%) as their data sources. In our experience, the manual modeling task is very time-consuming and error-prone. The experts should have the issues above in mind while planning to perform data collection activities.

- Derived data

In conformance with the EANA-MM, the derived data depends on the existence of primary data and specific operators that allow the data derivation. The principles of derived data are discussed in the section 2.2.8.1 and detailed in chapter 7, when we present a set of EA data derivation operators.

Finally, at the end of this activity, data sources, targeted EA layers, data collection techniques and type of used data should be defined. FUERSTENAU; ROTHE (2014) suggest storing the data in a central architecture repository so the data becomes reusable for other projects.

6.3.6 Build model

EA modeling typically involves creating abstract representations of enterprises: the business processes involved, the IT-infrastructure, as well as the relations between them. Given a specific analysis concern, an enterprise architect decides which aspects of an enterprise are relevant and should be represented in the model. Examples of aspects that are frequently included in enterprise architecture models are products, business processes, applications and IT-infrastructure elements, as well as their relations. As such, an enterprise architect gathers relevant information and transforms this into a model (LANKHORST, 2013).

According to LANKHORST (2013) , it is important to note that there is no such thing as an inherently good – or inherently bad – model. To assess the quality of an architecture model, we have to take into account for what purpose the model is created and who the target audience is. Its stakeholders also determine the quality of an architecture: we state that an enterprise architecture that is a ‘correct’ and ‘complete’ representation of the real-life enterprise that is being modeled, given the objectives of stakeholders, has a high external quality. For a good set of principles, best practices to create EA models, please refer to chapter of the Lankhorst’s book (LANKHORST, 2013).

In the EANA process, the type of analysis approach adopted will influence the modeling activity. For instance, if the model should reveal the links between business processes and application components that are in use, thus structuring the model around services sounds appropriate. If the flow of processes that are triggered by an event or activity is the central concern, business process descriptions and interactions must be captured. Those are two examples of concern oriented EANA approaches which require those components and relations semantics to be reflected in the model under design. Thus, according to the EANA approach adopted, we can have the following modeling approaches:

- Concern oriented modeling (data still has to be collected)

In the case of concern oriented approach, considering a given analysis concern, the experts can pick an NAI from the EANA library. As a result, the NAI’s component and relation semantics have to be present in the model under construction. Select an NAI results in a semantic contract with its specifications, that is, the modeling decisions follow the ones specified by each NAI.

In a second option, the concern oriented modeling, from scratch, a new NAI is proposed (See "Design an NAI" step). In the next step, we collect primary non-modeled data from the organization that fits with the new NAI information requirements and get back to the "Build model" step.

- Data oriented modeling

In this step, modeling decisions are made based on the data already available. Enterprise architecture is hardly ever performed in a green field situation: typically, business process models, information models, or infrastructure models about (parts of) the enterprise already exist (LANKHORST, 2013). As such, one should always check the validity of any existing models, and incorporate their information on an appropriate level of abstraction; domain-specific models provide more details about parts of the enterprise than an EA model (LANKHORST, 2013).

In the Scenario 3 of the data-oriented EANA approach, we already have a set of modeled data, and thus, a "get-what-you-can" strategy takes place, since the "semantics of components and relationships were already defined in the existent documents and models. It is in this context that we develop our empirical analysis: we have been provided with empirical datasets whose modeling decisions were already defined by their respective companies' experts. Thus, the discussion about the modeling construction was kept outside of the scope of our work. In this case, an analysis concern from library can be reused (Scenario 3) or proposed based only on the already modeled data (Scenario 4).

For both analysis approaches (concern and data oriented), **at the end of the modeling activities, we are supposed to have gathered the information to create, structure, and visualize the enterprise architecture model**. Data modelers should pre-process the dataset, removing duplicates, checking the consistency of the data and preparing the adjacency matrix. Thus, **the model can be finally converted and stored as DSM/DMM or adjacency matrices**.

The well-known network analysis tools like ORA, UCINET, and R combined with igraph package, support this modeling choice as a general input. With the help of one of these tools, the data analysis team cleans and reconciles the data. This includes removing duplicates, checking the consistency of the data FUERSTENAU; ROTHE (2014). A valuable advice from (LANKHORST, 2013) refers to not making the model more informative than necessary.

In the EANA context, besides the model itself, we still have specific network modeling decisions to make such as which network metric, method or model should be applied to the EA model. For this and the other decisions, the EANA-MM might work as a checklist to verify if all the NAI elements were already defined. The specific network analysis elements are discussed in the following.

6.3.7 Design NAI

Design an NAI is not a mandatory task in the EANA process, since several NAIs are already available in the library. However, if this is the decision, a good knowledge base about the analysis concern and also about network science is essential to that task. In addition, we have to acknowledge that there is no template or common process for such activity since it is an essentially innovative one. Nevertheless, the EANA-MM and the EANA library may help us in that direction. For instance, the red dashed concepts in Figure 6.2 can be starting points for some insights as we discuss in the following.

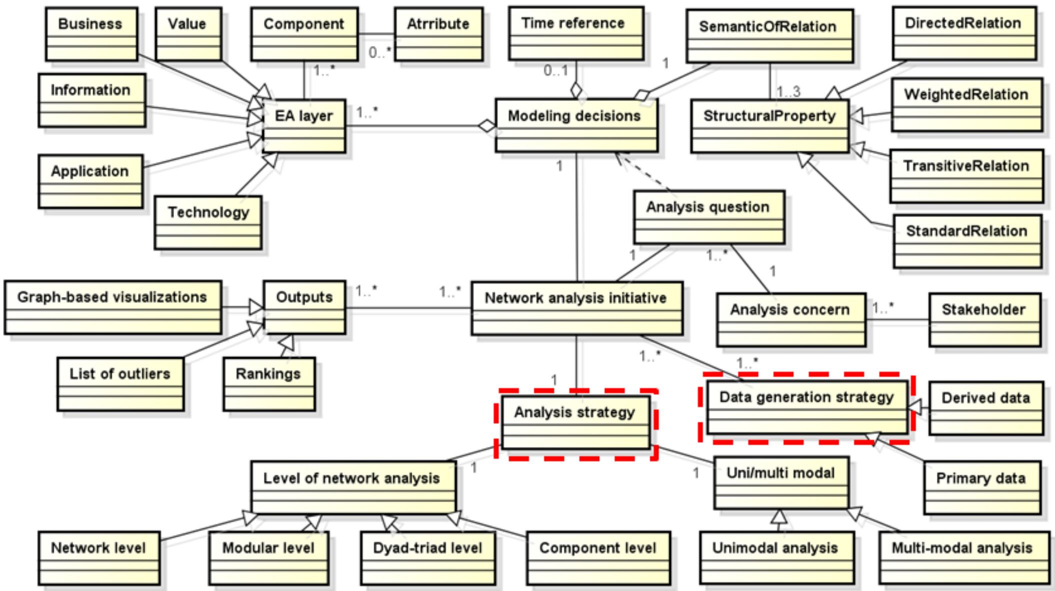

Figure 6.2: The EANA conceptual model

We describe, then, some general guidelines that might support the design process.

- Guideline 1 : Analysis strategy as a decision point

The first guideline is related to the selected analysis strategy (see Figure 6.2) , and more precisely, it is about defining a level of (network) analysis. This decision tells a lot about the kinds of analysis metrics and corresponding types of results that might be expected. We discuss how the three levels of analysis can be determinant in the NAI´s design in the following.

Network-level analysis. Network-level analysis (SCOTT, 1992) or macro analysis (FUERSTENAU; ROTHE, 2014) focuses on patterns of interconnections. They are typically visualized graphically or with the help of adjacency matrices. One of the most prominent coefficients for network analysis is the density of a network. Designers at this level of analysis should **aim to perform overall structural visual analysis or yet, to calculate network structural indicators** such as density and network degree of centralization, whose values then, in turn, should be contextualized in the analysis context.

Group level analysis. This is based on the types of edges and the (non-)existence of relationships. It is highly concerned with cliques, clusters and related group structures. NAIs applied at this level should concern about the **notion of grouping EA components for different analysis purposes**. For instance, clustering the application layer may reveal the degree of modularization among the information systems. A similar concept can be applied to other EA layers individually or combined.

Component level analysis. At this level, the focus is to **identify components that are important in terms of structure or study the attributes of those components in key structural positions**. The amount of component level analysis present in our EANA review outnumbers the two types mentioned before. In (Santana A.; Fischbach K.; Moura H., 2016a) a set of NAIs were classified according to the level of analysis adopted by the researchers in primary studies, resulting in a classification schema comprising four categories depicted in Figure 6.3.

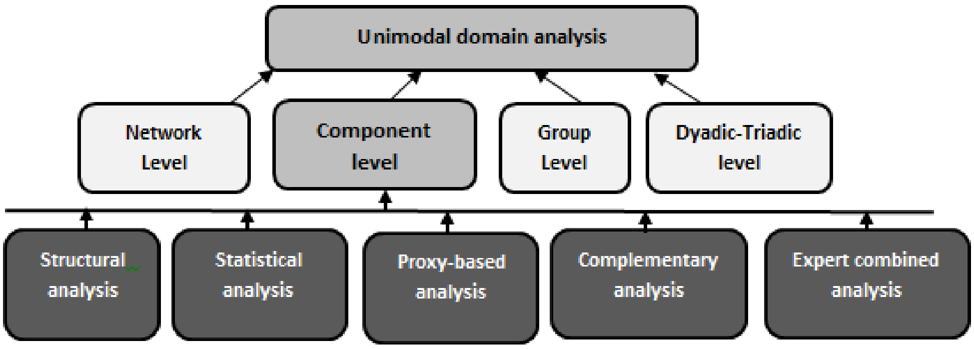

Figure 6.3: A classification schema of methods for EANA at the component level

If the choice is for an NAI comprises the analysis at the component level, the NAI designer might benefit from this classification schema to check what has been done so far in terms of NAIs. The existent NAIs at component level can be classified into the following categories:

Pure structural analysis. This category comprises works that try purely to identify architectures' structural components. In this category, papers highlight components with structural roles acting like bridges or hubs or are connected with other well-connected ones. Input for this analysis is limited to information about the relations among the components. Statistical models are excluded. NARANJO; SÁNCHEZ; VILLALOBOS (2014), which investigate the component's centrality degree, and BARTOLOMEI et al. (2012), which apply degree and betweenness centrality in their analysis, are examples of such studies.

Statistical analysis. This category comprises research aimed at building statistical models or finding correlations based on network measures. LEVINA; HILLMANN (2012), for example, use discriminant functions based on several network measures to classify business processes – core, information intensive, flexible, or automatable processes – in categories according to type.

Proxy-based analysis. The independent variables are composed of network measures only and the researcher is interested in using network measures as proxies for EA concerns such as the cost of change. The work of LANKHORST (2013) is an example of such an approach.

Complementary analysis. This category includes papers that use two complementary types of information as input to investigate an EA concern. Its main difference with the proxy-based analysis is the use of a component's properties (e.g. whether an application is self-developed, cloud-based, or off-the-shelf) together with network measures. For instance, POSTINA; SECHYN; STEFFENS (2009) use application attributes such as real-time requirements and operating cost together with outdegree centrality to evaluate the gap between the current and future EA states. In the work of LEHNERT et al. (2015), the Process-PageRank (PPR) algorithm ranks business process for improvement using degree centrality and a need for improvement (a business process attribute) to calculate the network-adjusted need for improvement. So far, there has been no discussion regarding the extent to which the structural information could be used together with expert knowledge (e.g. top-ranked applications perceived by the expert). Experts might want to know, for example, whether applications that occupy significant structural positions are hosted in-house or are outsourced. This discussion leads us to the next category of analytic methods.

Expert-combined analysis. This category includes papers that aim to investigate how the value extracted from network measures can be combined with tacit expert knowledge. In SANTANA et al. (2016c), this tacit expertise is combined with information obtained from structural analysis of the application architecture. We assume that one can use non-functional characteristics of EA components, such as business criticality or strategic relevance as judged by experts, with structural information. The second strategy under this category is to check whether a specific attribute value is present in components that occupy significant structural positions. For example, architects might be interested to know what kind of software (e.g. self-developed, cloud-based, or off-the-shelf) is running on applications with the highest eigenvector values.

To finish the analysis at component level guideline, we present in Figure 6.4 a decision schema about the NAI possibilities at this level. There are two decisive criteria to select a category: if a component´s attribute is considered in the analysis and if there is a substantial correlation (at least moderate for rho values) between the attribute values and network values of components. The rules are graphically explained in Figure 6.4:

Of course, the previous list of categories is not exhaustive for EANA at the component level. However, this schema helps us contextualize the kinds of NAIs that might be proposed.

- Guideline 02: Look at the data types and restrictions

If we have a delimited dataset with no possibilities of further data collection efforts, we need to stick with the given components and relations (the primary data already modeled). In this sense, the designer can use the EANA library as a reference to get to know which concerns and analysis techniques have been applied to these components and similar relations to get some inspiration for the new NAI.

A second option would be to look at indirect relations of components, as explained in section 2.2.8.1, to derive new networks, and consequently, forge new analyses perspectives which may result in additional NAIs. This idea is explored in chapter 7.

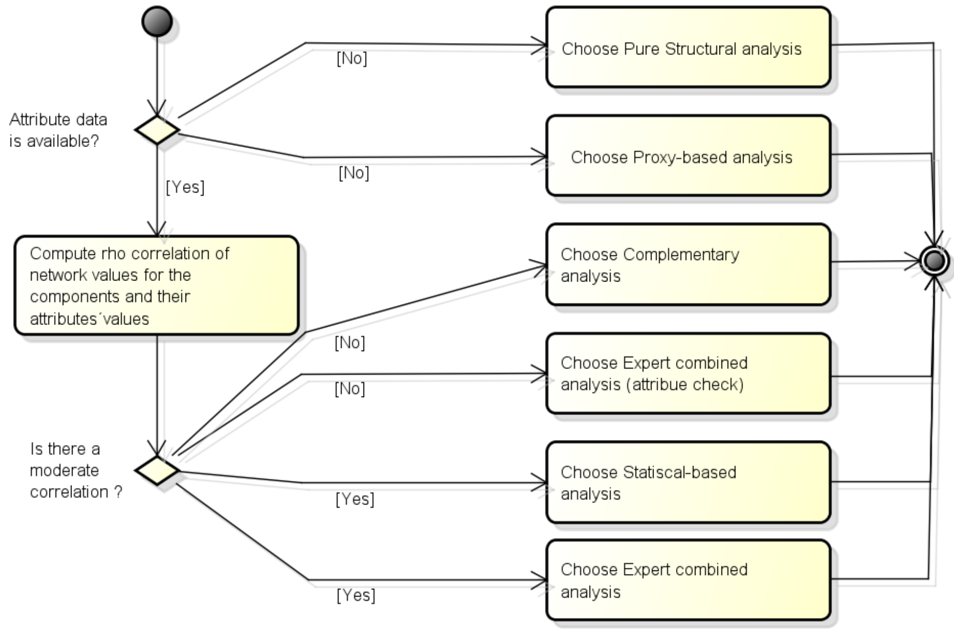

Figure 6.4: A decision schema to select a category of EANA at component level

- Guideline 03: Using the EANA library as a reference

The third guideline for NAI design is actually, more like a general advice. One can look back to the EANA library as a reference and search for similar concerns and modeling semantics for relations and components, and at the level of network analysis. For instance, given a dataset containing data related to applications and business process (a bi-modal network), the designer can search for analysis techniques applied to other bi-modal networks in the library to have a first impression about the analysis possibilities. In this sense, a consolidated body of knowledge about NAIs may be very useful.

6.3.8 Update EANA Library

Considering a public and available web repository, every time a new NAI is proposed, authors are invited to update the EANA library with all the information required for an NAI, as specified in the EANA-MM. As a subproduct of this thesis, the web library for EANA can be accessed at http://www.eanaresearch.org. Researchers and practitioners are encouraged to adopt the EANA library in their respective activities.

6.3.9 Run NAI

This is a major step since here we prepare a suitable metric, method or network model to perform the structural analysis with the EA data modeled as adjacency matrices. Since we did not implement any software tool to perform our analysis in this thesis, we used three different network analysis tools, in a complementary way, to import, export and analyze our data. Our user experience with those tools is described in Table 6.3.

Tool	Cost	Our user experience
UCINET[1] tool student edition for Windows (only)	$40 for single user student $150 for single user faculty	Main advantages: ▪ Perhaps, the most traditional network analysis tool. ▪ It offers a big set of network analysis measures at all network levels of analysis. ▪ It is easy to import or export data from/to several standard formats. Disadvantages: ▪ Most of the result manipulation is done with raw data (e.g. text files). ▪ Visualization of the network is relatively poor quality.

Continued on next page

Table 6.3: Network analysis tools used in this thesis

[1]https://sites.google.com/site/ucinetsoftware/pricing-and-licensing-information

Table 6.3 – continued from previous page

Tool	Cost	Our user experience
ORA-LITE[2] Student Edition for Windows (only)	Free for educational purposes during the first year	Main advantages: ▪ It is very easy and intuitive to manipulate. ▪ It offers a big set of network analysis measures at all network levels of analysis. ▪ It produces several reports in an easy way. ▪ It has a powerful network visualization editor. ▪ It is easy to import or export data to several standard formats. Disadvantages: ▪ The quality of the graphics resulted from clustering analysis algorithms was not good (at least, until the 2014 version).

Continued on next page

Table 6.3: Network analysis tools used in this thesis

[2]http://www.casos.cs.cmu.edu/projects/ora/software.php

Table 6.3 – continued from previous page

Tool	Cost	Our user experience
R[3] tool plus igraph package[4] for Windows	Free software	Main advantages: ▪ The visualization aspects of the outputs are the best feature of this tool combination. It provides a high degree of customization for the outputs (e.g. component and edges labels, colors, sizes, thickness) ▪ It also provides an important result format for our work, which is heat map. Disadvantages: ▪ It is a command line based tool.

Table 6.3: Network analysis tools used in this thesis

In addition to those tools, we also used the SPSS only to calculate Spearman correlations between network values and other variables, for two specific NAIs proposed in chapter 8. Datasheets in Excel (i.e. CSV files) worked as our data storage system.

We can have rankings, clusters, heat maps or visual graphs as output formats for the EANA process, as discussed in section 2.2.5. In the next step, we deal with the interpretation of those outputs to extract value for stakeholders.

6.3.10 Analyze and communicate results

After importing the data to a network analysis tool (e.g. ORA or R), one should start discussing and revising the findings with stakeholders by using network visualizations and measures as discussed earlier. After appreciating the feedback, the team may suggest an EA change or refactoring plan, or yet a monitoring report as a final output. In the case of a new NAI has been developed and validated, the EANA must be updated.

[3]https://www.r-project.org/
[4]https://www.r-project.org/

6.4 Methods for concern and data oriented EANA approaches

As discussed before, the sequence of the EANA activities presented may vary in terms of execution order according to the analysis approach adopted(concern or data oriented), generating 4 different analysis scenarios. We start showing the suggested workflow to perform EANA with the concern-oriented approach.

6.4.1 Concern-oriented analysis approach

The concern-oriented EANA approach´s workflow is depicted in Figure 6.5:

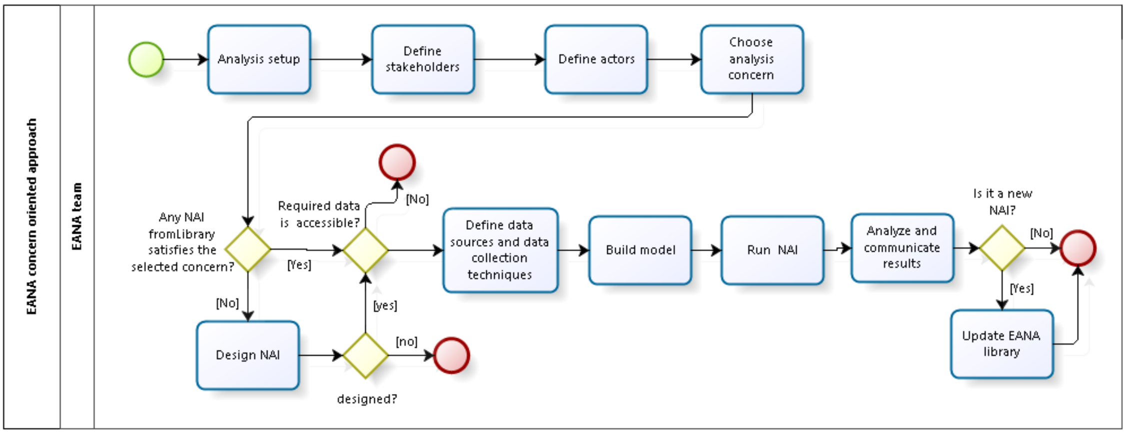

Figure 6.5: Activities of the EANA with concern oriented approach

As described in section 6.3, there are the main activities which we group now in four different phases: "Planning", "Data collection and modeling", "NAI execution", and "Results analysis and communication". In each phase, several activities may be carried out. For instance, in the “Planning” phase, we execute the “Analysis setup”, “Define stakeholders”, “Define actors” and “Define analysis approach” activitie. During this phase, we define the team members, the analysis concern and how it is addressed and which data sources are available (See section 6.3.1, 6.3.2, 6.3.3, 6.3.4 and 6.3.5 for details). We name the second phase as “Data collection and modeling”. Following the planned decisions, the data is collected and modeled as adjacency matrices at the end of this phase. Due to the complexity of the activity, the third phase consists in the “NAI execution”, on which network analysis metrics/methods/models are applied to the modeled data, producing the outputs for the final phase, “Results analysis and communication”. In Table 6.4, we briefly summarize the main activities per phase and their main inputs and outputs.

Phase/Activity	Input	Output
I/Analysis setup	EA data sources; EA models (if existent); Stakeholder request for EA analysis(optional)	Team definition; Overall check over the data sources (initial list) and EA models (if existent); list of available tools, grant access to data; terms of confidentiality (if needed).
I/Define analysis approach	Motivation	Definition of the analysis approach
I/ Define stakeholders	Selected analysis approach	List of stakeholders
I/ Define actors	Selected analysis approach	List of (internal) actors
I/ Choose analysis concern	EANA library, Stakeholder request for EA analysis(optional).	Selected analysis concern
II/Design NAI(optional)	Analysis concern; Overall check over the data sources (initial list) and EA models (if existent);	The new NAI or abort process.
II/ Define data sources	Analysis approach; Analysis concern; EA data sources list	Definition of the type of EA data to be used (primary and/or derived); targeted EA layers; EA data sources to be accessed; EA models to be used;
II/Define data collection techniques	Analysis approach ; Analysis concern; EA data sources list	Data collection techniques; data derivation methods;
II/ Build model	Definition of the type of EA data to be used (primary and/or derived); targeted EA layers; EA data sources to be accessed.; data collection techniques; data derivation methods;	Populated model converted to adjacency matrices.
III/ Run NAI	Adjacency matrices, NAI including a network analysis metrics/method or network model.	Network analysis results for the EA concern. Results can be presented in any of those formats: rankings, numerical values, Heat maps, DSMs and graphs.

Continued on next page

Table 6.4: Activities, inputs, outputs for EANA concert oriented approach

Table 6.4 – continued from previous page

Phase/Activity	Input	Output
IV- Analyze and communicate results	Network analysis results for the EA concern.	EA change or refactoring plan; Monitoring report.

Table 6.4: Activities, inputs, outputs for EANA concert oriented approach

Finally, in Figure 6.6, we present an overview of each phase, actors involved and tools that might support the analysis process:

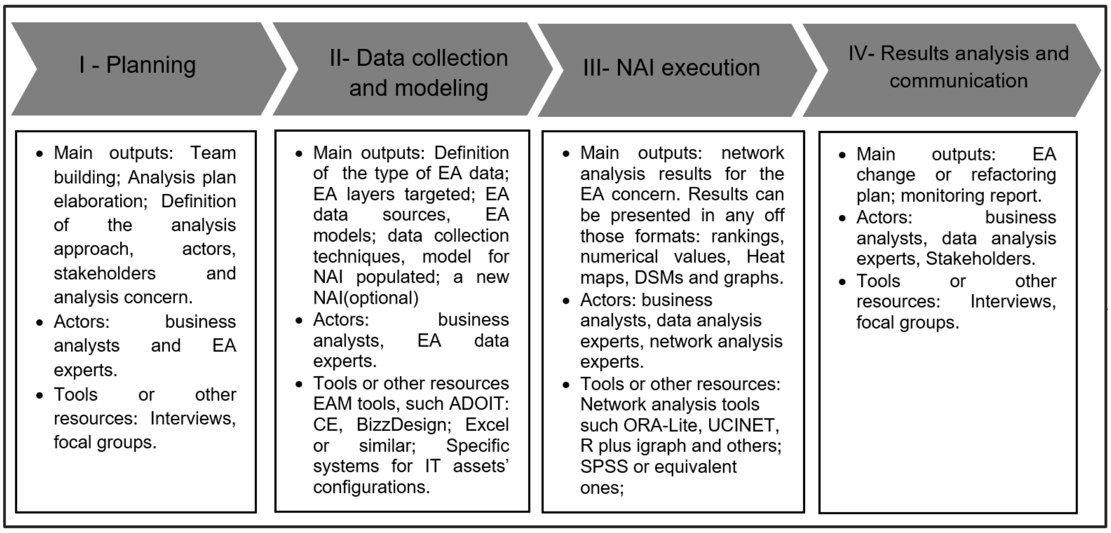

Figure 6.6: Phases of EANA data-oriented approach

6.4.2 Data-oriented analysis approach

This "mining" approach is mainly characterized by having a delimited dataset so that the definition of the data sources is shifted to the planning phase. Thus, two possibilities of analysis can be reached. In the first, aware of the data possibilities, EA experts and stakeholders can define a concern and thus look at the EANA library for suitable NAI candidates. Alternatively, experts can also decide to design an NAI. In this case, data derivation methods can play a major role. Stakeholders appear in the process after a first screening in the data-sources performed by the experts. This suggested workflow is depicted in Figure 6.7:

In Table 6.5, we briefly summarize the main activities per phase and their main inputs and outputs.

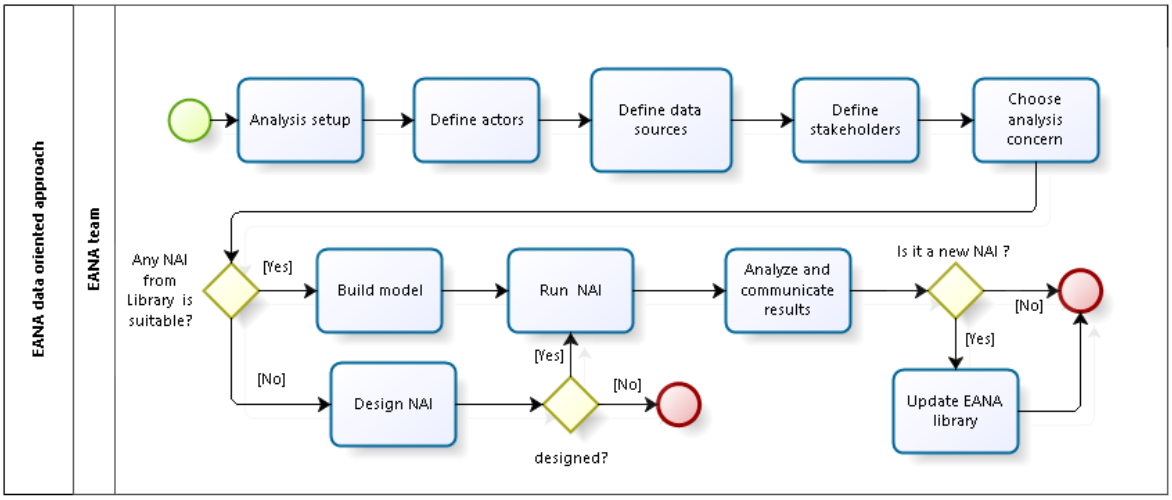

Figure 6.7: Activities of the EANA with data oriented approach

Phase/Activity	Input	Output
I/Analysis setup	EA data sources;	Team definition; Overall check over the data sources (initial list), available tools; grant access to data; terms of confidentiality (if needed).
I/Define actors	Overall check over the data sources (initial list), available tools	List of (internal) actors.
I/ Define data sources and EA models	Analysis concern; EA data sources list.	Definition of the type of EA data to be used (primary and/or derived); targeted EA layers; EA data sources to be accessed. EA models; Overview of analysis possibilities.
II/ Choose analysis concern	Stakeholder motivation; EA data sources and EA models; Overview of analysis possibilities; EANA library.	Analysis concern.
II/ Define stakeholders	EA data sources to be accessed. EA models; Overview of analysis possibilities.	List of stakeholders
II/Design NAI(optional)	EA data sources; EA models; Analysis concern	The new NAI or abort process.
		Continued on next page

Table 6.5: Activities, inputs, outputs for EANA data-oriented approach

Table 6.5 – continued from previous page

Phase/Activity	Input	Output
II/ Build model	EA data sources; EA models; Analysis concern	Model populated and converted to adjacency matrices.
III/ Run NAI	Model populated and converted to adjacency matrices, NAI (including a network analysis metrics/method or network model)	Network analysis results for the EA concern. Results can be presented in any of those formats: rankings, numerical values, heat maps, DSMs and graphs.
IV- Analyze and communicate results	Network analysis results for the EA concern.	EA change or refactoring plan; Monitoring report. Update EANA library(if necessary).

Table 6.5: Activities, inputs, outputs for EANA data-oriented approach

The EANA data-oriented approach and its phases, actors involved and tools that might support the analysis process are depicted in Figure 6.8.

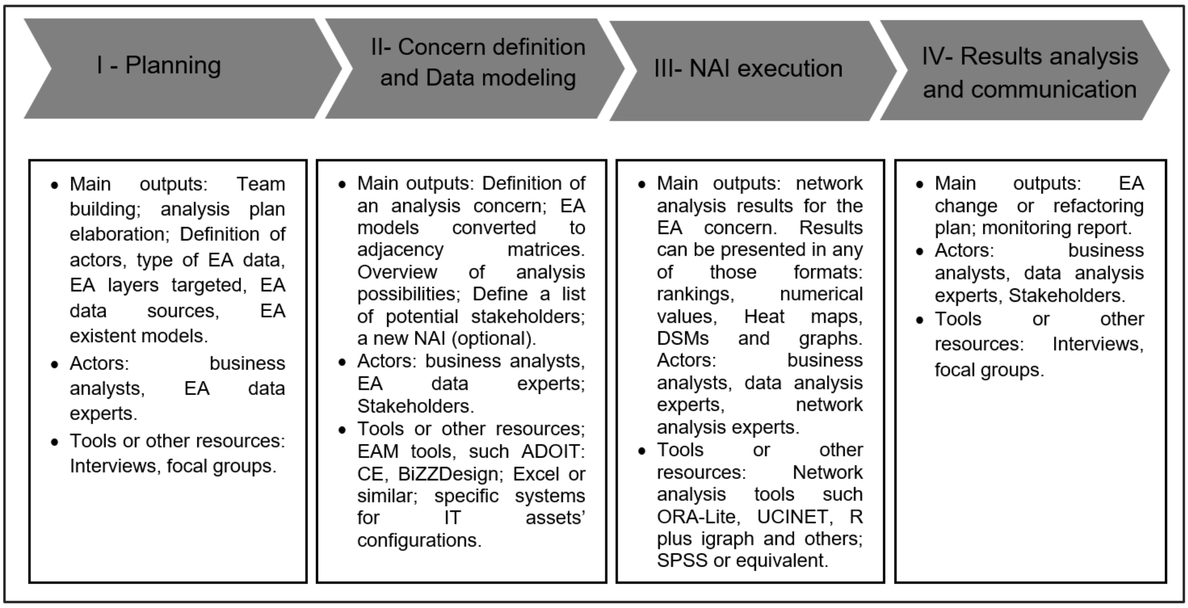

Figure 6.8: Phases of EANA data-oriented approach

In both EANA approaches (concern and data oriented), there are two optional activities: “Design NAI” and “Update library”. Far from being trivial, the designing of an NAI has no template or defined process, a knowledge intensive activity, which may use different support tools, to provide an original NAI (see section 6.3.7).

6.5 Chapter Summary

In this chapter, we presented a process view for EANA. To build our artifact, we ground its development in the EANA-MM concepts and EANA library both resulted from a thorough review of the literature on EANA presented in chapter 5. We also elaborate on the initial ideas proposed in ADDICKS; APPELRATH (2010), LANKHORST (2013), VENEBERG et al. (2014) and FUERSTENAU; ROTHE (2014), adding EANA contextual issues to them. Finally, in combination with that, we also use insights obtained from three empirical analysis performed by this thesis´author. Our goal was to design a process to guide researchers and practitioners in performing EANA.

The EANA process view, thus, is defined by two analysis approaches: concern oriented and data oriented. Both approaches are divided into ten main activities. We detailed those activities' inputs, outputs, actors and stakeholders and even indicated some tools that might support their execution.

Regarding the concern oriented approach, we believe that all NAIs which are part of the EANA library are already implicitly validated or at least, theoretically justified in their primary studies, thus constituting "ready-to-run" analysis possibilities for the experts. In this sense, our contribution was to consolidate, organize and provide a decision structure for choosing the right NAI according to the desired analysis concern, which now is complemented by the EANA process. As for the data-oriented approach, we demonstrate it in three real enterprise datasets while we create a strategy to derive EA network data and two NAIs, that are discussed in the chapters 7 and 8, respectively.

For all activities mentioned in the phases of our two methods, we acknowledge that it might be useful to refine some steps, through the splitting of activities into smaller activities, and/or customizing the proposed method for the company case under analysis. As in VENEBERG et al. (2014), we have left the assignment of roles to activities unspecified. However, it should be noted that this might be an improvement target for next versions of the methods. In general, domain experts from business, EA and network analysis areas should be involved and collaborate with them. The two EANA approaches also need to receive feedback from the research community and practice to become more robust. We believe that the EANA process view could benefit considerably if incorporated in specific EANA tools to be used by practitioners.

One last reflection that we would like to leave for EANA practitioners is based on the modeling guidelines of LANKHORST (2013). According to those authors, we would have to be able to answer the following questions before starting any modeling initiative:

- Is there a clear stakeholder for it?
- Is the analysis objective explicit?
- Will it create an enterprise architecture model help to reach this objective?

- Are the boundaries clear of what should be modeled?
- Are there realistic expectations regarding the results´application?

In the next chapter, we present data derivation operators which will allow us to design new EA viewpoints and consequently, new analysis possibilities.

7

Derivation viewpoints for EANA

In the last chapter, a process was introduced in order to guide experts in performing EANA. In this chapter, the focus shifts a bit, while an additional data generation mechanism to expand EANA possibilities is presented. Co-affiliation networks (described in section 2.2.8.1) and Design Structure Matrix (DSM) theory serve as theoretical support for designing of a set of data derivation operators which constitute the fourth and last component of the EANA framework.

This chapter is structured as follows. In section 2, topics and terms relevant for the understanding will be explained and the related work presented. The research approach is shown in section 3 and the resulting meta-model and its implications are detailed in section 4. In section 5, three examples illustrate how the derived composition operator can be applied in modeling practice. Finally, we draw conclusions for the chapter in section 6.

7.1 Context and motivation

Since Zachman's framework to the more recent releases of modeling languages such Archimate 2.1, viewpoints, a well-established concept from software architecture (STEEN et al., 2004), have represented a core concept for EA modeling process for nearly thirty years. Diagrams and related visual modeling work well for enterprise architecture views, which express the EA from the perspective of specific concerns and stakeholders (STEEN et al., 2004). They are also an essential result of any EA management initiative and are provided to different stakeholders (GRIGORIEV; KUDRYAVTSEV B. D.A GRIGORIEV; KUDRYAVTSEV, 2013). However, their representation´s power tends to diminish when complex systems are modeled. In some situations, performing only a visual analysis with these models is a crucial limitation, since it does not allow working with a large amount of data; as long lists or diagrams can not be processed by the human brain (GRIGORIEV; KUDRYAVTSEV B. D.A GRIGORIEV; KUDRYAVTSEV, 2013). Thus, matrix-based methods appear as an interesting modeling and analysis approach.

Indeed, the very nature of enterprise architecting requires a lot of matrices (goals, processes, capabilities-processes, processes, applications, etc) (GRIGORIEV; KUDRYAVTSEV B. D.A GRIGORIEV; KUDRYAVTSEV, 2013). Matrices are also actively used in enterprise architecture management. For example, TOGAF-recommended matrices: actor/role matrix, application/data matrix, application/function matrix (GRIGORIEV; KUDRYAVTSEV B. D.A GRIGORIEV; KUDRYAVTSEV, 2013), etc. Matrix representation is supported by the majority of EAM tools (GRIGORIEV; KUDRYAVTSEV B. D.A GRIGORIEV; KUDRYAVTSEV, 2013): matrix editor in ARIS design platform (DAVIS, 2008), matrix editor in IBM system architect (QUATRANI; PALISTRANT, 2006) and matrix manager in Casewise modeler (CASEWISE, 2016). Several fields such as system engineering, system of systems, project management, and product engineering have been applying DSM, a matrix based method, to successfully analyze complex systems. Despite the similarities of EA as a system with those fields, in the EA context, only a few works have explored this modeling approach. LANKHORST (2013) and SANTANA et al. (2016b) are some exceptions.

In this direction, thinking of EA as a robust and intuitive matrix, several layers and components can be modeled with it. Once we have that EA matrix, using data derivation concepts initially applied by BUUREN et al. (2004) and MATHEWS; KAESLIN; RYTZ (2014), artificial viewpoints can be derived, providing new analysis angles and new questions for the stakeholders. In this sense, we elaborate on the work of BUUREN et al. (2004), expanding the derivation operator possibilities and applying them together with network analysis measures on well-known EA viewpoints. In this research, we model these viewpoints as DSMs (or equivalently, networks) and thus, we propose a matrix of derived viewpoints to support EA analysis (EADV-MM) as a way to answer the following question: How can we derive a set of EA viewpoints using indirect relations among their components? Furthermore,which questions arise from those new viewpoints?

Therefore, the main contribution of this piece of research is indeed conceptual: the Enterprise Architecture Derived Viewpoints Meta-Model (EADV-MM) and derivation operators, aiming to reinforce a still lacking, concrete way of using DSM principles in EA analysis. Not limited to that, the applicability and utility of a subset of the derived viewpoints are demonstrated in three empirical cases.

7.2 Key concepts

7.2.1 EA viewpoints

Due to the plethora of frameworks and models to integrate the various architecture descriptions, as in (LANKHORST, 2004), we also advocate an approach on which architects and other stakeholders can define their own views and viewpoints of the EA. The ideas of using viewpoints in EA appeared first in IEEE 1471 standard for architecture description (IEEE, 2000) coming from the software engineering community (LANKHORST, 2013). Currently, there are several frameworks based on viewpoints such as the Zachman framework (ZACHMAN, 1987), Kruchten's 4+1 view model (KRUCHTEN, 1995), RM-ODP (RAYMOND, 1995), and TOGAF (The Open Group, 2011).

In this context, two first conceptual distinctions are needed: a view is a representation of a whole system from the perspective of a related set of concerns while a viewpoint establishes the purposes and audience for a view and the techniques for its creation and analysis (LANKHORST, 2013). Those concepts are contemplated in the IEEE 1741 standard definition for viewpoint, which should specify at least the following requirements described in Table 7.1.

Viewpoint's requirement	**Description**	**Instantiation in this work**
A view point name	A name or identification for the viewpoint	Any given name for the viewpoint. E.g Unimodal TecxTec.
Stakeholder(s)	Stakeholders to whom the viewpoint is aimed.	(STEEN et al., 2004) present a list of stakeholders that could benefit from its viewpoint model analysis for EA: architect, software developer, business process designer, product manager, CIO, CEO, customer, employer and others.
Concern(s)	Concern the viewpoint addresses	Analysis targets for stakeholders. Examples of structural concerns are: to find the key structural components, to measure the component´s change impact, etc.
The language modeling techniques	The modeling language techniques or analytical methods to be used in constructing a view based upon the viewpoint	Modeling choices for EA are quite diverse and numerous. EA should focus on the appropriate level of abstraction (STEEN et al., 2004). In this sense, a more abstract modeling approach can cover more real cases. Thus, we advocate DSMs or network graphs for that purpose.
Analytical methods	The method applied together with the viewpoint.	We explore network analysis metrics as an analysis toolset for EA viewpoints modeled as a graph or DSM (SCOTT, 1992; BARTOLOMEI, 2007).

Table 7.1: Viewpoint's requirements defined by IEEE 1741 (IEEE, 2000)

In support of TOGAF, Archimate defines 18 standard viewpoints to cover the enterprise architecture, focused on different sets of stakeholders. These are comprehensively documented on the Open Group Archimate website[1]. A comprehensive discussion of EA viewpoints can be found in STEEN et al. (2004).

7.2.2 DSM and related concepts

A DSM is a square matrix with thesame row and column labels which provides a simple, compact, and visual representation of a complex system and supports innovative solutions to

[1] http://www.opengroup.org/subjectareas/enterprise/archimate-overview

decomposition and integration problem (BROWNING, 2001). Depicted in Figure 7.1 is a typical design structure matrix, formed by elements or components. The presence or absence of a relation between each pair of components is represented by the "1" or blank cell (or "0"). With this model structure, we can represent connections among components from the same layer (e.g. application layer) and also with components from other EA layers (business and application layers). The values contained in the cells may have diverse semantics. For instance, it can represent the frequency of data exchange between two applications. Therefore, a matrix that has different values than "0" and "1" (binary DSM) adds semantic about the intensity or quality of that relation, besides its absence or presence itself.

	A	B	C	D	E	F	G
Element A	A	1				1	
Element B		B		1			
Element C	1		C				1
Element D				D	1		
Element E		1			E	1	
Element F			1			F	
Element G	1				1		G

Figure 7.1: Graphical representation of a binary DSM

DSM and Domain-Mapping Matrix (DMM) are the two central elements of a Multiple Domain Matrix (MDM). A DMM is a special kind of DSM which maps the relations between exactly two different domains (in the EA context, a domain is equivalent to a layer) of elements (in the EA context, an element is equivalent to a component). For instance, a DMM can map the "support" relation existent among application and business process components (see "DMM area" in Figure 7.2). An MDM extends the capabilities of the DMM by integrating multiple domains (or multiple EA layers) and reunites all single DSMs that are required for storing information about system's elements (in the EA context, the EA itself) and their relations in one single matrix system. In the EA context, an MDM may encompass more than one type of layer (e.g. technology, application and business process), together with different kinds of relations (e.g. "deployed on", "connected with" or "depends on") among components from those layers. An MDM enables the deduction of indirect (derived) relations among the modeled components, which in turn, allows us to derive new DSMs. We will discuss further those indirect relations in section 3, when we show mechanisms to derive DSMs. The concepts of DSM, DMMs, MDM and indirect relations are depicted in Figure 7.2.

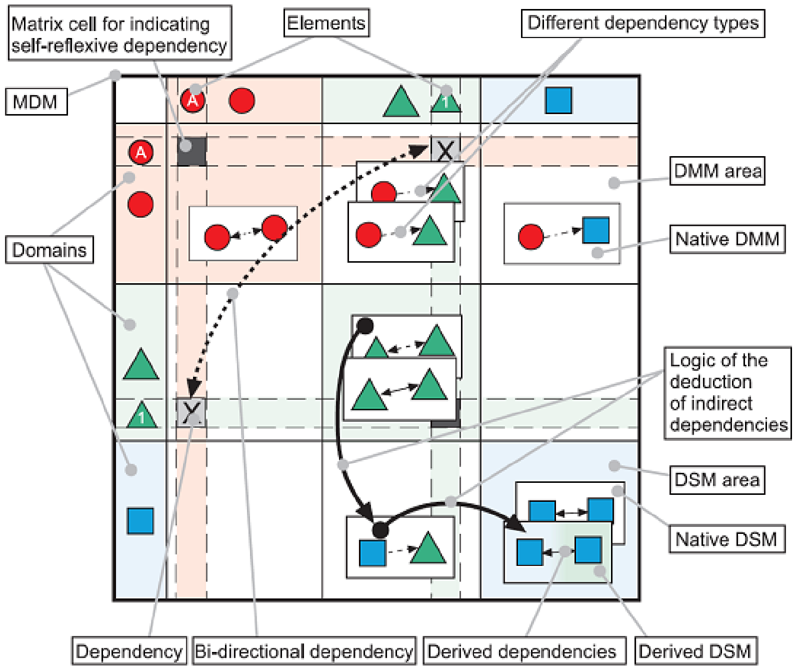

Figure 7.2: Representation of MDM's concepts – adapted from FURTMEIER; TOMMELEIN (2010)

As shown in Figure 7.2, we can model practically any system composed of components and relations with an MDM, that is, MDMs, DSM or equivalently, network graphs, are generic modeling choice candidates for EA analysis, since their concepts of components (objects) and their respective relations are fundamental parts of EA models as well and, in terms of abstraction, they lay one level up to the EA conceptualization.

For more information about DSM applications in other domains, as well as recent developments with domain DMMs and MDMs, please refer to the literature review of BROWNING (2016).

7.2.3 DSMs and EA

Since the 1990s, DSMs´ system representation and analysis techniques have led to their increasing use in a variety of contexts, including product development, project planning, project management, systems engineering, and organization design (BROWNING, 2001). In these different fields various other methods that rely heavily on matrices have also been developed: IBM's Business Systems Planning (BSP for strategic information system planning; Quality Function Deployment (QFD) for product design and total quality management (GRIGORIEV; KUDRYAVTSEV B. D.A GRIGORIEV; KUDRYAVTSEV, 2013). Technology has now advanced, especially in the area of databases, and can support a single conceptual graph represented as both a diagram and a matrix, so the user can benefit from the matrix representation in just those areas where it is an easier paradigm than graphical diagrams (GRIGORIEV; KUDRYAVTSEV B. D.A GRIGORIEV; KUDRYAVTSEV, 2013). Matrices are also actively used in enterprise architecture management. In this direction, following the intuitive idea of seeing EA as a system that can be represented as matrices, this piece of research transfers the DSMs, DMMs and MDMs concepts and related analysis methods originally from systems engineering to discuss them in the EA context. In this sense, we explore the applicability of the MDM and DSMs to model EA viewpoints and their potential to offer new insights of analysis, particularly using derived EA viewpoints combined with network analysis techniques.

Modeling viewpoints using matrices in EA context is not exactly a new approach. For instance, the Archimate itself already explored the EA analysis through views (bi-dimensional matrices). However, performing only a visual analysis with these models is a crucial limitation, since it doesn't allow working with a large amount of data; long lists or diagrams cannot be processed by the human brain. Thus, matrix-based methods outperform them in some situations (GRIGORIEV; KUDRYAVTSEV B. D.A GRIGORIEV; KUDRYAVTSEV, 2013). The novelty of this work is not modeling EA viewpoints as DSMs itself, but rather in:

- Explore MDMs' indirect dependencies, looking at the DSMs's theoretical background and network science to create new EA viewpoints.
- Complement the traditional visual analysis, often performed based only in EA viewpoints, with network analysis metrics such as centrality measures and clustering algorithms.

7.2.4 Related work

FURTMEIER; TOMMELEIN (2010) explored the application of MDMs as a mapping process approach in an analogous way (in terms of methodology) to what we do in this research. However, they focus on lean design context. Especially in the context of EA analysis, the DSM has been applied by a few papers. According to VAKKURI (2013) in his master thesis, the DSM seems to be a good choice for EA analysis for two reasons. First, past EA frameworks (The DoDAF Architecture Framework Version 2.02, 2010; The Open Group, 2011) propose using square matrices in modeling connections between system components. Secondly, many types of analysis approaches have been developed for the DSM. However, there are few literary works in which the DSM is used explicitly in an EA context (LAGERSTRÖM et al., 2014; SANTANA et al., 2016b; WALDMAN; SANGAL, 2007). In his work, VAKKURI (2013) focuses on application architecture and how it can be analyzed to select TO-BE states with the help of DSMs.

LANKHORST (2013) work with EA viewpoints, their use in communication, giving guidelines for their selection and use, and outlining a number of viewpoints on the Archimate language that can be used by architects involved in the creation or change of enterprise architecture models. In the same direction, STEEN et al. (2004) create a classification framework for EA viewpoints. They also discuss the importance of transformation models to convert models from any language to EA concepts (in this case, Archimate works as an integration language) and thus, generate viewpoints in a tool supported environment. Nevertheless, the authors did not propose any particular viewpoints nor perform any kind of analysis. Likewise, GRIGORIEV; KUDRYAVTSEV B. D.A GRIGORIEV; KUDRYAVTSEV (2013) emphasize the potential of using matrices to model EA and propose a modeling tool to integrate business process engineering with EA. However, those authors do not perform any analysis in their paper as well.

Regarding the use of indirect relations in DSMs and EA contexts, only a few papers can be cited: BALDWIN; MACCORMACK; RUSNAK (2013) defines an EA DSM to analyze the impact of indirect changes and identify the structural architectural arrangements with the "hidden structure method"; MATHEWS; KAESLIN; RYTZ (2014) consider multimodal dependencies among the enterprise components to form a dependency graph to deduce indirect dependencies between applications and IT-projects. Such indirect dependencies are useful for coordinating the development plan, they argue. Another interesting approach in this direction is presented by BUUREN et al. (2004). Those authors define a composition operator to generate stakeholder-oriented viewpoints. They formally define the composition operator and present two illustrative examples of its application. This piece of research is based on Bartolomei's thesis (BARTOLOMEI, 2007), FURTMEIER; TOMMELEIN (2010), and BUUREN et al. (2004) to present a set of essential EA layers and components, their relations and derivation operators that can be used to create new EA viewpoints, to be analyzed with mechanisms taken from the DSM and network science literatures. In the next section, we present our research design.

7.3 ARTIFACT´S RESEARCH DESIGN

7.3.1 Research characterization

As discussed in chapter 3, we have adopted the design science research method (HEVNER; MARCH; RAM, 2008). Our method parallels that described by PEFFERS et al. (2007a) and includes the following six steps described in Table 7.2:

DSR Phase	Research contextualization
Identify problem	In our context or environment, (HEVNER; MARCH; RAM, 2008), organizations have significant IT landscapes that make it hard to perform a holistic visual analysis with EA viewpoints that are also underestimated as knowledge sources for EA analysis.
Define solution objectives	Therefore, we aim to increase the analysis capability of EA viewpoints proposing new ones for EANA.
	Continued on next page

Table 7.2: DSR phases and contextualization (HEVNER; MARCH; RAM, 2008)

Table 7.2 – continued from previous page

DSR Phase	Research contextualization
Design and development	We model EA viewpoints as DSMs to benefit from the related analysis techniques such as network analysis. Those are our knowledge base (HEVNER; MARCH; RAM, 2008). Finally, we propose a set of derived EA viewpoints as an artifact.
Demonstration	In section 5, a proof-of-concept demonstration of a subset of the proposed artifact is described.
Evaluation	We evaluate our artifact with experts´ opinion and instantiation in practice.
Communication	We make the research exposition itself in this document (to the academic audience) and the analysis report for the experts (management audience).

Table 7.2: DSR phases and contextualization (HEVNER; MARCH; RAM, 2008)

Following a deductive rationale, we ground our work in the ideas of network science and along with DSMs to design a meta-model of EA viewpoints for performing EA network analysis. More specifically, we derive EA viewpoints, modeling them as DSMs, to propose derived EA viewpoints. We refer to this artifact as EADV-MM.

Next, we perform network analysis metrics such as centrality in these viewpoints built with empirical data provided by three organizations. Our hypothesis is that these new viewpoints are useful and represent new questions that can be used by the experts to perform EA analysis. We use quantitative methods (network analysis) and qualitative methods to evaluate the utility of the meta-model empirically through the analysis of expert opinion. The empirical analysis and validation of the artifact are presented in section 5. In the following, we present the EADV-MM, the used network measures and the data analyzed.

7.3.2 The EA data derivation mechanism

In this section, we review the EA data derivation mechanism and how we relate it with EA viewpoints and network analysis. An important element of a viewpoint is its data-source. We name as primary EA viewpoint, the viewpoint modeled by the researcher directly from the EA components and their connections. For example, in a viewpoint composed only of application components, a generic connection would be “application connects with application”. We represent this viewpoint as “APP x APP”. To build this, the EA data modeler might have extracted the data from documents or tools available at the organization, for instance. Finally, we see this viewpoint as a network. Thus, based on the EA layers adopted in this thesis and discussed in section 2.1, we have a set of primary EA viewpoint candidates, represented by each cell in Table 7.3:

	Value	Business	Information	Application	Technology
Value –VAL	Components of ValxVal				
Business –BUS	Components of BusxVal	Components of BusxBus			
Information-INF	Components of InfxVal	Components of InfxBus	Components of InfxInf		
Application-APP	Components of AppxVal	Components of AppXBus	Components of AppxInf	Components of AppxApp	
Technology-TEC	Components of TecxVal	Components of TecxBus	Components of TechxInf	Components of TecxApp	Components of TecxTec

Table 7.3: EA layers and viewpoint matrix

The above viewpoints cannot be taken as original ones. The five EA layers above are decently covered by the 18 Archimate viewpoints and also by ZACHMAN (1987). For example, the "landscape map" is one of them which covers components from business, application and technology layers. However, since Archimate is mainly concerned with EA modeling issues, it does not specify the use of network analysis to be used in combination with those viewpoints. Despite that, the application of network analysis with viewpoints modeled as networks also have some examples in the literature (e.g. MATHEWS; KAESLIN; RYTZ (2014)), so one cannot claim it as original as well.

In addition to the network analysis performed with the primary data discussed above, further questions can be elaborated with the help of data derivation. We name EA derived data, the artificially generated data based on primary through derivation operators, which in turn generate EA derived viewpoints. With this mechanism, we can now also create a broad range of more abstract models like the one shown in Figure 7.3, without losing a precise definition of the meaning of the models (BUUREN et al., 2004). The derivation mechanism is studied in network science as co-affiliation networks derived from affiliation networks, incidence matrix or bi-partite data (HANNEMAN; RIDDLE, 2005; BORGATTI; HALGIN, 2011). Analogously in DSM theory, derived DSMs can be generated based on DMMs (LINDEMANN; MAURER; BRAUN, 2008). In Figure 7.3, we illustrate the general concept of data derivation for viewpoints. In section 4, we elaborate on this derivation principle to propose a set of original operators to generate derived EA viewpoints.

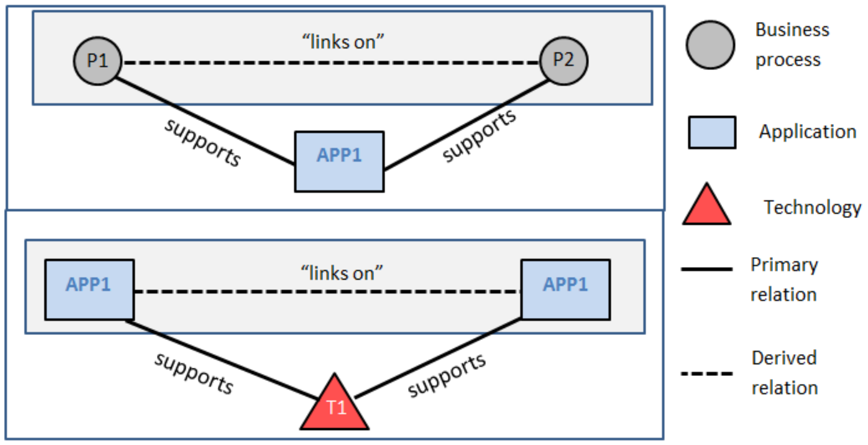

Figure 7.3: General concept of data derivation. Adapted from LINDEMANN; MAURER; BRAUN (2008)

7.3.3 Data collection

We collect EA data from three different organizations presented in section 3.4. We refer to the resulting datasets from these three organizations as Dataset01, Dataset02 and Dataset03 (due to anonymity reasons, we cannot be more specific or provide network figures with named nodes in this document).

Specifically, from Dataset01, we used the primary networks BoxApp and AppxTec as inputs; from Dataset02, the primary AppxTec and from the Dataset03, the BpxBu, BpxBo primary networks.

7.3.4 Analysis methods

We applied visual analysis (RAMOS et al., 2014) and network analysis metrics to investigate the value of the derived EA viewpoints. According to SCOTT (1992), different network measures can be used as proxies for various structural concepts. We use the set of measures described in Table 7.4.

The previous analysis is complemented with experts' evaluation of the utility and applicability of the new viewpoints. The analysis team was composed by one EA expert and one EANA expert for the cases of dataset01, dataset02, and dataset03. For the case of the dasaset03, two business experts were included in the analysis team as stakeholders. In the next section, we present the derivation operators and the EADV-MM.

7.4 DERIVATION OPERATORS AND THE EADV-MM

7.4.1 Derivation operators

In this section, we present four derivation operators which can generate five types of derivations described as follows.

Network measure	Contextualization of the measure
In-degree	In-degree centrality indicates the degree to which components depend on others.
Out-degree	Out-degree centrality indicates the degree to which components provide any support to others.
Betweenness	Betweenness can identify components that work as bridges and connectors [9] in the network.
Closeness	Closeness has a significant effect on how coupled a component is to all other components, not just to those that integrate with it directly (Scott, 1992). Outcloseness, in a directed network, can indicate the range of reachability of a component to the entire application network.
Eigenvector	Eigenvector centrality can be considered an extension of degree centrality that privileges components connected to other well-connected components. Thus, this measure can identify applications that together constitute global structural points.

Table 7.4: Used network measures

- Derivation by affiliation

This operator uses the traditional concept of co-affiliation networks (BORGATTI; HALGIN, 2011) as depicted in Figure 7.4:

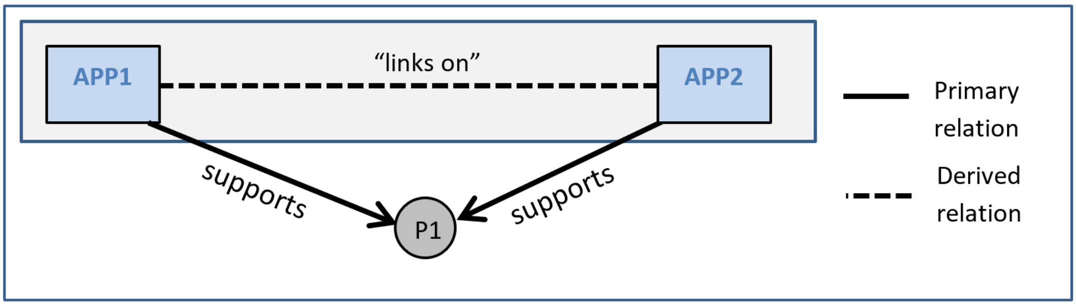

Figure 7.4: Visual representation for the derivation by affiliation.

In this case, the components "APP1" and "APP2" appear connected in the derived network AppxApp since they are co-affiliated with the component "P1" in the original network (AppxBp). Applications that support the same business process must appear clustered in the derived network. This might represent applications which belong to the same line of business, for example. The semantics of the artificial relation ("links on") has to be defined by the data modeler and/or the experts based on their knowledge of the real world (BUUREN et al., 2004).

- Derivation by intra-domain relation or "transfer relation".

In Figure 7.5, the artificial relation (dashed line) is created in layer A (e.g. business) based on an existent internal relation (double line) in layer B (intra-layer relation), and at last, thanks to another relation between the layers A and B (inter-layers relation). The dashed line is created based on the continuous lines which connect the component “BU1” (business unit one) with component “APP1” (application one); and “BU2” with “APP2” (both are inter-layers relations); and also based on the double line which connects applications “APP1” and “APP2” (intra-layer relation).

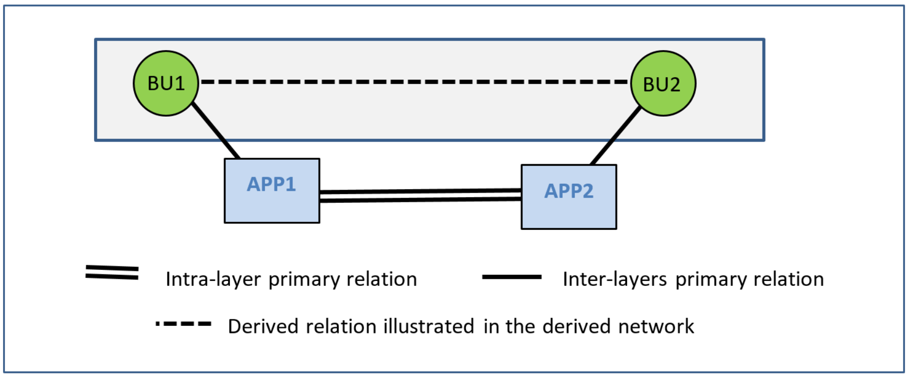

Figure 7.5: Visual representation for the derivation by intra-domain relation.

With this operator, we “transfer” or “project” an existent relation from layer B (Application) to layer A (Business). One possible aspect to be investigated with this new network configuration could be the information flow among business units supported by IT applications.

- Derivation by transitivity

Based on the concept of transitivity, this operator has its function illustrated in Figure 7.6:

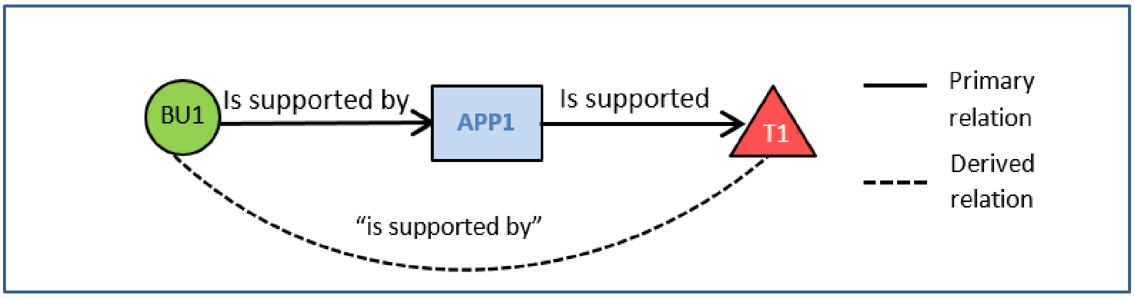

Figure 7.6: Visual representation for the derivation by transitivity.

The same concept is explored by BUUREN et al. (2004) labeled as “composition of relations”. The derived network might be used, for example, to graphically describe the traceability of the use of a specific technology across the entire company or by business units. In addition, network measures, such as degree and eigenvector centralities, can be applied to this derived network configuration as well.

- Derivation by attribute

In this case, we use more information than only the component itself and its relations to generate a derived network. For instance, we can have a unimodal network “TecxTec” and its component attribute “vendor” as depicted in Figure 7.7:

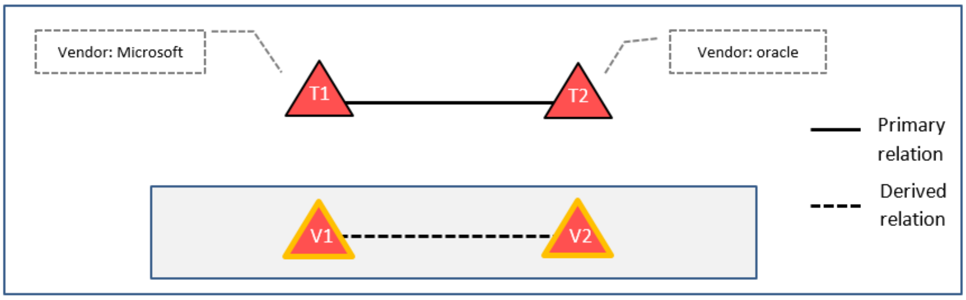

Figure 7.7: Visual representation for the derivation by attribute.

We decide to abstract the technology component itself and represent it only by its attribute “vendor”. Thus, it is possible to derive a new network labeled “VendorxVendor” from "TecxTec" network, considering the attribute “Vendor” of the technology component. In this case, instead of considering a relation between “T1” (e.g. an application server) and “T2” (e.g. a database server), we are looking thus for a relation between two technology vendors (e.g. how components of Microsoft and Oracle interact in the enterprise´s technology architecture). In the end, the derivation process builds a network of vendors, which could be further analyzed in terms of dependence or dominance of a particular vendor at the organization. In a similar way, we may apply the same operator with an “AppxTec” network, generating an “Appx Vendor” network. Thus, considering this operator, the resulted networks might be unimodal or bimodal according to the primary networks used as input.

At this point, we discussed some transformations that can be done with EA data already available (primary data). The four derivation operators and their respective transformation are summarized in Table 7.5.

The operators can be used in combination. For instance, the resulted network for the transformation sequence ↑(≡(BuxAppxTec) ,TecxTec) is a BuxBu unimodal derived network designed to reflect how business units relate considering their technology dependence. The analysis of each derived network needs suitable contextual information for the characterization of the new edges, nodes, subsets, and the entire network. For all operators, the derived relations among components need to be interpreted based on the architectural semantics, i.e. on our knowledge about the real world (BUUREN et al., 2004).

The underlying idea in applying this set of operators is that these new data configurations and interpretations will allow experts ask new questions and get new answers from the data already available at their organizations and leads to suggestions for improvement of the original EA viewpoints. In section 5, we chose some derivation operators to apply on three empirical datasets.

	Type of original network (input)	Type of derivation	Operator symbol	Example	Derived network (output)
1	Unimodal	Derivation by attribute (Type 1)	Φ	Φ(Tec x Tec)	Unimodal derived network (e.g., VendorxVendor)
2	Bimodal	Derivation by attribute (Type 1)	Φ	App x Φ(Tec)	Bimodal derived network (e.g. AppxVen)
3	Bimodal	Derivation by affiliation (Type 2)	≈	≈ (AppxBu)	Unimode derived network (e.g. BuxBu or AppxApp)
4	Bimodal and Unimodal	Derivation by intra-layer relation or transfer relation. (Type 3)	↑	↑ (AppxBu, AppxApp) or ↑ (AppxBu, BuxBu)	Unimodal derived network (e.g BuxBu or AppxApp)
5	Multimodal	Derivation by transitivity or composition (Type 4)	≡	≡ (TecxAppxBu)	Bimodal Derived network (e.g TecxBu)

Table 7.5: Operators used to derive EA network data

7.4.2 A meta-model for derived EA viewpoints

The EADV-MM has been created in the form of an MDM as described in Table 7.6. The layers and components used for the creation of the meta-model are the same adopted in this thesis, which were defined observing the components and layers targeted by researchrs in the EANA SLR. In the end, all of them are comprised by Archimate language. As we said before, several works already proposed primary viewpoints for EA. However, we advance on that, proposing derivation mechanisms applied over those primary viewpoints and, in addition, we discuss how to perform EA structural analysis with these new "data sources".

We group some possibilities of EA derived viewpoints in Table 7.6 which represents our partial EADV-MM instantiation. It has to be clarified that the presented list is not exhaustive. Each cell in the EADV-MM represents a subset of the entire EA, specifying a view, a method of analysis, which should be useful for a particular kind of stakeholder. Potentially, for each cell in Table 7, we can derive several DSMs and DMMs, considering the application of different operators. To illustrate this, we pick some cells of Table 7 to show the derivation mechanism embedded in each operator. First, at the top of Table 7.6, the network "BpxBp" is resulted from the application of the affiliation operator (≈) over the DMM "GoalxBp" (symbolically, ≈(GoalxBp)). Generally speaking, in this new DSM, business processes which support the same business goals appear linked to each other, even if in the original BpxBp DSM they did not have an operational linkage, therefore, having a new semantic for the relation.

	Value	Business	Information	Application	Technology
Value–VAL		GoalxGoal from ↑(GoalxBp, BpxBp) BpxBp from ≈ (GoalxBp)			
Business –BUS		BusxBus from (≈BuxBp)		BusxBus from (≈BuxApp)	BuxBu from ↑(≡(BuxAppxTec), TecxTec)
Information-INF				BoxBo (from ≈BoxApp)	
Application-APP					AppxApp from ↑(AppxTec, TecxTec)
Technology-TEC				TecxTec from ↑(TEcxApp, AppxApp) TecxTec from (≈TecxApp) (technology ecosystem) AppxVendor (from Appx$F$Tec)	TecxVendor (from Tecx$F$Tec) VendorxVendor (from FTec x FTec)

Table 7.6: The EADV-MM partial instantiation

The same operator is applied in the second example, resulting in the DSM BusxBus (resulted from ≈BuxBp). This new DSM takes into consideration the interaction between the business units with business processes to link business units which are responsible or use the same set of business process. In this case, the higher the number of business process shared by a pair of business units, the higher the weight of the interaction between these two business units. This could be numerically demonstrated in the DSM, or visually in a graph with a thicker edge between those particular two components. The third example uses the transfer operator (↑). In this case, we have one DSM and one DMM as inputs (AppxApp and TecxApp, respectively) to produce a TecxTec DSM. In the case we do not have a map of the technology dependencies, one could be inferred, considering the relations among the applications and projecting them to the technology layer. Finally, we used the derivation by attribute operator (Φ) to produce a DSM of technology vendors (VendorxVendor), using as input the DSM TecxTec and its component attribute “vendor”.

Those transformations described previously are just a few examples of what can be achieved applying the operators. Since each cell corresponds to an EA viewpoint, a number of analysis perspectives can be drawn to support and potentially improve the EA analysis (VAKKURI, 2013; SCOTT, 1992). For instance, one might want to investigate the influence of a particular vendor or technology in the technology ecosystem (technology layer) or a ripple effect provoked by a component change. A second analysis perspective is to group or cluster components according to their inter- relations. A cluster of technologies or vendors may be identified in order to understand their relations and influences in the internal software ecosystem. As a third one, either as a visual analysis or as performed by LAGERSTRÖM et al. (2014) with their "hidden structure method", we might want to analyze the distribution of components over the network in order to understand dependencies and impacts of changes, for example.

The choice of the analysis perspective and its analysis method for each derived EA viewpoint must be decided case-by-case, taking the semantic of components and relations and stakeholders 'concerns into account. Following our EANA-MM, for each cell in the EADV-MM, the EA analyst must specify the following decisions: the concern to be analyzed; a meaningful operational question to be answered with the viewpoint´s support; a metric/method/model to be applied (e.g visual analysis, centrality degree or clustering algorithms); modeling choices about the edge´s semantic, weight or direction; the network level of analysis (component, group or network level), the EA layers targeted (one of the five considered EA layers presented previously), used components´ attribute(s); the type of network (or matrix) unimodal/bimodal (DSM/DMM); the type of used data (in this research, derived data), and finally, the stakeholder or potential user for the derived EA viewpoint. The set of cells depicted in the EADV-MM, together with the derivation operators may result in several derived viewpoints possibilities, which in turn, can be designed following these modeling decisions specified by our EANA-MM. Altogether, those three elements work like parts of a strategy to derive data for EA network analysis.

In this section, we presented some approaches For EA data derivation. As a proof-of-concept for the derived viewpoints meta-model, we perform a structural analysis over four derived viewpoints in section 5.

7.5 Empirical evaluation of the artifacts

Our analysis started using the derived EADV-MM to brainstorm various analysis scenarios during two discussion sessions of around one hour each, with EA practitioners responsible for the three datasets presented in section 3.4. The choice of which derived viewpoints should be used was made considering the data availability for our research. For instance, we did not have access to data related to the EA value layer. We divide the derived viewpoints analysis according to the three used datasets as described in the following.

7.5.1 Using Dataset01

Business Objects (BOs) are data entities processed or manipulated by applications during their business processes execution. The first organization (Dataset01) provided us with information about all BOs and their respective supporting applications. At first hand, we did not have data related to direct interactions among the BOs, and we wanted to check how these entities were indirectly related considering the technical support given by applications.

Therefore, we applied the derivation by affiliation operator (≈) with the BoxApp network as input, resulting in a BoxBo derived network. The derived BoxBo has 127 BOs and 1440 relations with the following semantic: if the same application APP1 supports two or more BOs, then these BOs must appear connected in the derived BoxBo network (derivation by affiliation).

Considering this network, the expert can visually analyze the level of modularization of Bos. Each cluster of BOs represents a set of interrelated BOs which are supported by the same set of related applications. In the ideal modularization level, related BOs should appear clustered in the derived network and the network topology would have several isolated clusters, each of them representing Bos manipulated by an application module.

The generated BoxBo network is depicted in Figure 7.8, produced by the "circular layout for groups" algorithm from ORA tool[2] (23 isolated BOs were removed). In Figure 7.8, one can notice the presence of several clusters formed by BOs which are manipulated by the same set of applications. We asked the EA expert of the related company to comment on those clusters during a specific workshop for presenting the analysis results. According to him, the visualized component clusters were expected and are overall aligned with the desired design for the applications´ modularization.

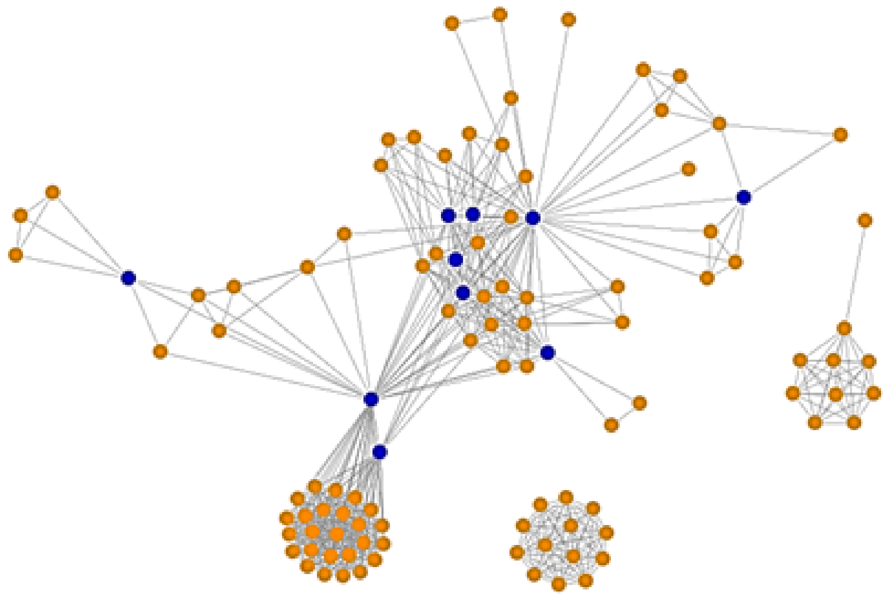

Figure 7.8: BoxBo derived network and highlighted betweenness outliers.

Some entities will eventually be manipulated by more than one application of different modules. In this case, at the component level, the betweenness centrality becomes an important

[2] http://www.casos.cs.cmu.edu/projects/ora/

indicator of components that act as brokers between clusters. The blue colored components in Figure 7.8 represent the top 10 betweenness values in the BoxBo derived network. According to the expert, the highlighted individual components also represent important BOs for the EA that must require special managerial attention.

Continuing with the Dataset01, we apply again the affiliation operator in the AppxTec network which has 183 components to derive a TecxTec network. The idea behind the design of this derived network is that every time we have a set of technologies working to make an individual application works, those technologies together constitutes a local tech ecosystem. If we look at the whole application architecture, those local tech ecosystems may or may not be reinforced by other applications in the network. In the end, we can build a visual map of the technologies and check how they are related to each other, what we call the technology ecosystem of application architecture. Analyzing Figure 7.9 at the network level, it is possible to identify a very dense network, with a core-periphery architecture where a bunch technologies shape the core such as MS Windows SQL Server, MS Windows Server, Oracle, C, among others (blue nodes). At the core of the network, there is no predominant technology stack (i.e. a set of components or layers in a software offering that provides broad functionality), instead, a diversity of technologies connected to each other.

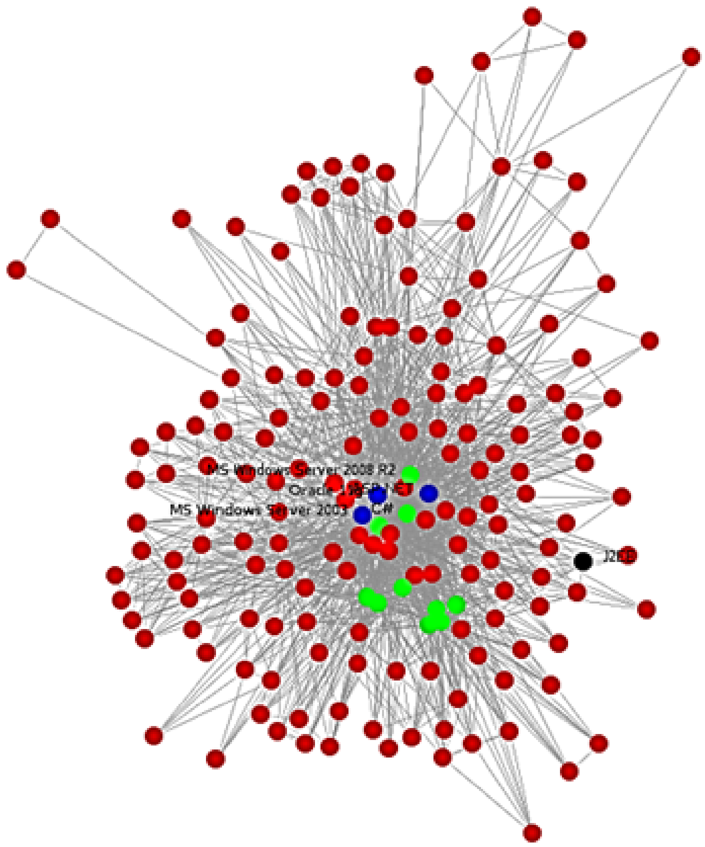

Figure 7.9: TecxTec derived network and highlighted eigenvector outliers.

However, we can find some clusters located in the periphery. One of them, at the right part of the network, contains the J2EE component (a black node), which is linked to other commercial related technology components such as TopLink, GlassFish, EntireX,Oracle, LDAP/ADS, LOG4J, SOLARIS and Java. This association is a technology stack commonly found in the market and also represented here. We call this association a local tech ecosystem.

Still, in the network periphery but less connected, we can find less used technologies like Coremedia, Crystal reports, Business communication server and others. However, it is not possible to make any qualitative assumptions about the value of the peripheral technologies (i.e. if they are critical despite not well connected to other applications) only with structural data. Regarding this issue, the expert suggested to combine the structural information with the expert knowledge in future analyses of the role of those technologies in the architecture.

We also apply the eigenvector centrality to identify the top 10 most structurally well-connected components in the overall network (technologies that are very used and that are connected with other well-used technologies). These components are Host, z/OS, Client/Server, Natural, ADABAS, MS Windows Server 2008 C,VB.NET,.NET Framework 3.5, .NET Framework 4.0; and are shown colored in green in Figure 7.9.

7.5.2 Using Dataset02

Regarding Dataset02, we analyzed a VendorxVendor network derived from a TecxTec one. In this case, we apply the derivation by attribute operator (F). The TecxTec network has 169 components. With this network, we explore the vendors' relationships, important information for technology migration strategies adopted by the company. With this sort of vendors´ map, the company can follow the market's moves of vendors, to monitor external technology discontinuities and updates that could impact the organization´s agility to change. The VendorxVendor network has 27 components and is depicted in Figure 7.10.

We also analyzed the frequency of vendors' association (e.g. how often Apache technologies are used in association with IBM ones). The results are shown in the heat map depicted in Figure 7.11. The dark blue cells, mainly located at the bottom of the matrix, represent the high frequency of associations between a pair of technology vendors. According to the experts, the highlighted components describe the importance of key vendors in the analyzed EA.

Heat maps (DSMs), network visualization algorithms and network centralities act complementary in this case, highlighting technology associations that might be useful for experts in the migration platform programs, for example.

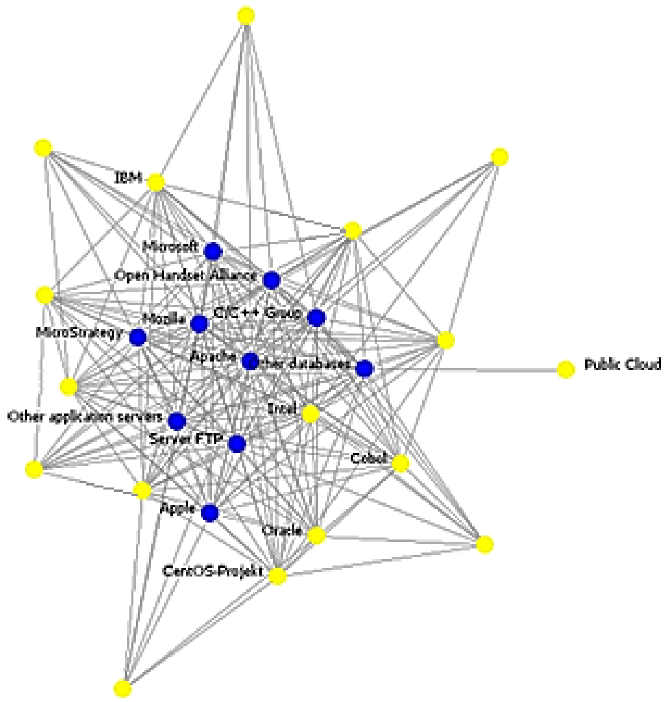

Figure 7.10: VendorxVendor derived network and eigenvector outliers highlighted

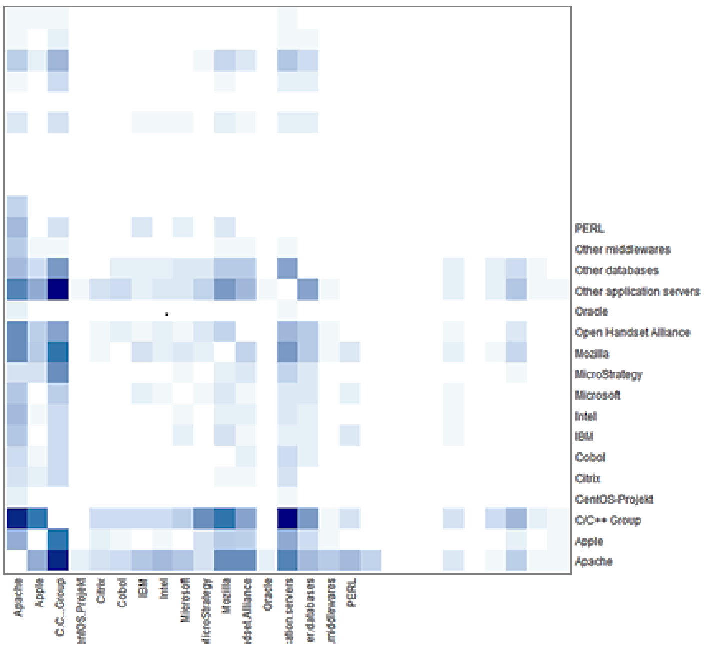

Figure 7.11: VendorxVendor heat map

7.5.3 Using Dataset03

Organization business units (BUs) execute different workloads depending on each process phase. We consider two phases for analysis purposes: PROD (phase I) and KONZ (phase II), and we pose this operational question: can we identify key structural BUs considering the business process interaction in different process phases? Thus, our analysis concern here is to find the key BUs, which might be important when it comes to involving the right people (BUs) in the decision-making process about the business process changes. As input, we took the BuxBp network and applied the transfer operator(↑) to generate a derived BuxBu network. With this derivation method, if two processes BP1 and BP2 are connected, an artificial connection is created between their respective BUs in the BuxBu network. The hypothesis H1 formulated by the experts is then broken down as follows.

H1.1: In Phase I, the focus of the project management unit (BU1) should be fairly continuous as they manage all activities. H1.2: In Phase II, the focus will be more on the technology people, with a ramp up to production and logistics and possibly to purchasing. H1.3: Overall, in Phase II, design engineers will be fairly central, as they function as a sort of information hub around which all technical concept design focuses.

As we did with Dataset02, we combined two types of network analysis outputs: the network metrics´ rankings described in Table 7.7 and the heat maps depicted in Figure 7.12. With the heat map, it was possible to check the high intensity of the information flow from B1, BU5, and BU7, which is spread out in Phase I. There was intense activity inside BU1, as can be seen in the dark blue cells, confirming the importance of BU1 for Phase I (thus supporting H3.1).

Figure 7.12 also suggests a strong interaction from BU5 to BU1. This might confirm that both BUs together are the most active BUs in Phase I in terms of process interactions. From Table 7.7, we notice that BU1, BU5, BU2, and BU7 also appear in different rankings of network measures, reinforcing our visual analysis results from Figure 7.12.

TOP Out-degree BUs		TOP eigenvector BUs		Most recurrent BUs	
BU5	Product management	BU1	Project management	BU1	Project management
BU1	Project management	BU5	Product management	BU13	Prod. management
BU7	Total vehicle integration	BU2	Controlling	BU2	Controlling
BU3	Quality	BU7	Total vehicle integration	BU5	Prod. Management
BU2	Control	BU13	Production preparation	BU7	Total vehicle integration

Table 7.7: Network analysis at the component level for BuxBu Prod (Phase I)

For H1.2, we obtained the following results: BU13, responsible for production aspects, became imperative in Phase II (detected by high in-degree centrality components and eigenvector centrality); Purchasing (B15) had importance detected by high in-degree centrality and also was among the most recurrent outliers; Validation and integration (BU7, BU10, BU3) aspects received focus in Phase II, detected by eigenvector centrality, most recurrent outliers, and out and in-degree centralities. Although identified by the experts a priori, logistics did not appear as a focus in Phase II. In conclusion, we found H1.3 to be partially supported.

For hypothesis H1.3, integration, validation, and preparation for production activities were the main BUs in Phase II. So, we can conclude that H1.3 was also supported.

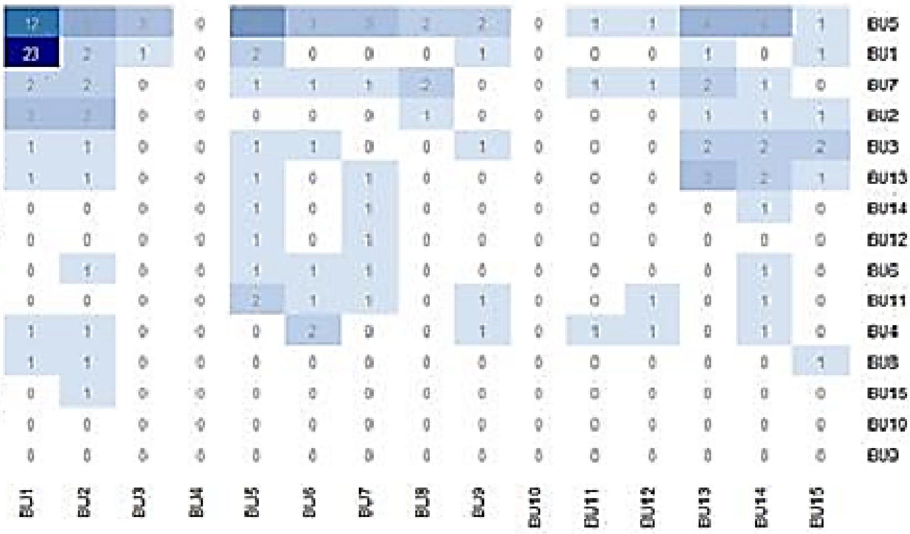

Figure 7.12: Heat map for BuxBu network

In Table 7.8, we summarize the analyzed viewpoints, targeted concerns and respective analysis methods according to the taxonomy defined by the Santana A.; Fischbach K.; Moura H. (2016a).

In the end, for each of the three analyzed datasets, we generated different derived viewpoints, composed by different components and relations, and therefore, with different analysis concerns. These new viewpoints together with the applied analysis methods represent additional perspectives of analysis for the EA experts. In the next section, we present our final considerations.

7.6 Chapter Summary

In this chapter, an EA derived viewpoint meta-model (EADV-MM) and four data derivation operators were proposed. The operators transfer concepts already applied in network science and DSM body of knowledge for the EA analysis context. With the approach suggested in this chapter, it is possible to derive cross layer viewpoints, amplifying the analysis possibilities for EA experts. In sections 4 and 5, we presented and applied the derivation operators instantiating them with a subset of the meta-model, resulting in those four derived networks. In the present concrete case, the relations between components of BoxBo, TecxTec, VendorxVendor and BuxBu networks, implicit in their original networks, were created to offer some new analysis perspectives.

We highlighted some individual components and emerged clusters which were evaluated by EA experts related to a dataset. We claim that the presented derivation rules can be used for

Derived network	Analysis concern	Operational question	Analysis methods
Unimodal BoxBo derived from ≈(AppxBo)	Application's modularization of Bos	Are related BOs manipulated by a specific and modularized set of applications?	Visual cluster analysis (cluster level) and betweenness centrality (component level)
Unimodal TecxTec derived from ≈(AppxTec)	Technology ecosystem; key technologies	What are the key structural technologies that support the application architecture and how they are related with each other?	Eigenvector centrality (component level), topological network analysis (network level)
Unimodal VendorxVendor derived from ΦTecx ΦTec	Vendors ecosystem; key vendors	What are the key structural vendors and what is the frequency of their associations in the technology ecosystems?	Visual analysis (network level), eigenvector centrality and heat maps (component level)
Unimodal BuxBu from ↑ (BuxBp)	Evolution of the role of BUs along the business process execution.	What are key structural BUs considering a specific business process interaction set?	Heat maps and network analysis metrics

Table 7.8: Network analysis at the component level for BuxBu Prod (Phase I)

the creation of "non-standard/implicit" stakeholder-oriented visualizations. The case of BuxBu network is a clear example of the usefulness of derived data: the company did not have the relations between the BUs mapped at first hand but only the relations between BUs and the BPs on which the BUs take part, and the network of BPs itself (BuxBp and BpxBp networks). Nevertheless, despite the absence of the primary data about the BUs, it was possible to confirm experts 'perceptions and highlight key structural BUs during the business processes execution. Less pragmatic but still intuitive, the derived TecxTec and related networks bring to the analysis the relations among technologies and among vendors and may help understand their influence and evolution in the enterprise in a time series fashion.

Despite the new insights generated based on the derived data, the importance of combing the structural knowledge with the expert tacit knowledge to improve the methods´performance was also mentioned by the experts.

Despite the high number of possible combinations of viewpoints allowed with the EADV-MM and operators, we only explored a subset of them. The list of derived viewpoints presented is not exhaustive. It is important to clarify that EA data modelers still have to define the meaning for the derived relations of each particular derived viewpoint. In general, the experts need to analyze the potential of those derived components and relations further, making sense of them and validate their utility and implications for EA managerial reality. This problem should be approached by testing alternative viewpoints, measures and visualization methods, and seeing which has the best "value" for the EA analysis (SOSA; EPPINGER; ROWLES, 2007). This can foster further research using the respective DSMs and DMMs together with network analysis, to support the development of EA analysis tools. The manual and time-consuming information collection process employed is a downside of our approach. If it can be done automatically, then this approach is very viable, in our opinion, especially if standard model languages such Archimate could be automatically translated to this more generic modeling approach (DSMs).

8

Cognitive-structural diagnosis analysis and attribute check analysis methods

In this chapter, we present our last two DSR contributions. We also describe the context and motivation (section 8.1), artifacts' research design (section 8.2), the two proposed methods (section 8.3) and finally, the results obtained with the application of the artifacts with three empirical datasets (section 8.4).

8.1 Context and motivation

As discussed in Section 2.1.5.1, BUCKL et al. (2009) present a schema for EA analysis classification with five dimensions: body of analysis, time reference, analysis technique, analysis concern, and self-referentiality. With the "analysis technique" dimension, they define three categories: expert-based, rule-based, and indicator-based. They point out that the expert-based analysis techniques depend on the experience and expertise of the executing person. That is, one or more experts – the enterprise architects – analyze properties of the EA along appropriate architecture views, such as, for example, reports or graphical visualizations. BUCKL et al. (2009) found only two papers in this category, which they discuss in their literature review. In these two cases, the experts analyze the EA models based on their tacit knowledge and used no analytical tools or other knowledge sources in drawing their conclusions. This scenario is depicted in cycle I in Figure 8.1. In addition to the expert-based techniques, BUCKL et al. (2009) also discuss two other types: rule-based, performed at an increased level of formalization and potentially automated (e.g., EA analysis based on ontologies); and indicator-based, used to assess quantitatively architectural properties such as complexity or costs. The authors make no mention of the possibility of simultaneously combining two types of analytic techniques, as we aim to do in this chapter. This is essentially what is done in the cycle III, after the cycle II modeling of EA as a network. Considering the EA network analysis paradigm only, there were no analysis strategies that combined at the same time expert (i.e., knowledge about components) and structural information (Santana A.; Fischbach K.; Moura H., 2016a). This analysis scenario is represented in cycle IV in Figure 8.1.

The goal with the expert combined EA analysis is to minimize the subjectivity of informal

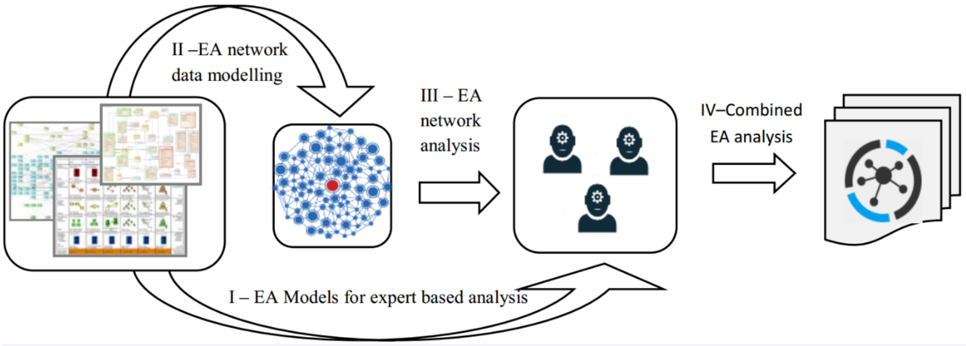

Figure 8.1: Combined EA analysis. Cycle I: EA models are analyzed with expert-based techniques; Cycle II: EA models are converted into EA network data; Cycle III: EA network analysis is performed using network measures; Cycle IV: Combined EA analysis

analysis done by experts by combining it with components' structural information (indicator-based analysis). To that end, we introduce and evaluate two methods for EA network analysis at the component level: cognitive-structural diagnosis analysis and attribute checking. We use cognitive-structural diagnosis analysis to provide complementary information for expert-based EA assessments. With attribute checking, we make considerations about components with certain attributes of interest that occupy important structural positions. With the introduced methods, we aim to answer the following research question: How can we combine expert information with network measures to support EA analysis?

As contributions, We aid cognitive-structural diagnosis analysis by assessing the architect's perception of components together with network measures' output, which would allow experts to validate or refine their knowledge about their EA components.

8.2 Artifacts´research design

We use the outputs of applying quantitative measures from network science along with information provided by EA experts from three sample cases to design two artifacts for performing EA network analysis based on the combined data. More specifically, we compare the main components identified by the measures with those pointed out manually and subjectively by the experts. We refer to this artifact as cognitive-structural diagnosis analysis. Our hypothesis is that network measures will confirm the majority of components perceived as important by the experts and thus validate their perceptions, while also shedding light on potentially important components that were at first overlooked by the experts. This method is explained in Section 8.3.1. We also operationalize a method to verify which kinds of components occupy important structural positions in the network, while observing these components' features. We call this second artifact attribute check analysis. Both artifacts are validated empirically. The rationale behind these two methods is that when an enterprise has a set of applications like the one depicted in Figure 8.2, which represents the architecture of one of our cases, it may be intuitive for a person to point out some applications based on their license´s price, number of users, or any other intuitive heuristic. However, as can be seen from the colored components, different centrality measures can highlight different structural aspects that are not possible to detect simply by "eyeballing" the network. We believe that this additional source of information may generate important insights for the expert.

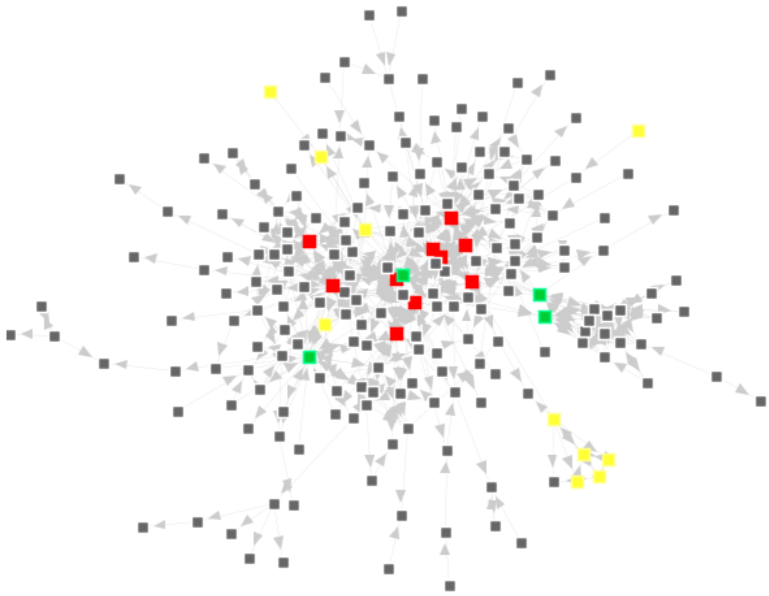

Figure 8.2: An application architecture modeled as a network (red nodes represent the ten highest eigenvector centrality values, yellow ones have the ten highest closeness, and green ones have the 10 top betweenness values).

The analyzed datasets were already showed in chapter 3. The dataset01 and dataset02 come with the same semantics for the relations between the applications, which is that one "application provides-data-to another application." As depicted in Figure 8.3, if an application A sends data to a second application B, we link A with B, using a directed and unweighted relation from A to B. The first two cycles of the EA combined analysis proposal (see Figure 8.1) are related to external knowledge acquisition and network data modeling activities, respectively. Regarding the first cycle, we obtained the data with the attributes' values (and components' inter-relationships) in datasheets.

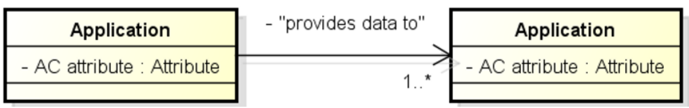

Figure 8.3: Semantic adopted for modeling the Dataset01 and Dataset02

In Dataset01, we refer to "type of development." Its possible values are "standard" (representing purchased commercial software), "self-developed" (internally developed applications), and "others." In Dataset02, "core application" is a qualitative attribute that materializes the expert´s perception about whether an application belongs to the core set of the organization's applications and thus has top priority from a business point of view. In the middle column, we describe the ratio of maintained values for each component attribute or relation.

Subsequently, we modeled manually the relations among applications with the help of the ORA tool. In general, this task might benefit considerably from a software plug-in capable of converting data from architecture models to networks. Once the network models were finished, we used the UCINET software to run the chosen network measures in Cycle III.

Table 8.1 shows the set of network measures we use, chosen according to two criteria: (1) they are well-known centrality measures often employed in related studies; and (2) they address different aspects of the network. FREEMAN (1978) argues that centrality degree, betweenness, and closeness are related to various aspects of centrality such as control, brokerage, or independence. Eigenvector centrality is the fourth measure of centrality. Hence, the measures of centrality then offer four different ways of identifying how network structure might differentiate the roles of components. Our two networks are two unimodal networks formed by (only) applications components. We used the UCINET software to calculate those measures with our data.

Network measure	**Contextualization of the measure**
In-degree	In-degree centrality (applicable as our networks are directed ones) indicates the degree to which applications depend on data provided by others.
Out-degree	Out-degree centrality (again, applicable as our networks are directed ones) indicates the degree to which applications provide data to others.
Betweenness	Betweenness can identify applications that work as bridges and connectors [9] in the application network.
Out-closeness	Closeness has a significant effect on how coupled a component is to all other components, not just to those that integrate with it directly [8]. Outcloseness, in a directed network, can indicate the range of reachability of an application to the entire application network. We apply out-closeness in our datasets.
Eigenvector	Eigenvector centrality can be considered an extension of degree centrality that privileges components connected to other well-connected components [9]. Thus, this measure can identify applications that together constitute global structural points.

Table 8.1: Network measures used in this research and their contextualization to the case studies

8.3 Cognitive-structural diagnosis analysis and attribute check analysis artifacts

We present a new category for EA network analysis: the expert-combined analysis. In addition, to demonstrate the idea behind the category, we propose two methods – explained as follows – that both try to take advantage of combining structural and expert knowledge.

8.3.1 Cognitive-structural diagnosis analysis

As pointed out earlier, EA analysis still depends strongly on human expertise due to the lack of development of proper tools. In the case of vast IT landscapes, human perception – which is limited and subjective – might decrease the confidence in the quality and coverage of EA analysis. Applications that support strategic business processes, for example, are natural candidates to be listed by the experts as important EA components. Experts might not notice other applications that do not fit these criteria, despite that the applications play important structural roles in the information flow. Hence, considering EA analysis as a complex process, one might want to use additional knowledge sources in order to add confidence to the analysis. Combining the information generated by applying network analysis measures (structural criteria) together with expert knowledge (subjective, in nature), we believe we can design a more robust method for EA analysis. One important assumption for this method is that network analysis and expertise are both important factors for the decision of identifying critical components.

To apply this method, given a component C (e.g., application, business process), we first identify an Attribute of Interest (AI) (e.g., application's availability, business process strategic value) or any other EA analysis concern for this component. Second, we store the network measure's outputs for the components in the matrix Network Measure Value (NMV) (e.g., eigenvector and degree centrality values for all the components). Thus, we investigate whether there is a correlation, at least at a moderate level (rho >0.3, Spearman correlation, p<0.05), between AI values and those in the NMV matrix. If there is no correlation, one should review whether this is due to the fact that the given attribute of interest is naturally independent of the component's structural position or that the expert might have wrongly missed including a component's structural context in the analysis altogether. If there is at least a moderate correlation, we select the number nc of components with the highest NMV values (we call these outliers, which form a set with size=nc).

From this outliers set, the goal is to find the components that might validate the expert's perception about "important components" (true positives) and/or add new ones not perceived previously by experts (false positives). Thus, we list those components for further evaluation and analysis report. Typical questions that may be answered with this method are:

- Can the "x" highest-ranked network measure values (outliers) identify the "y" most-important components as per a certain attribute (like strategic value)?
- Can the "x%" highest-ranked network values identify the "y" components with "high" and "very high" availability (attribute) values?

For both questions, two parameters need to be defined. One is the amount "x" of selected components from the outliers' set to be used in the cognitive-structural diagnosis analysis. "x" can be an exact amount or a percentage of the outliers' set ("x%"). The second is "y": the amount of components from the set of components of interest (components already identified as important ones by the expert). Thus, we compare the subsets of x and y components and compute recall and precision values, which results in the four areas Figure 8.4 depicts.

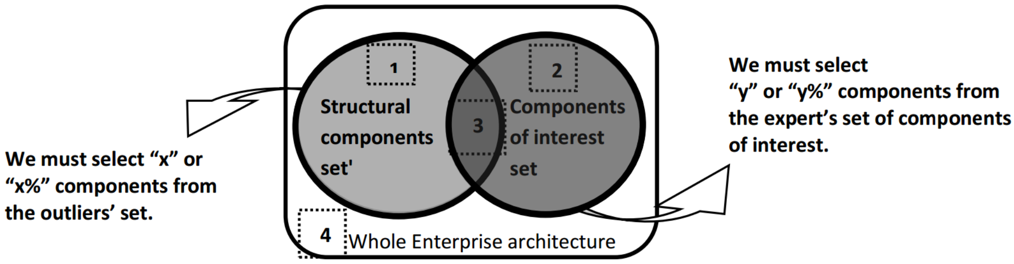

Figure 8.4: Cognitive-structural diagnosis analysis: defining the samples to compute recall and precision

Precision and recall are not considered as the central performance indicator for the method, but they are used to acknowledge that network measures overlap with the expert knowledge at a certain level. However, we are more interested in the groups generated with the F-measure: true positives, representing reinforcement of the components' importance from a structural and expert-based (non-structural) perspective (Figure 8.4, Area 3); false positives (Figure 8.4, Area 1); and false negatives (Figure 8.4, Area 2).

The true positives represent components classified by experts as "strategic," "critical," or any other intrinsic value of interest that had its structural importance confirmed. The false positives represent components with no particular value for the attribute of interest (e.g., "core application" = no) but detected by network measures as having structural importance; they must be discussed further to review and validate the expert's classification with respect to the detection of this structural importance detected (which may represent valuable information). The next group represents components with no structural relevance to the network measures but is important for the experts for different reasons (false negatives). Table 8.2 contextualizes the recall and precision results. The pseudo-algorithm in Figure 8.5 summarizes the method.

Attribute: Business Criticality; Values of interest: Very critical and Critical Used network measures: e.g.: outdegree and betweenness		
	Important to the expert	**Not important to the expert**
Structurally important	True positives	False positives
Not structurally important	False negatives	
Recall =	How many relevant items are selected?	
Precision =	How many selected items are relevant?	

Table 8.2: Outputs of the cognitive-structural analysis method

Cognitive-structural diagnosis analysis – algorithm

```
Set m = number of applied network measures
Set c = number of components in the application architecture
Set y = number of components of interest (classified by the experts)
Set NMV_{c,m} = the matrix to save m network measures values for all the c components
Set CNMc = the list of network measures correlated with an attribute of interest
For each network measure NM_i (NM_1=eigenvector, NM_2=closeness centrality,…, NM_m)
  For each component C_j (C_1=APP_1, C_2=APP_2…, C_n=APP_c)
   NMV_{i,j} ← value of NM_i for C_j
From the component type C, select an attribute of interest AT_C
For each network measures NM_{i(i=0, m)}
  If AT_C and NMV_i are correlated (Spearman) then
    CNM_c ← CNM_c + NM_i
Calculate Recall and precision (set of "y" components of interest and set of "x" or "x "%-highest outliers for each CNM)
Return true positives set, false negatives set and false positives set
```

Figure 8.5: Cognitive-structural diagnosis analysis method – pseudo algorithm

8.3.2 Attribute check analysis

Our assumption for this method is that combining each network measure's contribution will allow us to identify components that play the different strategic roles indicated in Table 8.1. In this sense, the goal of this method is to take the outliers among the components (i.e., those with highest values for any centrality measure) and check whether they have specific attributes or features of interest, such as type of software license (e.g., SaaS, self-developed), type of application hosting (e.g., cloud, internally, etc.), type of technology (mainframe, client-server, distributed, etc.), vendor (e.g., IBM, Oracle), and so on (see Figure 8.6, which presents the algorithm for the attribute check analysis method). Any of these features might influence an expert in the case of changes, acquisitions, mergers, and so on.

Attribute check analysis – algorithm

```
Set m = number of applied network measures
Set c = number of components in the application architecture
Set y = number of components of interest (classified by the experts)
Set NMV_{c,m} = the matrix to save m network measures values for all the c components
For each component Ci (i=0, nc)
For each measure NMj (j=1,m)
   NMV_{c,m} ← Value of NMj for Ci
Calculate the x or x% of outliers from NMV_{c,m}
From the component type C, select an attribute of interest AT_C
Check the particular values for ATc among the outliers
```

Figure 8.6: Attribute check analysis method – pseudo algorithm

More specifically, this method would answer a typical question such as: How are “individual” (self-developed) and “standard” (procured) software distributed over the most important (i.e., top 10%) structural components? One could also use this method to verify the presence or dominance of a specific software supplier in key structural applications. Another possibility would be to create triggering events that might warn experts when key components assume specific values for a target attribute (e.g., availability).

In general, this method is suitable for contexts in which the attribute of interest does not have at least a moderate correlation with the network measures. The method’s output is descriptive and qualitative information that may enhance the expert’s knowledge about the EA and support architects in their decision processes. In the end, though, the conclusions are the responsibility of the expert.

All in all, we have introduced two exploratory and generic approaches that might be applied to any set of network measures and any target attribute when the expert needs or wants to compare his tacit knowledge with the structural dimension of the architecture. It is worth mentioning that past work using network analysis did not explore this potential, focusing mainly on the pure structural value of the EA components.

8.4 Empirical evaluation of the artifacts

8.4.1 Cognitive-structural analysis

As results of the cognitive-structural diagnosis analysis of the availability concern, it was possible to identify that the five components pointed previously out by experts are related to important structural positions in the application architecture(see Table 8.3). In addition to that, given their structural importance, four additional components were suggested for further analysis. With this information, architects might become aware that there are four components with occasional downtimes (since they have only moderate availability) that are well connected and may thus affect other components in case of failure. The expert confirmed this insight, for instance, in the case of an application that manages accounting-related data supplied on a daily basis to several connected applications; in case of a longer downtime, these applications may thus no longer operate on current data. According to the expert, this information can be used to elaborate adequate mitigation plans.

Regarding the business criticality concern, we found a moderate Spearman correlation with outdegree (rho=0,379, $p<0.05$) and betweenness (rho=0,313, $p<0.05$) measures, which were selected.

Following the algorithm, the next step is to look for the critical and very critical applications (y=150) and the 150 highest outliers (x=150) for each selected measures. Both measures overlap in 118 of a total of 165 highest outliers analyzed. Finally, we calculate the recall and precision for the outliers. Among the highest 150 outliers of each selected measure, we can find most (80.66%) of the “critical” and “very critical” applications.

Attribute of interest: Availability. Values: high or very high Network measures: Outdegree, eigenvector and betweenness. 10% outliers selected		
	High/very high	**Normal**
Structural important	5	4
Not Structural important	0	42
Recall =	1 or 100%	
Precision =	0.55 or 55%	

Table 8.3: Recall and precision values for availability

On the other hand, the method suggests a set of potential candidates for further analysis (32 false positives). It is worth mentioning the importance of looking through the 32 false positives to verify whether their business criticality can now be reevaluated by the experts, considering their structural importance. For example, one of the false positives is an application that maintains operationally significant data such as bills, receipts, and other records. Malfunctioning of this application and the corresponding impact on the connected applications may result in major operational issues in, for example, logistics processes.

We can also perform the cognitive-structural diagnosis analysis with a percentage of the outliers´ set. Rather than using x=150 as before, we now investigate the outputs of the two network measures selecting the 20-percent highest outliers to identify among them critical and very critical components. Table 8.4 shows the results. The sample of 20 percent of the outliers from each measure resulted in 60 different components, with an overlap of 29 components that belong to both outliers´ set.

In this case, we have a good precision rate when we consider 20 percent of the highest outliers as a sample. Note that we allowed the experts to decide the percentage of selected outliers (10%, 15%, or 20%), as there is no predefined percentage. In this case, 20 percent of the outliers resulted in 60 different components, which means we are considering only part of the 150 “high” and “very high” critical applications (components of interest set). For this piece of the application architecture, the selected network measures produce a precision rate of 93.33% (i.e., business criticality of components is present among the top 20% outliers of betweenness and outdegree), while the low recall value is due to the reduced sample size that was selected to be analyzed (60 of 202 components).

Attribute of interest: business criticality. Values: critical or very critical. Network measures: outdegree and betweenness. 20% outliers		
	Correct	**Not correct**
Selected	56	4
Not selected	94	48
Recall =	56/(56+94) = 37.33%	
Precision =	=56/(56+4)= 93.33%	

Table 8.4: Recall and precision values for business criticality, using outdegree and betweenness

For the “core application” attribute, we identify most of the core components among the outliers set. However, many (58) false positives are returned (see Table 8.5). The experts should look at the selected components individually, especially the false positives, and also investigate the criteria used to classify the components as core ones. For example, among the false positives is a quality management application that maintains specifications against which to evaluate product quality. Quality management will not work properly if these specifications are not provided to other related applications; while this may not interrupt business operations, it may put reputation at risk, since products of insufficient quality may get into distribution.

Attribute of interest: core application. Values: yes and no. Network measures: indegree, outdegree, eigenvector and betweenness. 63 outliers		
	Correct	**Not correct**
Selected	44	58
Not selected	19	107
Recall =	=44/(44+19)= 69.84%	
Precision =	=44/(58+44)=43.13%	

Table 8.5: Recall and precision values for “core application” using indegree, outdegree, eigenvector and betweenness

We also applied a variation of the cognitive-structural named diagnosis analysis method, implemented in SANTANA et al. (2016b), to analyze the following models created with the dataset03: BpxBp, BoxBo and BuxBu.

First, we aimed to answer the following operational question: can the business process outliers (central points of the network BPxBP) also be identified as elements of the critical path defined by experts? Our EA analysis concern, then, was to identify key structural components of an EA layer (in this instance, the business process layer). We used as input for this analysis the BPxBP network (primary data) extracted from Autocompany’s documents (dataset03). Our hypothesis H1 is that network measures will be able to identify the main components belonging to the critical path already defined by Autocompany’s experts. Additional components will also be identified and may have their importance validated.

The experts highlighted seven BP components in the critical path (the data had to be anonymized). Following the algorithm of the diagnosis analysis method defined in Section 3, we selected the TOP 15 outliers generated for each of six network measures. We then took the most recurrent components among all measures, using the voting strategy, on which everytime a component was voted as an outlier by a metric, a vote was computed for that component. At the end, the analyst might want to analyze the x% most voted components or a specific amount y of most voted components. This resulted in the selection of 21 distinct components (BPs), among which it was possible to identify successfully, from a universe of 102 BPs, the seven BPs that constitute the critical path defined in Autocompany’s documents. This selection also included components considered for further analysis by Autocompany’s experts, who classified all of them as important BPs. As one expert remarked about these complementary BPs, for example, *"BP64 confirms the overall tech concept. BP68 formalizes this concept and turns it into a planned bill of material, and BP73 makes a document that contains these three aspects. Thus, this cluster makes total sense."* In short, the experts found the results consistent with their expectations and very helpful. Thus, we found support for H1.

For the BoxBo network, the operational question here is: can network measures identify the most important BOs (according to the experts' opinions)? Thus, the analysis concern was about key components of business object components, entities produced by business processes that might be handled by other business process components. The BOxBO network is a primary network that aims to represent how a BO is related to other BOs .

The experts classified 12 BOs as critical. Applying the diagnosis analysis method, we first selected the TOP 15 (X=15) components voted by each network measure from a universe of 70. From among all voted components, we chose the 19 with the most votes (Y=19), seeking to verify whether the 12 critical BOs were among the BOs in the ranking (H2 hypothesis). As a result, nine (9 of 12) critical BOs were identified among the 19 ranked components.

The experts indicated that identifying this set was a good result. As one said, "This is a very nice result. We recreated a discussion that we had when we designed the critical path for the BOs. At the time, and until now, we were unsure what the critical path indeed included." We still had ten additional BOs identified as critical by the network measures (10/19) that had not been mentioned a priori by the experts (see yellow squares in Figure 8.7). One expert stated that "these ten components might take part in the main information flow in case of a broader selection." Therefore, we decided to build the Ishikawa diagram in Figure 8.7 to check this information visually:

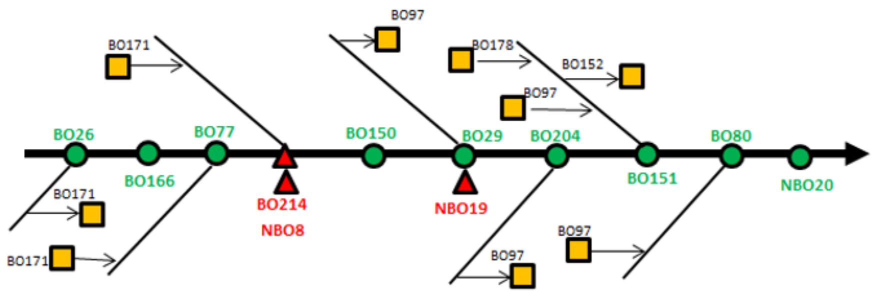

Figure 8.7: Ishikawa diagram for true positive, false positive, and false negative BOs

The red BOs (triangles) in Figure 8.7 were not identified by our method. We can divide the yellow BOs (squares) into two subsets. The first is components with high values for global centrality measures. This is the case for BO152 (high betweenness), BO171, BO97 (high eigenvector), BO178 (high betweenness), and BO97 (high eigenvector). These BOs appear in the surrounding areas of the significant BOs (green circles). The second subset comprises components with high values for local centralities, such as BO57 (high in-degree centrality value) and BO91, BO99, and NB05 (total degree centrality). This second group is not connected with the important BOs identified by the experts but is identified as important due to local importance of the components (high number of in-connections or out-connections to other local BOs) and may be worth of the experts' attention. We claim that these results support H2.

8.4.2 Attribute check analysis

Our final analysis is the attribute check for the attribute type of development from dataset01. This attribute can assume three values: "self-developed," "standard," and "others." For this kind of analysis, we do not need a correlation among the attribute and network measures, since we are performing an exploratory analysis among the outliers that looks for some specific values of interest. We choose the eigenvector and betweenness centralities for the reasons explained in Table 8.1. Considering the 10 percent eigenvector outliers that had the type of development identified as a value (18 applications), we find that among the 18 most important components in terms of eigenvector centrality, almost all are "self-developed" software components and only one is "standard." Regarding the outdegree values, we have only three standard software components among the 19 highest ones.

Altogether, the fact that the most structurally relevant components (pointed out by eigenvector centrality) and the greatest providers of interfaces (applications with high outdegree) belong to the "self-developed" category might indicate a less external dependency on software vendors. On the other hand, the more self-developed components are among the most well-connected ones (as it is present in this case), the more likely may be a variety of self-developed interfaces, which in turn may reduce the architectural agility and impede any integration with partners outside the organization.

For the dataset03, experts wanted to investigate to which extent the manipulation of the main BOs (business objects) was supported by IT components. Then, the BoxBo network was extracted from documents on which experts described the most relevant BOs for the product development process. In that document, 12 BOs are highlighted as critical ones by the experts. The attribute check analysis is applied to investigate if the outliers indicated by the network measures are among these 12 critical ones and, also, if they are supported by IT tools of the company. To perform the method, first, we run the six network centrality measures to detect central BOs and select the TOP 15 components outputted by each of those measures. We use a list of the 19 most recurrent nodes among the ones outputted before.

Among the 19 most recurrent nodes, we identified 9 (9 of 12) critical BOs . Thus, we looked at a list containing all BOs subdivided in smaller (sub) BOs . This list also has the status of the IT support for each sub-BO. From a total of 54 sub-BOs related to the nine critical BOs identified, only 18 are supported by IT tools. In this case, considerable part (61%) of the sub-Bos is not supported by IT tools. In terms of management, this indicator might serve to foster a discussion about the improvement of management and control functions along the product development phases. In Table 8.6, we present the percentage of TOP ranked BOs and their sub-BOs supported by IT tools.

Sub-BO's manipulation is supported by IT tools	Amount	%
Yes	18	33,33%
No	33	61,11%
Not informed	3	5,55%
Total	54	100%

Table 8.6: Percentage of TOP ranked BOs and their subMS Completion Criteria supported by PLM

Experts might take our results to analyze and implement the IT support for the product development process, in order to enhance its management and control functions.

8.5 Chapter summary

In line with the objectives we formulated earlier, the two methods presented in this chapter to combine network measures and expert knowledge as support mechanisms for EA analysis, were applied in three empirical datasets.

We found that, for the analyzed context, business criticality of components is correlated with the network measures we used, which allowed us to validate the experts' initial, inherently subjective perceptions for most components (88.66%), on one hand, and on the other hand suggest a set of potential candidates for further analysis. We found similar results for the cognitive-structural diagnosis analysis of availability, in which all five components of interest were identified by the method and, given their structural importance, four additional components were suggested for further analysis. Regarding the cognitive-structural diagnosis analysis for core application, 63 of 228 applications, the method identified 69.84 percent correctly. However, since the network measures are individually correlated with the attribute of interest, each of them also brings applications with similar structural values (false positives) to further analysis. The precision rate (43.13%) calls for an individual reconsideration of the selected false positives, which were suggested by, at least, one of the four network measures. We need to highlight the importance of false positives since they represent components that have network values similar to the true positives but were not classified as components of interest by the experts.

We also tested the attribute check approach with type of development values from dataset01 and IT support from dataset03. With this kind of descriptive analysis, it was possible to identify structural positions in the main architecture occupied by self-developed applications and critical BOs that were not supported by IT tools. Simple checks like this can warn experts about undesired configurations and help them monitor the EA evolution. It should be noted that both the cognitive-structural diagnosis analysis and the attribute check analysis method accept as inputs virtually any network measure and EA concern.

There are limitations of this study. As for the methods' validity, the definition of the attribute of interest is the most important threat to the construct validity. To facilitate the comparison of results from different organizations, in future studies, we may create a minimum criteria set to define an application as “TOP”, for example. A second threat may be the correctness of the modeled network, which depends on the quality of EA documentation and on the process used to build the network. We recommend researchers always double-check their modeled networks. All in all, we do not propose to generalize the results for all organizations, since the definitions of some concepts of interest (e.g., business criticality and core application) we employed in this study may suffer from an interpretive bias and may vary across organizations. Thus, further validation by analyzing application architectures from additional organizations (with different sizes and structures) is desirable for future research.

Despite the originality of the approach, our analysis is concentrated at the application domain - a strategy followed in other works (e.g., Santana A.; Fischbach K.; Moura H. (2016a); LANKHORST (2013); SIMON; FISCHBACH (2013) Pestina et al., 2009). As other layers, such as business architecture, are also subject to inherent relations and dependencies (cf.Aptrop, 2015), displaying a business capability dependency map), we see the use of our approach not as generally limited to the application architecture but also valid elsewhere in the EA to compare subjective perception with structural information in a given layer (e.g., business capability degree with strategic value assigned). Interlayer dependencies, though, have not yet become a subject of our analysis. In addition, because the attribute check analysis method is an intuitive approach, further examples are needed to provide more evidence of its utility. As for the cognitive-structural diagnosis analysis, it would benefit from future research that includes a case-by-case analysis of the false positives outliers pointed out by each network measure. This analysis consumes considerable time and effort of representatives from the end-user organization. Although we did not have the chance to perform it for all false positives detected, we do recognize its importance.

Finally, we conclude by stressing again that our cognitive-structural diagnosis approach should be seen as a heuristic to support the expert in identifying key applications. In this sense, a supportive approach should be adopted regarding the benefits of the two proposed methods. They should be seen as a guide or supporting instrument in a specific context more than as substitutes for expert assessments. Nevertheless, we believe that the more complex the application landscape, the more the pure “EA analysis by hand” approach can benefit from our methods’ support, as they can both substantiate and complement the former.

More details about the empirical studies which used the cognitive-structural diagnosis analysis and attribute check analysis can be found in SANTANA et al. (2016b,c).

9

EVALUATING THE ARTIFACTS

Along the last chapters, we have presented a set of six proposed artifacts. In this chapter, we take a holistic perspective to understand how those artifacts integrate a coherent whole. The goal of this chapter is thus to present the evaluations of the six proposed artifacts: the EANA-MM, the EANA library, the EANA process, the EA data derivation strategy, and cognitive structural analysis and attribute check analysis methods. First, we present the evaluation criteria in Section 9.1. The EANA-MM and library are evaluated in section 9.2, the EANA process in section 9.3, the EANA data derivation strategy in section 9.4, and finally, the cognitive-structural diagnosis analysis and attribute check analysis evaluations are presented in section 9.5.

9.1 ARTIFACTS AND EVALUATION CRITERIA

First, we describe the proposed artifacts' relationships in Figure 9.1. The dashed lines represent a relation of instantiation of one artifact by another. For instance, the relation between the EANA-MM and EANA library is a directed one and represents the instantiation (use) of the EANA-MM (artifact source of the relation) by the EANA library (artifact target of the relation) as extensively discussed in chapter 5. Another example is the EANA process which may instantiate the two previous mentioned artifacts. As we discuss later, this is due to the utilization of the EANA-MM and EANA library as a knowledge base for the EA process design. The EANA process, in turn, it is instantiated during the design of three other proposed artifacts (EA data derivation strategy, cognitive structural analysis and attribute check analysis methods). Thus, we may say that the EANA-MM is used by all other artifacts directly or indirectly and that the EANA process is directly used by three other artifacts in their design phase. Finally, those three artifacts are not instantiated by other artifacts but, instead, we instantiate themselves in this thesis. Altogether, those six artifacts constitute our contributions for the EANA knowledge base.

Besides explaining the relationships among the proposed artifacts, Figure 9.1 depicts the artifact´s evaluations represented by the numbered black circles. Each relation between two artifacts developed by us (dashed lines) accounts for an instantiation (evaluation at instantiation level, PRAT; COMYN-WATTIAU; AKOKA (2014)) of the source artifact performed by the targeted artifact and receives a label such in 2, 4, 5, 6, 7, and 9. For those cases, if an artifact

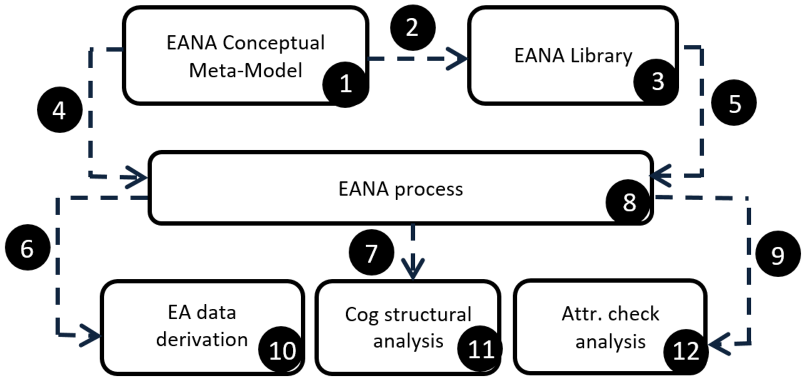

Figure 9.1: Artifacts´ evaluations performed n this thesis

B instantiates an artifact A, we consider in this thesis that A is implicitly evaluated by B. For instance, in the evaluation two, the EANA library uses the EANA-MM in its conception (it instantiates the EANA-MM, according to PRAT; COMYN-WATTIAU; AKOKA (2014)). Nevertheless, we also perform evaluations with EA experts (evaluation with secondary participants, PRAT; COMYN-WATTIAU; AKOKA (2014) or evaluation with external position, CLEVEN; GUBLER; HÜNER (2009)) for the artifacts with the inner circles labeled as in 1 ,3 ,8 ,10 and 11. Altogether, twelve evaluations are performed for the six proposed artifacts. In Table 9.1, those internal and external evaluations and their respective evaluation criteria are described.

Evaluation	Description	Artifact	Evaluation method (Prat et al, 2014)	Evaluation criteria	Section
1	In this evaluation, we survey the experts from the field regarding their opinion regarding the evaluation criteria.	EANA-MM	Survey	Efficacy/Usefulness, Understandability,Generality, Utility, Level of details	8.2
2	In this case, we discuss how the EANA-MM was adopted in the EANA library building.	EANA-MM	Logical reasoning and instantiation	Efficacy and applicability	8.2
3	Similarly to the evaluation 1, we surveyed the researchers about the EANA library.	EANA library	Survey	Efficacy/Usefulness, Utility, Applicability	8.2

Continued on next page

Table 9.1: Methods, evaluation criteria for the proposed artifacts

Table 9.1 – continued from previous page

Evaluation	Description	Artifact	Evaluation method (Prat et al, 2014)	Evaluation criteria	Section
4 and 5	We describe how the EANA-MM and library take part as knowledge base during the EANA process design.	EANA-MM and EANA library	Logical reasoning and instantiation	Efficacy and Applicability	8.3
6,7,8 and 9	For these evaluations, we demonstrate theoretically how the EANA process can be instantiated in three EANA scenarios (evaluation 8). For a fourth scenario, we explain how the EANA process was used during the design of three other artifacts (the EA data derivation strategy, the cognitive-structural method, the attribute check analysis method) (evaluations 6, 7 and 9).	EANA process	Logical reasoning and instantiation	Efficacy and Applicability	8.3
10	We instantiate the EA data derivation strategy with three datasets and evaluate it with EA experts.	EA network data derivation strategy	Logical reasoning and instantiation	Efficacy, applicability and utility.	8.4

Continued on next page

Table 9.1: Methods, evaluation criteria for the proposed artifacts

Table 9.1 – continued from previous page

Evaluation	Description	Artifact	Evaluation method (Prat et al, 2014)	Evaluation criteria	Section
11	We design and instantiate the cognitive-structural diagnosis analysis method with using three empirical datasets and evaluating the expert opinion about the results.	Cognitive-structural diagnosis analysis	Logical reasoning and instantiation	Efficacy, applicability and utility.	8.5
12	Another artifact which is a instantiated with empirical data and EA expert opinion.	Attribute check analysis	Logical reasoning and instantiation	Efficacy, applicability and utility.	8.5

Table 9.1: Methods, evaluation criteria for the proposed artifacts

The first evaluated artifacts are the EANA-MM and EANA library as described in the next section.

9.2 EANA META-MODEL AND EANA LIBRARY EVALUATIONS

During the EANA library design, we instantiated our EANA-MM to classify several NAIs found in the literature (74 in total). Therefore, the inductive approach adopted in the EANA-MM building process guarantees that the EANA-MM artifact has several real instantiations. Therefore, we can support its efficacy (Evaluation 2, Table 9.1). In addition, while designing the EANA process, we also used the conceptual elements of the EANA-MM to support the EA analysis in three different scenarios (Evaluation 4, Table 9.1). Nevertheless, we wanted to complement the EANA-MM evaluation process with an external evaluation, asking experts from this particular field regarding their perceptions about our EANA meta-model and library (Evaluation 1, Table 9.1). In this direction, we invited by email 15 experts who were chosen based on the purposive sampling method among authors found in our dataset of final papers produced in Santana A.; Fischbach K.; Moura H. (2016a). From that sample, we got 3 of 15 invitations accepted (20%) after two rounds of email invitations sent with an interval of 15 days between each. Despite the small sample size, this is a very highly qualified set of respondents, who co-authored the main works in the EANA field.

We thus classify the evaluation approach adopted for the EANA-MM according to the mechanism discussed in chapter 3, as described in Table 9.2:

Dimension	Value				
Approach	Qualitative		Quantitative		
Artifact focus	Technical	Organizational		Strategic	
Artifact type	Construct	Model	Method	Instantiation	Theory
Object	Artifact		Artifact construction		
Time	Ex ante		Ex post		
Level of evaluation	Abstract artifact(Survey)		Instantiation (by the EANA library and EANA process)		
Relativeness' of evaluation	Absolute	Relative to comparable artifacts		Relative to absence of artifacts	
Secondary participant	Student	Practitioner		Researcher	

Table 9.2: Classification for the evaluation method of the EANA-MM

According to Table 9.2, we evaluate the efficacy, understandability, generality, utility, level of details of the EANA-MM in an abstract way, with a qualitative approach, involving practitioners and researchers which answered the survey(Evaluation one, Table 9.1). We also evaluate the efficacy and applicability of the EANA-MM through its instantiation while we classified several NAIs which compose the EANA library (Evaluation 2, Table 9.1).

In the case of the EANA library, we theoretically demonstrate its efficacy, exemplifying how it can be used as an essential component in the EANA process (Evaluation five, Table 9.1). We also survey EA experts about the EANA library itself (Evaluation 3, Table 9.1). Thus, the classification for its evaluation method described in Table 9.3.:

Dimension	Value				
Approach	Qualitative		Quantitative		
Artifact focus	Technical	Organizational		Strategic	
Artifact type	Construct	Model	Method	Instantiation	Theory
Object	Artifact		Artifact construction		
Time	Ex ante		Ex post		
Level of evaluation	Abstract artifact(Survey)		Instantiation (by the EANA library and EANA process)		
Relativeness' of evaluation	Absolute	Relative to comparable artifacts		Relative to absence of artifacts	
Secondary participant	Student	Practitioner		Researcher	

Table 9.3: Classification for the evaluation method of the EANA library

According to the Table 9.3, the evaluation approach is qualitative and the artifact is evaluated in a direct abstract way (Evaluation three, Table 9.1) by researchers and practitioners; and also indirectly through its instantiation by the EANA process (Evaluation five, Table 9.1).

9.2.1 Evaluation results

As stated before, for both artifacts, besides the logical reasoning, we employ a survey (CLEVEN; GUBLER; HÜNER, 2009) as an evaluation method. By conducting a survey, information is collected through interviewing representatives of a certain target group. Questions can relate to subjective and objective issues, whereas the answers are always subjective and only limitedly verifiable (CLEVEN; GUBLER; HÜNER, 2009). The overall intention of this survey was to gain further insights into the efficacy, generality, utility, understandability, clarity, consistency, and the level of detail of the introduced meta-model and library, in the view of experts. We also asked about personal data to characterize the respondents.

We start describing the age and experience related information of the respondents, which were 33, 49 and 29 years old. In Figure 9.2, we present their industry and academic experiences:

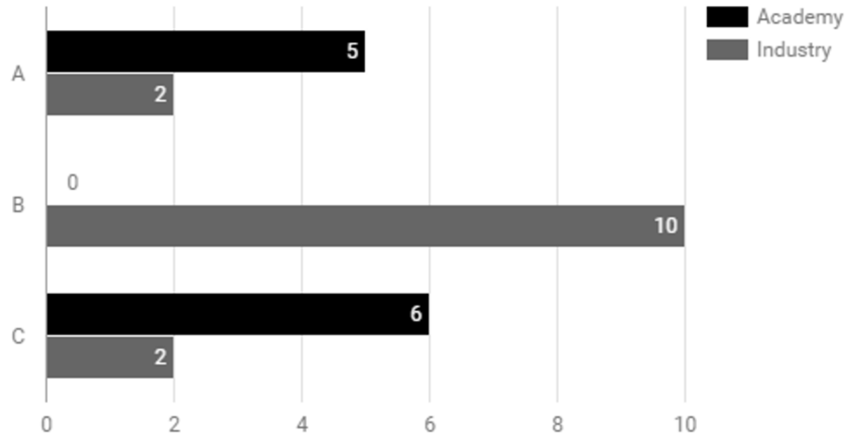

Figure 9.2: Experience in academy and industry of the respondents

According to the sample, we have respondents with at least five years of academic experience in the field combined with some industry experience, while one respondent has ten years of exclusive industry experience, altogether composing a minimal and interesting opinion set in terms of representativeness.

For the EANA-MM artifact, under the goal dimension (see Section 3.5), we take the respondents' perception about four criteria: efficacy, understandability, usefulness and generality (See Table 3.6, Section 3.5.1, for more details about the criteria). Complementary, in Table 9.4 we show the answers from the respondents. More information about the survey instrument and its questions can be found in Appendix B.

Selected Criteria	Related affirmation/question	Evaluation
Efficacy	This meta-model allows classifying existent network analysis initiatives. The meta-model also helps in identifying which information one needs to collect/organize to develop new analysis initiatives.	The respondents agree (3/4) or strongly (1/4) agree that the EANA-MM is effective in providing the EANA research classification conceptual elements.
Understandability	One can clearly understand the concepts which integrate the EANA meta-model and the relations among them.	Two of the respondents observed that if the final user is a practitioner not familiarized with the EANA concepts, they can not ensure that the model understanding is fully achieved, since "in the current form the method and analysis are quite research/academic." For the research audience, although, no problem was pointed by the respondents.
Generality	With this meta-model, it is possible to classify a broad variety of network analysis initiatives.	All the respondents agreed with the generality of the EANA-MM to classify EANA research.
Utility	One could use this meta-model to frame any method developed for EA network analysis research or to document new ones.	According to all respondents, the model is useful from the academic point of view. The usefulness depends on the user, however. Some level of adaptation in required before the model can be adopted by practitioners.

Continued on next page

Table 9.4: Survey result for the EANA-MM

Table 9.4 – continued from previous page

Selected Criteria	Related affirmation/question	Evaluation
Level of details	The level of details of the meta-model is suitable to its goal, that is, to classify or to help to design network analysis initiatives for the EA context.	According to the respondents, "The level of detail is quite appropriate". Another additional feedback was offered to pose the model as "useful to guide initial attention. However, the analysis layer could be more detailed".

Table 9.4: Survey result for the EANA-MM

In conclusion, the efficacy, generality and level of details of the EANA-MM were acknowledged by the respondents. This is an important external evaluation of the research work at hand. From the practitioner point of view, the EANA-MM utility and understandability have to be improved. Despite the concepts being suitable for the research community, respondents also affirmed that if the model is intended to be used in companies (utility criteria), then "the people there must be fully well versed in the EANA terminology, or yet, the analysis might be used as the back end of a specific tool". In this sense, the need for the EA analysis team to pursue a minimal EANA expertise and/or a specific tool which could embed the EANA-MM is highlighted (understandability criteria).

In line with PRAT; COMYN-WATTIAU; AKOKA (2014) and in favor of the efficacy of our EANA library, we reason that, by its very construction, it inherits the efficacy argued in the primary studies from the sample of 29 papers collected in Santana A.; Fischbach K.; Moura H. (2016a), which is the base for the EANA library. Notwithstanding, we asked the respondents about the utility and efficacy criteria as well. The evaluation is described in Table 9.5.

Selected Criteria	Related affirmation/question	Evaluation
Efficacy	Organizing the existent network analysis initiatives for EA as a GQM library helps the expert to focus on a specific concern and capture the relevant data for the analysis .	All respondents support the usefulness of the EANA library´ support with the GQM decision structure, reinforcing that it is an “interesting approach”.
Utility	Once each analysis initiative is implemented, the whole set of NAIs could form a library which can be used as a toolbox for experts to analyze several EA structural concerns.	The respondents (2/3) agree with the utility of the EANA library. The third is not sure about it, since some goals are fuzzy, according to him.
Applicability	Once each analysis initiative is implemented, the whole set of NAIs could form a library which can be used as a toolbox for experts to analyze several EA structural concerns.	One respondent agrees with the statement, while the other two affirmed that the implemented EANA library maybe would generally be used in organization practice. In this case, they conditioned the use of EANA library to particular types of organizations such as “highly networked and complex companies,” that are much suited to such analysis [EANA] than a small company.”

Table 9.5: Survey result for the EANA library

In conclusion, the efficacy and utility of the EANA library were attested with the respondents perceptions. As for the EANA-MM, there is a recommendation to adapt the terminology of concepts and methods identified in this artifact to the organization reality.

Overall, we consider that both artifacts were positively evaluated by the respondents and also find their efficacy demonstrated when applied by other artifacts in this thesis, as we discuss in the following.

9.3 The EANA process view evaluation

In chapter 6, we presented a process view for EANA as a DSR artifact. Its development is grounded in the EANA-MM concepts and EANA library both presented in chapter 5. As discussed in chapter 6, the EANA process is composed of two other sub-processes: the concern-oriented and data-oriented analysis. Both sub-processes differ from each other mainly in terms of the order on which a set of ten activities is executed. In chapter 6, we detailed those activities' inputs, outputs, actors, stakeholders and even indicated some tools that might support their execution.

In this chapter, we are interested in verifying if the EANA process is feasible, i.e. if it can be instantiated in practice. Thus, we claim the efficacy of the EANA process as a DSR artifact, arguing about its reasonability and contextualizing the steps we took in order to design three of our artifacts: the EA data derivation strategy, the cognitive-structural method, the attribute check analysis method. This is how we demonstrate the efficacy of the EANA process.

The first subprocess is concern-oriented and has a main flow which uses the EANA library as a reference to perform the analysis. In this thesis, we do not implement such approach (scenario one). In its alternative flow (scenario two), there is a design of an NAI by the analysis team. Again, this scenario was not present in our thesis. The two EANA process scenarios are depicted in Figure 9.3 despite already being discussed in Section 6.4:

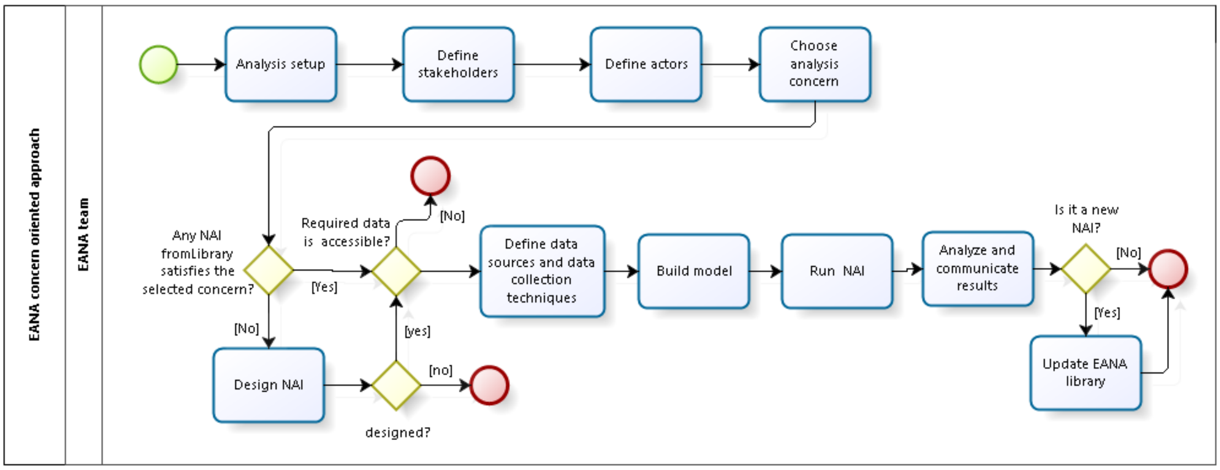

Figure 9.3: Activities of the EANA with concern oriented approach

In the data-oriented EANA subprocess, there is a different starting point. First, we look at the data to decide to apply a suitable NAI taken from the EANA library (third scenario of the EANA process). This third scenario is similar to the first one, on which the analyst picks an NAI from the library and follows the modeling specifications of the selected NAI. In the fourth scenario, the design of a new NAI is performed based on the available data. The third and fourth scenarios for EANA process are depicted in Figure 9.4, also recovered from Section 6.4:

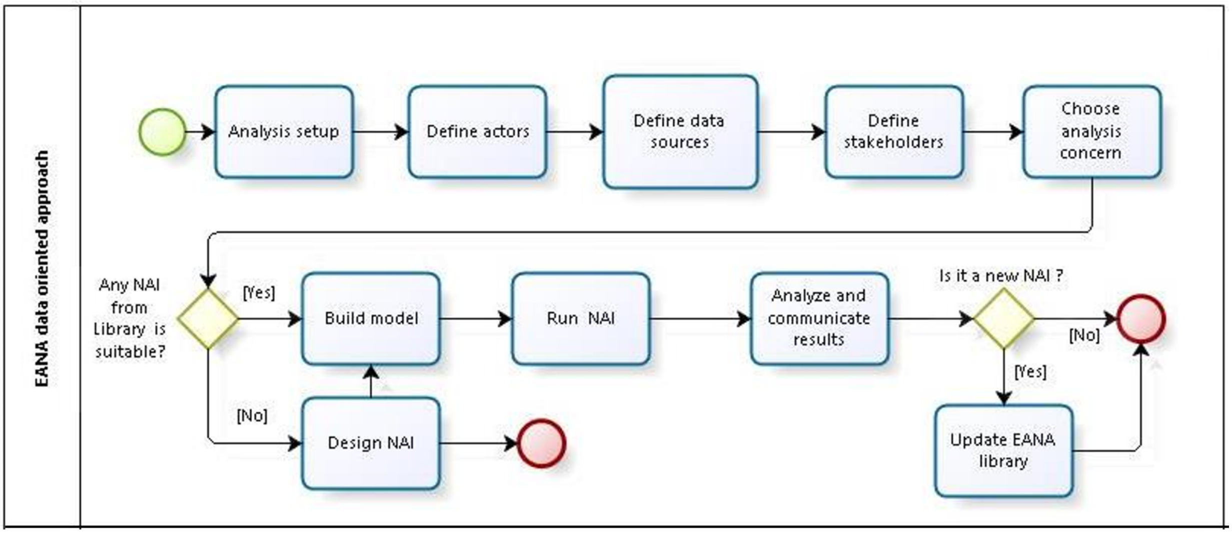

Figure 9.4: Activities of the EANA with data oriented approach

Since we did not perform any of the first three scenarios (there is no instantiation of them in this thesis), therefore, we reason about their efficacy (Evaluation 8, Table 9.1). In this sense, the evaluation methods for the three cases consider each abstract artifact directly, performing an ex ante evaluation (PRAT; COMYN-WATTIAU; AKOKA, 2014). The evaluation methods applied to the three referred scenarios are categorized in Table 9.6.

Dimension	Value				
Approach	Qualitative		Quantitative		
Artifact focus	Technical	Organizational		Strategic	
Artifact type	Construct	Model	Method	Instantiation	Theory
Object	Artifact		Artifact construction		
Time	Ex ante		Ex post		
Level of evaluation	Abstract artifact		Instantiation		
Relativeness' of evaluation	Absolute	Relative to comparable artifacts		Relative to absence of artifacts	
Secondary participant	Student	Practitioner		Researcher	

Table 9.6: Classification for the evaluation method for the scenarios one, two and three

The evaluation of the three mentioned scenarios is qualitative, it considers the EANA process as a method and uses logical reasoning (PRAT; COMYN-WATTIAU; AKOKA, 2014) to directly evaluate the abstract artifact (the EANA process itself), before its instantiation (ex ante evaluation).

On the other hand, we instantiate the data-oriented subprocess three times (scenario four of the EANA process, Figure 9.4) to create three other artifacts described in Sections 8.4 and 8.5. Therefore, the evaluation method used for the three instantiations is categorized in Table 9.7.

According to the Table 9.7, the data-oriented EANA process (its fourth scenario) is also qualitatively evaluated (Evaluations 6, 7 and 9, Table 9.1) though its three instantiations resulted in three different artifacts. Those instantiations involved practitioners in their design process which indirectly corroborate to evaluate the EANA process´s efficacy.

<table>
<tr><th>Dimension</th><th colspan="5">Value</th></tr>
<tr><td>Approach</td><td colspan="2">Qualitative</td><td colspan="3">Quantitative</td></tr>
<tr><td>Artifact focus</td><td>Technical</td><td colspan="2">Organizational</td><td colspan="2">Strategic</td></tr>
<tr><td>Artifact type</td><td>Construct</td><td>Model</td><td>Method</td><td>Instantiation</td><td>Theory</td></tr>
<tr><td>Object</td><td colspan="2">Artifact</td><td colspan="3">Artifact construction</td></tr>
<tr><td>Time</td><td colspan="2">Ex ante</td><td colspan="3">Ex post</td></tr>
<tr><td>Level of evaluation</td><td colspan="2">Abstract artifact</td><td colspan="3">Instantiation</td></tr>
<tr><td>Relativeness' of evaluation</td><td>Absolute</td><td colspan="2">Relative to comparable artifacts</td><td colspan="2">Relative to absence of artifacts</td></tr>
<tr><td>Secondary participant</td><td>Student</td><td colspan="2">Practitioner</td><td colspan="2">Researcher</td></tr>
</table>

Table 9.7: Classification for the evaluation method for the fourth scenario

Putting it in other words, the EANA process is partly evaluated at an abstract level with logical reasoning (Evaluation 8, Table 9.1) and partly at the instantiation level through three instantiations(Evaluations 6, 7 and 9, Table 9.1). In the next Section, we describe the results of both evaluation methods described in Tables 9.6 and 9.7.

9.3.1 Evaluation results

Regarding scenarios one and three previously introduced, since they converge to the same end (they follow a specification of an NAI provided in the EANA library), they are strongly based on third-party NAIs which already had their efficacy demonstrated, theoretically or empirically, in their respective primary studies from which we built the EANA library. However, the activities executed in those primary studies were seldom described. To fill this gap, we believe that the sequences proposed in the EANA process and described in Section 6.4, in combination with the EANA-MM, provide clear guidelines for implementing any of the NAIs already available at our library. For instance, we select one NAI from the EANA library, the "Coupling" metric presented in the [1S_ID_0079], and present the information extracted from that work (which is organized and classified according to our EANA-MM) as described in the following:

- Analysis strategy

The NAI should be applied at the component level, using application components from the application architecture. No application´s attributes are required for it;

- Network metric

The "coupling" concept is operationalized by the closeness centrality. We calculate closeness centrality for a component by summing the value of the reciprocal path values between it and each other component in the network. The path value is the sum of the log (size) of the components encountered along the path.

- Analysis concern targeted

Its analysis concern is “EA cost of change”. The operational question to answer that concern is “How to measure the flexibility of the application layer based on the cost of change of its components?”

- Data strategy

The NAI requires primary data to build an unimodal network. The authors collected the data through documents and interviews.

- Modeling decisions

In order to build the EA model as input for the NAI, the relation meaning adopted is “application depends on application”. The relation is directed and has the transitive property considered in the analysis.

- NAI’s outputs

The NAI produces a list of applications with high coupling values. Grounded on the previous collection of information, the analyst would have to gather, model, and analyze it in a reasonable sequence. To contribute to that, we took inspiration from FUERSTENAU; ROTHE (2014) and also took our experience in executing EA analysis during our research, to propose the concern-oriented EANA process as a guide for practitioners and researchers to use the EANA library.

In the second scenario mentioned earlier, an NAI is designed from scratch with the freedom to collect the necessary data (alternative flow described in Figure 9.4). In this case, our assumption is that the set of activities proposed in our EANA process in combination with the definition of the necessary elements of the EANA-MM are suitable to support the design of a new NAI. In other words, in this scenario, the EA expert knows the required parameters to design the NAI (provided by the EANA-MM) and the sequence of activities to perform the EANA with it (provided by the EANA process). For instance, one should start by composing a team in the analysis setup step, defining data sources and data collection techniques, building the EA model, designing the NAI according to the EANA-MM and so on. Despite the innovative character of this scenario, we think that a minimal baseline to guide the EANA is provided by the EANA process and the EANA-MM.

The fourth EANA process scenario is instantiated three times in this thesis, resulting in the creation of the EA data derivation strategy, cognitive structural analysis and attribute check methods. Their respective evaluations are discussed briefly in the following.

• EA data derivation strategy design as an EANA process´s instantiation

This artifact was already instantiated in chapter 7. In this section, we use process lens to look at that instantiation and highlight the executed EANA process’s activities as a way of evaluating its efficacy.

Following the process flow described in Figure 9.4, in the analysis setup activity, we defined the analysis team, which was composed by one EA expert and one EANA expert for the cases of Dataset01, Dataset02 and Dataset03 described in Section 3.4. At the end of the planning phase, the available data sources (Dataset01, Dataset02 and Dataset03), used documents, actors, EA scope, and supporting tools were all defined during an initial meeting at each company.

After checking the available data, during a second meeting, the EANA expert explained the data-oriented approach to the team and the possibility of applying EA data derivation methods (see Section 7.4) used initially by BUUREN et al. (2004), and widely adopted in the DSM community. This discussion was detailed in Sections 7.1, 7.3 and 7.4. For the case of Dasaset03, two business experts were also included in the analysis team as potential stakeholders. Following the EANA process, we then considered as analysis concerns: 1- how BOs were related considering the technical support of a specific application for their manipulation; 2- technical and vendor ecosystems analysis; 3- the BUs interaction along the business process execution. We used a set of tools (ADOIT, ORA Lite, R plus igraph package and Excel) to extract and manipulate the related datasets to build and convert the EA models to adjacency matrices. The "Concern definition and Data modeling" phase ended after the execution of the previous activities. Those activities were mainly described in Section 7.3.

Finally, as outputs of the "NAI execution" phase, graphs, heat maps and ranking lists were produced with the support of the mentioned tools. The outputs shown in Section 7.5, were reunited in a report and presented during the final meeting with the analysis team and stakeholders in the "Results analysis and communication" phase. Figure 9.5 depicts how the EANA activities are transversely described in details along the chapter 7:

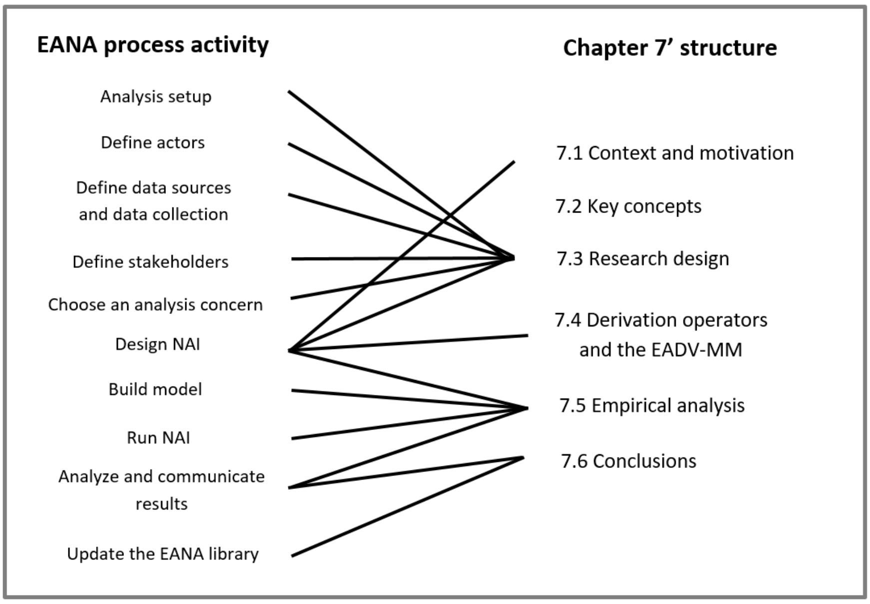

Figure 9.5: Distribution of the EANA activities of the EANA process from chapter 7

- The cognitive-structural diagnosis analysis and attribute check analysis design as EANA process's instantiations

The design and implementation activities of these two artifacts were performed in parallel. They are very similar and that is the reason we concentrate their evaluation in the same section. We instantiated the data-oriented EANA process with three datasets provided by three companies.

Following the data-oriented EANA process (Figure 9.4), the team composition (one EA expert and one EANA expert) and available data sources (three datasets) were defined in a first meeting. We then delimitated the EA scope using application layer data (applications) from Dataset01 and Dataset02, together with their respective components´ attributes (business criticality, availability, "TOP"). We also used information (business objects) and business (business process) layers in the case of Dataset03. For the three datasets, the stakeholder (final user) of the designed NAIs was the EA expert already a member of the analysis team. For the Dataset03, specifically, two additional (business) stakeholders joint the team due to the type of components under analysis (business process and business units).

Considering the data available, the team had the idea of combining the tacit EA knowledge provided by the EA expert with the structural information provided by network analysis applied to the datasets. Particularly, in the case of dataset01 and dataset02, experts wanted to combine their knowledge about important application components (according to their perceptions) in terms of business criticality, availability and strategic value, with key components in terms of structure, pointed out by network analysis metrics. Such a combination method was not available in any of the NAIs from the EANA library and the decision thus was to design a suitable NAI. As discussed in Section 6.3 ("Activities of the EANA process"), there is no template for such a task. However, we take some categories of EANA techniques, and more specifically, types of analysis at component level presented in Section 6.3.7 ("Design NAI") as an initial reference. We discussed the design of the two artifacts in chapter 8. In practice, EA analysis depends critically on human cognitive abilities (e.g., expertise) (HEVNER et al., 2004) to produce effective results (SIMON; FISCHBACH, 2013). We believe that, particularly in organizations with dozens or even hundreds of business processes supported by a similar number of applications, one might want to use additional knowledge sources to add confidence to the analysis. As our design artifact, we develop a method to combine expert knowledge (subjective by nature) about EA components (e.g. critical Business Units (BUs), Business Processs (BPs), and BOs) with network measure´s outputs (structural criteria).

Back to the data-oriented EANA process, after this design step, we built the EA models as adjacency matrices using the support tools (ADOIT, ORA Lite, R plus igraph package and Excel). We also use the same tools to apply network metrics behind the cognitive-structural analysis and attribute check analysis methods to Dataset01 and Dataset02. A variation of the cognitive-structural analysis was applied to the Dataset03. Finally, in the "Analyze and communicate results" phase, for each instantiation (each dataset), a meeting was set with its respective analysis team to discuss the results, considering the graphs, heat maps, and ranking lists produced. Overall, the design of the three artifacts discussed in this Section was contextualized regarding the EANA process´ activities to evaluate the later artifact. The adherence of the three design processes to the prescribed data-oriented EANA activities was detailed in this Section as a means to demonstrate the efficacy of the fourth scenario of the EANA process.

It has to be clear that the artifacts resulted from the instantiation of the data-oriented EANA process are themselves original contributions of this thesis and have specific evaluations which are described in Section 8.5. The detailed results of the design and implementation of the cognitive-structural analysis and attribute check analysis methods are presented in SANTANA et al. (2016b,c)

In the next Section, we evaluate the EA network data derivation strategy itself.

9.4 EA network data derivation strategy evaluation

As presented in chapter 7, with the EADV-MM, combined with a set of operators to derive artificial EA viewpoints, our aim is to provide a way to produce new analysis angles and new questions for the stakeholders.

For the EA network data derivation strategy as a composed artifact (considering the EADV-MM and the derivation operators), we applied a qualitative ex-post evaluation (made after the instantiation), which involved the instantiation of part of the EADV-MM and derivation operators with real data of three datasets. We asked practitioners from the organizations which provided the data (secondary participant evaluation) about the utility of the results produced with the artifact. We also make a distinction of the proposed artifact from the ones already provided in the literature (relative to comparable artifact). The evaluation categorization is described in Table 9.8:

<table>
<tr><th>Dimension</th><th colspan="5">Value</th></tr>
<tr><td>Approach</td><td colspan="2">Qualitative</td><td colspan="3">Quantitative</td></tr>
<tr><td>Artifact focus</td><td>Technical</td><td colspan="2">Organizational</td><td colspan="2">Strategic</td></tr>
<tr><td>Artifact type</td><td>Construct</td><td>Model</td><td>Method</td><td>Instantiation</td><td>Theory</td></tr>
<tr><td>Object</td><td colspan="2">Artifact</td><td colspan="3">Artifact construction</td></tr>
<tr><td>Time</td><td colspan="2">Ex ante</td><td colspan="3">Ex post</td></tr>
<tr><td>Level of evaluation</td><td colspan="2">Abstract artifact</td><td colspan="3">Instantiation</td></tr>
<tr><td>Relativeness' of evaluation</td><td>Absolute</td><td colspan="2">Relative to comparable artifacts</td><td colspan="2">Relative to absence of artifacts</td></tr>
<tr><td>Secondary participant</td><td>Student</td><td colspan="2">Practitioner</td><td colspan="2">Researcher</td></tr>
</table>

Table 9.8: Categorization for the EA network data strategy's evaluation method

As for evaluation method, we employed the analysis and logical reasoning with the artifact´s instantiation in real examples (PRAT; COMYN-WATTIAU; AKOKA, 2014). The analysis criteria were efficacy and utility.

9.4.1 Evaluation results

In chapter 7, we specifically demonstrated the use of the derivation operators with three empirical datasets (efficacy). With this instantiation performed in chapter 7, it was possible to highlight some individual components and emerged clusters which were evaluated by EA experts of each dataset. We claimed that the derivation rules we presented could be used for the creation of “non-standard/implicit” stakeholder-oriented visualizations

Another situation on which the data derivation process might be useful is in the absence of primary data about some component relations. For instance, we used BuxBp and BpxBp networks to derive a BuxBu network, which in turn, allowed us to confirm experts´ perceptions and highlight key structural BUs during the business processes execution. A third example was the derived TecxTec and related networks which brought to the analysis the relations among technologies and vendors, and may help understand their influence and evolution in the enterprise with time (e.g. with a longitudinal analysis). Thus, the utility of the artifact remained proven. For the sake of the readability and organization of this document, we decided to encapsulate the full instantiation of the EA network data strategy in chapter 7. For more details about the artifact outputs and results, the reader should refer to Section 7.5.

9.5 Cognitive-structural diagnosis analysis and attribute check analysis evaluations

Previously we discussed that, as a mean of evaluating the EANA process efficacy, we designed two methods which combined network measures and expert knowledge to analyze EA at the component level (SANTANA et al., 2016b,c). It is important to note that those two methods, the cognitive-structural diagnosis analysis and the attribute check analysis, are themselves two new artifacts. The cognitive-structural diagnosis analysis method minimizes analysis subjectivity while validating important components previously identified by the experts. The method also suggests important structural components be further analyzed which were neglected by experts at first hand. The attribute check analysis offers further contributions by helping in the investigation of particular attributes of applications in important architectural positions. In this Section, we discuss the evaluation of these two artifacts performed with EA practitioners after their instantiations with empirical datasets.

For both methods (or artifacts), we applied a qualitative analysis to verify if the methods´ outputs could support the EA expert in validating his perception about the key structural components of the application architecture. Thus, we instantiated the method using three empirical datasets provided by companies. An EANA expert and one EA expert took part in the analysis providing the input for the methods and validating their outputs as well. We illustrate in Table 9.9 the categorization of the evaluation method applied to these two artifacts:

Our artifacts’ evaluation method comprehends an analysis with logical reasoning, after the instantiation of the two artifacts (PRAT; COMYN-WATTIAU; AKOKA, 2014). As for evaluation criteria, we took the efficacy of the artifact in producing the expected results and its utility for the experts.

Dimension	Value				
Approach	Qualitative		Quantitative		
Artifact focus	Technical	Organizational		Strategic	
Artifact type	Construct	Model	Method	Instantiation	Theory
Object	Artifact		Artifact construction		
Time	Ex ante		Ex post		
Level of evaluation	Abstract artifact		Instantiation		
Relativeness' of evaluation	Absolute	Relative to comparable artifacts		Relative to absence of artifacts	
Secondary participant	Student	Practitioner		Researcher	

Table 9.9: Categorization for the EA network data strategy's evaluation method

9.5.1 *Evaluation results*

We were able to perform the cognitive-structural diagnosis analysis and attribute check analysis (artifact's efficacy was attested) and thus demonstrated how they can allow experts to refine their knowledge of the application architecture (artifact's utility). We performed five instantiations for the cognitive-structural diagnosis analysis artifact (regarding five different EA models), and two instantiations for the attribute check analysis, each of them analyzing different EA analysis concerns as described along the chapter 8.

The results of the cognitive-structural diagnosis analysis demonstrate the method's potential as an auxiliary tool for experts that can help them refine their tacit EA knowledge. Moreover, it was possible to investigate how the values of a certain attribute (type of development and IT support) are distributed among the most important structural points of the network with the attribute check analysis.

Therefore, as detailed in chapter 8, with the help of the experts, the efficacy and utility of both artifacts were demonstrated for the analysis cases. Nevertheless, the artifacts need to be applied in different EA contexts as a way of adding robustness to our results. The chapter is closed in section 9.6.

9.6 Chapter summary

In this chapter, we presented all evaluations (twelve in total) and evaluation criteria applied to the six proposed artifacts. The evaluations were categorized according to the schema shown in Section 3.5 ("Evaluation methods").

Besides being instantiated by other proposed artifacts, the EANA-MM and EANA library were evaluated by three EANA experts regarding their efficacy, usefulness, understandability, generality, utility and the level of details. In general, the EANA-MM and EANA library were well accepted by the respondents ("Indeed, the results output are impressive") but their general suggestions are to "keep in mind that they need to made simple for companies to apply them". In our opinion, the evaluation results reinforce the both artifacts as important contributions to the field.

A subset of the EANA process (scenario 4) was instantiated with three empirical datasets as a way to evaluate its efficacy and applicability. We claim the efficacy and applicability of the scenarios one, two and three considering that they have the same set of activities performed in the instantiated fourth scenario (only in a different sequence) and are strongly conformed with the EANA-MM. Nevertheless, the EANA process should benefit from its instantiation in different companies, EAs, NAIs and experts.

The EADV-MM and derivation operators were also demonstrated with empirical data and evaluated with expert opinion, proven to be a useful approach for generating derived EA viewpoints as initially explored by (BUUREN et al., 2004). The efficacy, applicability and utility of the artifact are detailed in chapter 7 through its instantiation with three datasets.

At last, we showed that combining expert and structural knowledge is a useful tool to assist experts in EA analysis. Due to their very design, the cognitive-structural diagnosis analysis and attribute check methods accept as inputs virtually any network measure and EA concern. A supportive approach should be adopted regarding the benefits of the two proposed methods. They should be seen as a guide or supporting instrument in a specific analysis context more than as substitutes for expert assessments.

Testing and revising the proposed artifacts concepts have occurred through expert review, with preliminary versions exposed to practitioners in specific meetings. The artifacts' evaluation had fundamental participation of EA experts directly (evaluating the results) in seminars/meetings or indirectly (taking part in the EANA process as actors) as shown in Table 9.10.

Artifact	Experts' participation
EANA-MM and EANA library	Three researchers from the EANA field were invited and participated in the evaluation survey.
EANA process	At least, one EA expert and one EANA expert were involved during three instantiations of the data-oriented EANA process.
EANA network data derivation strategy	Results from instantiations with dataset01 and dataset02 were evaluated by one EA expert and one EANA expert from each company. For the dataset03, one EA expert, one EANA expert and two business stakeholders took part in the evaluation process.
Cognitive-structural diagnosis analysis and attribute check analysis	Similar to the previous case.

Table 9.10: Participation of the experts during the artifacts´evaluations

At last, we believe that the involvement of experts evaluating the artifacts positively, in complement with other several artifacts ´instantiations, support our overall claim about the reasonability, applicability and utility of the proposed artifacts. Nevertheless, we recognize that all artifacts might benefit from additional instantiations and evaluations in different industries, companies, experts and so on.

So far in this thesis, we applied and evaluated all artifacts individually inc a compartmentalized way. One interesting approach would be examining a broader instantiation and evaluation that could encompass some of he proposed artifacts is combination (e.g EANA-MM, EANA library and EANA process instantiated together in one analysis case). We discuss this possibility in the next chapter.

10

MAKING SENSE OF THE EANA FRAMEWORK

In the previous chapter, we saw six artifacts being proposed and evaluated. After went in details through each one of the artifacts individually, at this point, a holistic view of how they relate to each other, how they communicate and fit together might be still missing. Therefore, in this chapter, the pieces (artifacts) of the framework are putted together in perspective and explained as a whole unit: the EANA framework. In section 10.1, we present this integrated view of the artifacts. In section 10.2, our goal is to suggest a life-cycle for the framework´s application. We discuss possibilities of evaluation in section 10.3. We conclude the chapter with some considerations about the framework in section 10.4.

10.1 AN INTEGRATED VIEW OF THE ARTIFACTS

The oxford dictionary [1] defines framework as "a basic structure underlying a system, concept, or text". In this direction, what we consider as our framework is the union of basics elements to compose the theoretical base (system) which supports the network analysis to the EA analysis context.

Our framework is composed essentially by four main elements: the EANA-MM, EANA library, EANA process and EANA data derivation strategy. These elements together, in our opinion, represent a comprehensive set of concepts, metrics, methods and process that may help experts to perform network analysis in the EA context. In addition, along our research we also produced side contributions to EANA: two analysis methods that were added to the EANA library. However, those two methods themselves do not represent central components in the framework itself.

In Table 10.1, we generally describe the purpose of each of the four artifacts:

[1] https://en.oxforddictionaries.com/definition/framework

Artifact	Use
EANA-MM	As discussed in chapter 5, it **is used to understand the concepts and information requirements involved when performing EANA**. It allows classifying the past and future research according to their constructs. Therefore, it was used to structure the EANA library. And finally, it may serve as an initial reference to think about NAI to be designed during an instantiation of the EANA process, since it contains several decision points need to perform it.
EANA library	Also discussed in Chapter 5, the EANA library consolidates the existent knowledge about the metrics, methods of network analysis applied to the EA context. **It may be used as an initial reference catalog of NAIs for practitioners**.
EANA data derivation strategy	This artifact has an auxiliary role inside the framework. With the data derivation operators, one can produce new analysis viewpoints from primary data already available at the companies. **In other words, experts can apply the derivation operators to expand their possibilities of EA analysis**, given an existent EA dataset.
EANA process	This is one of the core elements of the framework. **The EANA process is a proposal to guide practitioners during the application of the EANA.** Two approaches for applying the EANA process were specified, together with their inputs, outputs, and tools: the concern-oriented and the data-oriented, both presented in chapter 6. During the EANA process execution, practitioners may use one or more of the previous artifacts in association (e.g. applying the concern oriented analysis approach and collect data in conformance with a specific NAI from the EANA library).

Table 10.1: Survey result for the EANA-MM

The relations among the artifacts and the framework composition are depicted in Figure 10.1. The dashed lines delimit the scope of the EANA framework, composed by the same four artifacts described in Table 10.1 and their respective relations. The two artifacts in the periphery, cognitive structural analysis and attribute check analysis, are independent methods developed in chapter 8, through the instantiation of the EANA process, and represent two new NAIs which are given back to the EANA library.

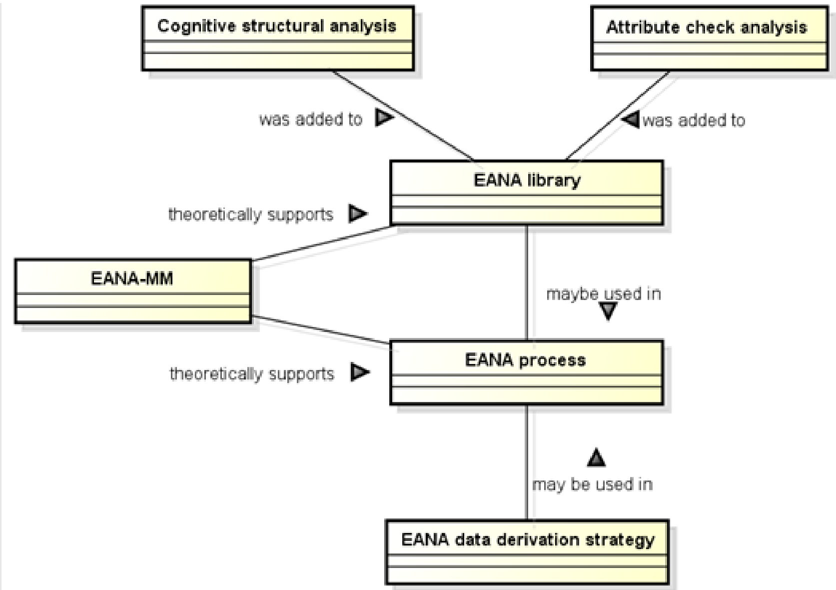

Figure 10.1: The EANA framework and internal relations among the proposed artifacts

Figure 10.1 depicts the holistic view about the designed artifacts and their role in the EANA framework. Since this was not discussed yet, one may wonder how would be the instantiation and evaluation encompassing the four artifacts at the same time. This is a valid question, since we have applied and evaluate the artifacts, individually, as isolated research compartments. In other words, in the previous chapters, only parts of the framework were instantiated and evaluated, one by one (not the framework, itself, as a whole). The designs of the artifacts EANA-MM and the two proposed analysis methods were performed independently and almost simultaneously, for example. We discuss the framework overall instantiation in the next section.

10.2 Using the framework from the beginning

The scenario described in the previous section leaded us to think about the following question: how to start using the EANA framework? In order to answer that question, we came up with the activities depicted in Figure 10.2. The novelty on it compared to what we already saw in the EANA process is the inclusion of two activities: enterprise diagnosis, which prescribes some pre-conditions for using the EANA framework; and the EANA seminar, a training step which might be necessary for specific organizational contexts. The sequence of activities to instantiate the entire EANA framework was based on our experience running experiments during this research and it is explained as follows:

As a first step to apply the proposed framework in its current stage (especially regarding the absence of specific software tools to support it), some pre-conditions must be guaranteed,

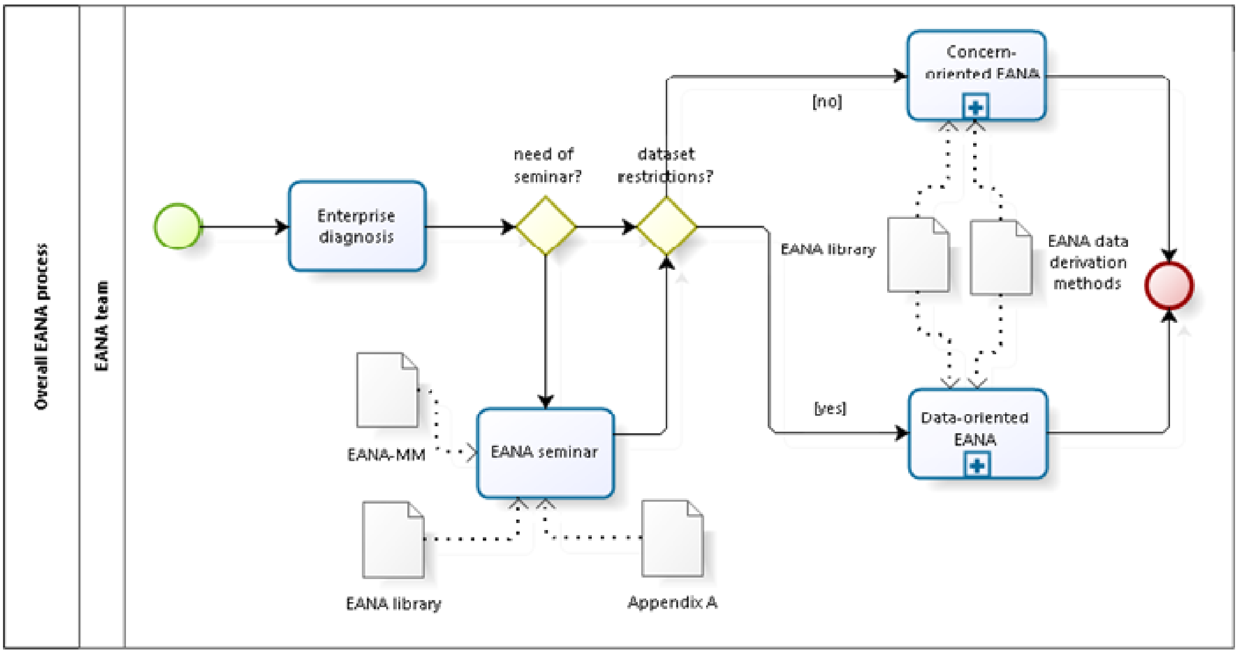

Figure 10.2: The EANA framework´s instantiation guidelines

like the ones as follows:

- Suitable context for the EANA framework instantiation

We suggest that the findings presented in this research would be more valuable to enterprises with complex IT landscapes. Despite not having a precise size to define those scenarios, taking the studied organizations as a reference, organizations with at least a hundred components in any EA layer could benefit from the presented artifacts. As two additional references, the works of FUERSTENAU; ROTHE (2014) andBALDWIN; MACCORMACK; RUSNAK (2013) analyzed 400 applications and 103 applications, respectively. Maturity is EA practices and/or use of EA tools, are desirable but not mandatory for the companies to instantiate the framework. However, naturally, a minimal knowledge regarding EA must be present among the analysis team.

- Network science expertise

Naturally, the presence of network analysis skills constitutes an essential pre-condition to be satisfied. They are useful to understand its applicability in the EA context. In this direction, knowledge about the main centrality metrics, clustering algorithms discussed in the Appendix A should be present in the team.

The EANA-MM presented in chapter 5 is also an important reference, since it describes the fundamental elements needed to perform EANA. Therefore, an overview about the concepts of NAI, analysis strategy, data generation strategy and other related ones can raise awareness about the analysis possibilities and foster an important conceptual alignment that may also help in understanding the results later.

Finally, the EANA library may be presented to the team as a reference catalog since it contains analysis concern candidates and their respective NAIs. In the end, we need to consider that one of the existent NAIs may be useful for the EANA team.

Thus, we need to check if this knowledge is available at the enterprise. In one of our three analysis cases, for example, it was necessary to introduce network analysis concepts for the team, in order to better explore the analysis potential. In this sense, we needed to set a specific team meeting which lasted about one hour. After what we call "EANA seminar", two members (not experts in EA) were able to suggest initial ideas of how to apply EANA in their context. In this seminar, images depicting network metrics results (like the ones presented in section 2.2) are strongly encouraged to be used. In the other two analysis cases, the team members were already familiarized with network analysis techniques (the two already had contact with the subject in the academy). Nevertheless, the seminar was useful to promote a conceptual alignment among members. To sum it up, the presence of an EANA expert in the team is recommendable, perhaps, mandatory, to generate the alignment about the subject, at the current development stage of our framework. In both mentioned cases, I played this role.

- Network science expertise

Once we have a suitable analysis case and the understanding about the EANA concepts and possibilities, the next check is the access or availability of network analysis tools. As mentioned in Section 6.3, due to the absence of a specific software tool for the EANA context, in this thesis, we complemented functionalities and resources from three different general purpose network analysis tools: UCINET, ORA-LITE and R tool combined with igraph package, to import, export and analyze our data. Generally speaking, at least, all of those tools are intuitive and I would not consider them a significant barrier to perform EANA, since they have tutorials and consolidated user communities. All the versions used in this thesis were student related or cost free ones. Costs for software licenses might need to be checked by the enterprise as well. As for the EANA-MM and EANA library, the tools can also be presented and discussed during one or more EANA seminars, if necessary.

Altogether, the three pre-conditions´ checking is performed during the so-called "enterprise diagnosis" activity. If necessary, an EANA seminar may be set to discuss the last three pre-conditions. In case we pass this checking, the next decision is to choose which EANA process approach to implement, considering the existence of any dataset restrictions. This decision point is also depicted in Figure 10.2. From this point on, the discussion is the same as the one done in chapter 6.

10.3 Evaluation of the overall framework

As discussed before and described in details in Table 9.1, we conducted a DSR cycle for each of the six proposed artifacts, while the one cycle DSR encompassing the entire EANA framework is still missing. One question that arises then is how to know if the overall proposal, including the pre-conditions checking, performs well? Nevertheless, we consider the sequence of activities proposed in Figure 10.2 as a reasonable suggestion of how to start using the EANA framework, since those activities were implicit performed during our research. By the time the empirical analyses of our artifacts were done, we did not have the entire EANA framework specified yet. Therefore, the overall EANA framework instantiation, i.e. presenting the EANA-MM, EANA library, network analysis metrics, in a planned and linear workflow starting from the beginning; is a natural candidate for future research. We believe that this limitation, however, does not exclude the possibilities of using individual artifacts from the framework such as EANA-MM and EANA library, individually, as we did along this thesis, when suitable.

10.4 Chapter summary

In this chapter, we presented a holistic perspective about the EANA framework, showing their components, relations and how they fit together to compose a coherent whole unit. Under this view, we briefly discussed the possibilities for overall EANA framework instantiation presenting the artifacts involved on that and related pre-conditions. We see the lack of support of a software tool for data modeling or conversion, and analysis, as the big challenge to instantiate our approach in practice. Notwithstanding, I believe that in the future, with the use of EA models exported from market standard tools and automatically converted to graph-based models, together with the use of a specific purpose and self-contained network analysis tool, the need of the EANA expert during the framework´s instantiation will be reduced while the framework´s applicability will be increased.

11

Final conclusions and future works

This chapter presents our final considerations about this research. A summary of findings and contributions to theory and practice are presented in Section 10.1 while in Section 10.2, we discuss the research limitations, gaps and opportunities.

11.1 Summary of findings

The aim of the present research was to investigate how structural metrics and methods can be applied over EA models to generate information value for EA stakeholders. Our initial motivation had its roots in the work of Simon and Fischbach (SIMON; FISCHBACH, 2013), which advocates that analysis initiatives could represent more reliable and quantitative indicators to describe EA structure, helping experts guide its analysis process and evolution, allowing them to go beyond their perceptions, reducing subjectivity.

Later, in our broader SLR discussed in chapter 4, we saw that, among the five most applied EA analysis paradigms, the structural analysis was present in 18 of 120 papers (or 15%). In line with that, graph-based EA models were present in 21 papers of 120 (or 17.5%). These numbers reinforce the emerging role of structural analysis techniques in the EA analysis context.

We also saw that, for many years, structural analysis has been applied in related fields such as product engineering, complex systems and system engineering. Applying the structural analysis is somehow consolidated at the micro level on those fields resulting in a useful paradigm to understand systems, components and their relations. Our assumption was that, in an analogous way, EA analysis, at a medium or macro level of abstraction, potentially could benefit from that too.

In fact, we have found an initial set of NAIs in our SLR in chapter 5 (74 initiatives, in total). Nevertheless, we observed a lack of theoretical ground for EANA after our initial exploratory research, which was confirmed later by a broad EANA SLR. After these two rounds of analysis, we delineated the research problem around the following issues:

- There was no common language shared by researchers.
- There was no clarity about the extent of application of NAIs.

- Implementation of techniques and methods was not clear in the papers.

Putting it in other words, before this thesis, structural analysis in EA context was not consistently explored. Therefore, we aimed to clarify **the fundamental constructs of EA network analysis and consolidate a mapping of which initiatives did already exist together with their analysis concerns**. In a second moment, our focus was **to describe how this body of knowledge could be used by researchers and practitioners** to enhance their knowledge base and practice, respectively.

Accordingly, grounded in the guidelines for SLR (KITCHENHAM, 2007) and the DSR method (HEVNER et al., 2004), we performed two broad literature reviews and designed six DSR artifacts. During this work, we brought theoretical evidences and expert feedback to design and evaluate the artifacts and improve the EANA knowledge base, identifying its conceptual components, analysis concerns, activities and its application in practice. These enumerated contributions met the research questions proposed for this thesis, which were explained in the following:

- *RQ1 -Which EANA initiatives are already available? and RQ1.2 what are the concerns and approaches analyzed in EANA?*

Our second SLR described in chapter 5 shed light on the existent EANA works. Out of 5739 papers returned by the queries done at five (5) digital libraries (Scopus, IEEE, Engineering village, EBSCO and AISeL were analyzed), twenty-nine (29) primary studies were selected identifying a set of seventy-four (74) NAIs. Up to now, we did not find any other work that lists such an extensive number of NAIs like in the present research.

We analyzed those metrics and methods to describe the common analysis goals aimed in the selected papers, which resulted in a set of seven categories (7) for twenty-four (24) analysis goals already discussed in chapter 5.

We list and organize all the knowledge produced about the metrics and methods extracted from the selected papers according to our EANA-MM, contextualizing them with EA analysis concerns. We used the GQM model to provide a decision structure for helping the architects through the NAI´s choice. At the end, we consolidate all this information in our EANA library (a catalog of NAIs).

Currently, a portal containing all the information obtained during this research was developed (http://www.eanaresearch.org). In the portal, the information about the 74 NAIs is available to be checked by any researcher. In addition, other researchers can contribute inserting new NAIs into it, including their analysis concerns, papers that address the NAIs and related information. Each NAI uses data related to one upto five EA layers and their respective components. The scope of the layers and components adopted in this thesis is described in Table 2.2.

We hope this can facilitate the metrics choice by researchers with none or low experience in metrics adoption, by only showing a subset of all NAIs available according to their analysis needs. Basically, the researcher can choose which category is relevant for his/her evaluation, and the portal informs a list of NAIs.

In addition to the researchers and architects, according to STEEN et al. (2004), EA analysis outputs may be consumed by users varying from employees, software developers, business process designers, product managers, CIOs and even CEOs. With that in mind, user candidates for the EANA library also range accordingly to the diversity of NAIs and analysis concerns found.

The EANA-MM and EANA library were evaluated by EANA experts and also instantiated by other proposed artifacts, as discussed in chapter 9. Therefore, we claim that clear and satisfactory answers to our first research question were given in this thesis.

- *RQ2 - What are the information requirements to perform EANA? and RQ3 - How EA network analysis initiatives can be classified?*

The answers for these questions are based on the EANA-MM and EADV-MM designs. First, with the EANA-MM, we define a meta-model which has the core conceptual elements such as analysis strategy, data strategy, analysis concern, NAI, modeling decisions and so on. Those are conceptual elements which can be used to describe or classify virtually any NAI, as we did with the seventy-seven (74) NAIs already available at the EANA library. Thus, we claim that the EANA-MM can work as the information requirements to classify EA research, also allowing researchers and practitioners to be aware of the elements needed in the design of any NAI.

With the instantiation of the EANA-MM by the EANA library and also by the EANA process, both described in chapter 9, we demonstrate its efficacy and utility. In addition, three EANA researchers also evaluated positively the meta-model regarding the efficacy/usefulness, understandability, generality, utility, and level of details criteria. An issue that is important to point out is that EANA expert researchers, such as Daniel Fürstenau, Mario Sánchez and Prince Mayurank Singh, which worked in well cited papers related to EA analysis and network science, took part in the evaluation survey. They all also gave us positive feedback about the present research.

We saw two types of data used by EANA NAIs: primary and derived data. Analysts can collect primary data in loco or derive it through a set of proposed derivation operators. Considering only the selected studies from our EANA SLR, all of them applied only primary data in their analysis. In this sense, we expanded the possibilities of EA data analysis, with the EADV-MM, which is basically a matrix of derived viewpoints that can be used in combination with data derivation operators, as proposed in chapter 7, to derive the new data for EA analysis. This data generation strategy was evaluated by experts in chapter 9 regarding its efficacy and utility. The derived data was incorporated in the EANA-MM under the conceptual element "data strategy".

Therefore, we claim that RQ2 and RQ3 are consistently answered.

- *RQ4 -How can experts be guided to perform EANA?*

The fourth and last research question lead us to think about some guidelines to perform EANA. Therefore, we answered that question defining a set of main activities and applying them during the design of three DSR artifacts. We named it EANA process and described it in chapter 6. The process itself is divided in two sub-processes, the concern-oriented EANA and the data-oriented EANA. The former was demonstrated theoretically, with logical reasoning while we gave implementation recommendations to justify its efficacy. The latter was instantiated during the design of the EA data derivation strategy, cognitive-structural analysis and attribute check analysis and had criteria such as efficacy and utility evaluated by experts as discussed in chapter 9.

The EANA process represents the first attempt to formalize a set of activities for EANA, resulted from our insights based on the EANA-MM conceptual elements and our experience acquired during the instantiation of three other artifacts. It brings together a set of ten activities in a rational sequence, their expected outputs and suggestions of tools that may support the analysis process. Nevertheless, we recognize that the process certainly will benefit from further implementations in different contexts. With that in mind, the answer for RQ4 was provided. We consider the enterprise architect as a capable direct user of the EANA process.

All in all, the present study makes several noteworthy contributions to theory and practice:

- We design an **EANA meta-model** to classify existent EA network analysis research, the main modelling decisions for EANA, bringing a common set of concepts to be shared by the research community.
- In addition, we list and organize all of the knowledge produced about network analysis in EA so far, and contextualize them with EA analysis concerns as a goal-question-metric based library (**EANA library**). Experts may benefit from such a catalog of NAIs in their daily EA analysis activities, especially if there is a support tool to implement them.
- Third, we present the EADV-MM and a set of derivation operators as a **data derivation strategy**.
- We also propose a generic process for EANA (**EANA process**). The first contribution of its type in EANA research. The process can guide practitioners.

Altogether, those artifacts combined constitute what we call the **EANA framework**. Additionally, the cognitive-structural analysis and attribute check analysis are two new DSR artifacts which can be added to the EANA library as new NAIs and also can be incorporated by analysis tools to be developed.

Finally, the literature review performed in chapter 4 also produced important side contributions: the state-of-art of EA analysis, its main modeling choices, most frequent analysis concerns, most applied techniques and research gaps pointed out by the primary papers.From the SLR developed in chapter 5, another important contribution made is the specific agenda for EANA research. In Figure 11.1, we depict all the outlined research contributions in a timeline fashion (from its top to its bottom).

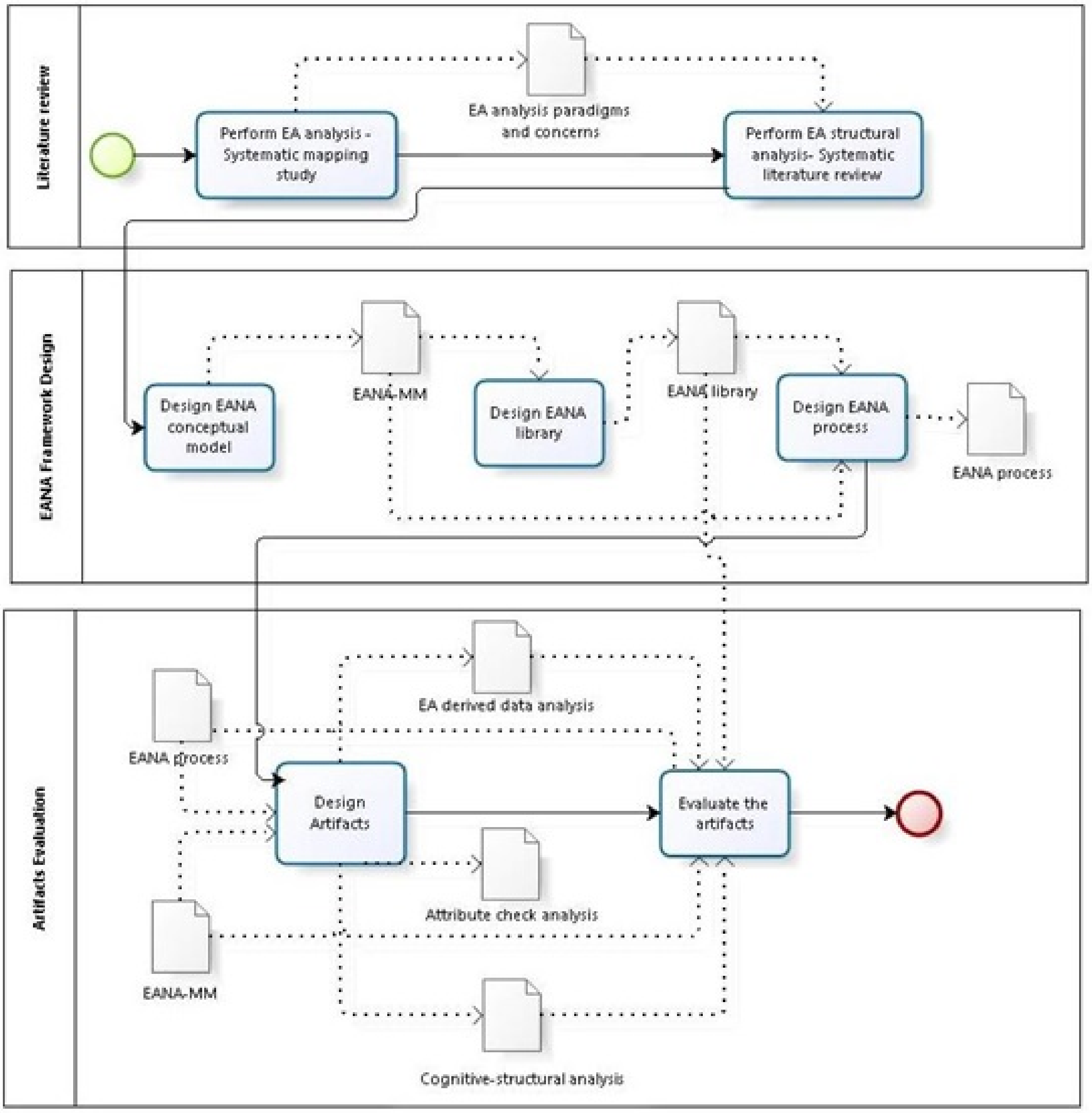

Figure 11.1: Thesis' contributions and the research flow

Therefore, in Table 11.1, we use the taxonomy of (GREGOR; HEVNER, 2013) to classify our proposed EANA framework and methods. Our contributions are placed in the Levels 1 and 2 in the descriptive and prescriptive knowledge categories (GREGOR; HEVNER, 2013).

Altogether, we consider that the general research question, *How to perform analysis of EA components and their relationships supported by network measures?*, was successfully approached with the literature researches, artifacts´ development and evaluations presented throughout this thesis. The presented results represent the state of the art about what is necessary to perform EA analysis and how to perform it, considering the network-based analysis paradigm.

<table>
<tr><th colspan="4">Classification of the contributions of this research</th></tr>
<tr><td>Levels</td><td>Level 1. Situated implementation of artifact(instantiations, software products or implemented processes)

In this thesis:
Two proposed analysis methods; EA network data generation strategies.</td><td>Level 2. Nascent design theory—knowledge as operational principles/architecture (constructs, methods, models, framework, design principles)

In this thesis: EANA-MM, GQM-based EANA library; EANA process.</td><td>Level 3. Well-developed design theory about embedded phenomena.

In this thesis: no contributions at this level.</td></tr>
<tr><td>Type of Knowledge</td><td colspan="2">Descriptive knowledge
(clarification, cataloging, measuring)

In this thesis: GQM-based EANA library</td><td>Prescriptive Knowledge (constructs, Models, methods/techniques, instantiation)
In this thesis: EANA-MM; EANA process; Two proposed analysis methods; EA network data generation strategy.</td></tr>
</table>

Table 11.1: Knowledge contribution classification (adapted from GREGOR; HEVNER (2013)

With the outlined contributions, we pave the way for the research on this specific field, providing a fundamental knowledge base for future research. We believe that this research is the first and necessary step to systematically understand the application of structural analysis in the EA context. Based on the list of network metrics and methods created and the relative success of structural analysis in correlated fields, the EANA thus can be considered a useful and promising instrument to analyze EA, and consecutively, improve the EAM.

We suggest that the findings presented in this research would be more valuable to enterprises with complex IT landscapes. Despite not having a precise size to define those scenarios, taking the studied organizations as a reference, organizations with at least a hundred components in any EA layer could benefit from the presented artifacts.

11.2 Research limitations and future work

As with any empirical field study, the evaluation of the proposed artifacts has limitations. First, despite being evaluated with data from three different organizations and their experts, we argue that the results still need to be applied in different contexts. In this sense, we would like to see studies applying the artifacts to different EA sizes and structures to reinforce their efficacy and utility, promoting thus their enhancement. Nevertheless, we must warn that performing and evaluating the EANA process in its current format is very difficult, since the availability of experts at the companies to discuss the results is scarce. However, a suitable tool might help to achieve results in easier and quicker way.

As discussed in chapter 10, despite the evaluation of each artifact separately, the overall EANA framework instantiation (i.e. presenting the EANA-MM, EANA library, network analysis metrics, in a planned and linear workflow starting from the beginning) is still lacking. Such evaluation may reinforce the usefulness of the artifacts also used in combination among them.

As for future research, the present work offers in the following a wide range of topics to be further investigated:

a) **To apply the EANA process, the EANA library, the data derivation strategy, the cognitive-structure analysis and the attribute check analysis** in different case studies, and possibly, in a integrated and systematic way, according to the prescriptions presented in chapter 10.

b) **To survey industry experts about the EANA library' coverage and utility**, in order to build a prioritized list of NAIs which may be interesting for practitioners to foster the improvement of existing EA analysis tools.

c) **To investigate the benefits of an EA longitudinal analysis with the structural indicators**. Longitudinal studies that trace the structural indicators might contribute with additional insights to EAM, regarding EA complexity and evolution over time, in a dynamic fashion analysis.

d) **To explore transformations from standard modeling languages like UML and Archimate to graph representations**. The design of automated converters may promote the integration of EANA library with existent EA tools.

f) **To include EANA in the portfolio management practices**. Projects are the executors of organizational change and hence in charge of the managed evolution of the application landscape in the context of enterprise architecture (EA) management (BUCKL et al., 2009). Evaluating a portfolio is a complex process wherein organizations need to explore the value of the future performance of the technology as well as the tradeoff between this value and the risk (POPKIN, 2005). According to Fischer et al (FISCHER; MATTHES; WITTENBURG, 2005), IT projects can be prioritized considering their architectural impact. As we did with our cognitive-structural analysis and attribute check analysis methods, we hypothesize that structural information can be combined and used at the portfolio management level as one additional element in the decision process. For instance, projects which impact key structural components of any EA layer, could receive more attention or priority from the stakeholders.

g) **To investigate systematically and transfer other structural analysis techniques from related fields to EA analysis**.

According to Schütz et al. (SCHÜTZ; WIDJAJA; GREGORY, 2013), there are possibly various analysis approaches from other fields which might be transferable to EA. For example, in our EANA SLR described in chapter 5, research from system of systems theory (SoS), system thinking theory, system engineering and others which did not specifically mention the expression "enterprise architecture" did cover one of the EA subsystems – application architecture – and presented some network measures/metrics. In those cases, we investigated if those measures had potential to be listed as a contribution to EA analysis research. In line with that, in Santana et al.(SANTANA et al., 2016b), we were inspired by the DSM research, often applied in product engineering field, to bring some concepts to analyze EA. As a result, the paper was well accepted by one of the pioneers from that field at the DSM conference, who is now interested to take part in related future cooperation.

In Popkin (POPKIN, 2005), there is a discussion about how the disciplines of EA and system engineering deal with different scopes and abstraction levels but still are closely related to each other. For that author, systems engineering has a mature understanding of risk analysis and failure and success rates in delivering products on time and within budget while measuring the performance of enterprise architecture initiatives is still in its early development stages. Up to date, in our view, the latter can benefit from the established analysis know-how of the former.

Similarly, for Nightingale and Rhodes (NIGHTINGALE; RHODES, 2004), the Enterprise Architecting practice has clear extensions from software architecting practice. In fact, we found in chapter 5 that some clustering methods, extensively used in the software architecture research, were used in EANA community to approach EA modularity concerns. In line with that, one anonymous reviewer from the EDOC 2016 conference recommended to look closer at the cognitive software architecture complexity (WANG, 2009) to search for further insights. All that being said, we encourage researchers to look closely at those related fields to verify opportunities of reusing their mature structural analysis approaches in the EANA context.

h) **To investigate overlaps with new IT architectural paradigms**

The analysis of service oriented architectures was not explicitly approached by our identified NAIs. Thus, we are not able to confirm if there are any peculiarities regarding the application of the EANA approaches in this emerging type of IT environments. ALWADAIN et al. (2011) say that there is no uniformity with regard to the SOA elements and their level of details in the EA frameworks. In addition, although ArchiMate has a particularly strong focus on service-orientation and was adopted by the Open Group and the TOGAF body, it is still incomplete with regard to the SOA elements such as service descriptions and the representation of specific types of services such as enterprise services (ALWADAIN et al., 2011). Therefore, how to proceed with EANA in this context is an open research question.

More recently, the advance of Internet of Things and Industry 4.0 (SCHEER, 2013) may represent even more complex scenarios of fully connected enterprises, with their EA components and interactions. Industry 4.0 connects embedded system production technologies and smart production processes to pave the way to a new technological age which will radically transform industry and production value chains and business models (SNIDERMAN; MONIKA; COTTELEER, 2016). The exercise of imagining the possibilities for the EANA research in this context may be quite exciting.

Finally, we already described wider specific research agendas in chapters 4 and 5 for EA analysis and EA network analysis, respectively.

REFERENCES

AALST, W. van der. Data Science in Action. In: **Process Mining**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016. p.3–23.

ADDICKS, J. S. Enterprise architecture dependent application evaluations. **2009 3rd IEEE International Conference on Digital Ecosystems and Technologies, DEST '09**, [S.l.], p.594–599, 2009.

ADDICKS, J. S.; APPELRATH, H.-J. A method for application evaluations in context of enterprise architecture. **Proceedings of the 2010 ACM Symposium on Applied Computing - SAC '10**, [S.l.], p.131, 2010.

AHLEMANN, F. et al. (Ed.). **Strategic Enterprise Architecture Management: challenges, best practices, and future developments**. [S.l.]: Springer Heidelberg Dordrecht London New York, 2012. 300p.

AIER, S.; SCHÖNHERR, M. Model driven service domain analysis. In: INTERNATIONAL CONFERENCE ON SERVICE-ORIENTED COMPUTING. **Anais...** [S.l.: s.n.], 2006. p.190–200.

ALWADAIN, A. et al. Where do we find services in enterprise architectures? A comparative approach. In: AUSTRALASIAN CONFERENCE ON INFORMATION SYSTEMS (ACIS 2011), 22. **Proceedings...** [S.l.: s.n.], 2011.

AMBLER, S. Enterprise Architecture: reality over rhetoric. **Dr. Dobb's Journal**, [S.l.], 2010.

ANABY-TAVOR, A. et al. Insights into enterprise conceptual modeling. **Data & Knowledge Engineering**, [S.l.], v.69, n.12, p.1302–1318, 2010.

ANDERSEN, P.; CARUGATI, A. **ENTERPRISE ARCHITECTURE EVALUATION: a systematic literature review**. 2014.

ANTUNES, G. et al. Using ontologies for enterprise architecture analysis. In: ENTERPRISE DISTRIBUTED OBJECT COMPUTING CONFERENCE WORKSHOPS (EDOCW), 2013 17TH IEEE INTERNATIONAL. **Anais...** [S.l.: s.n.], 2013. p.361–368.

BALDWIN, C. Y.; MACCORMACK, A.; RUSNAK, J. Hidden Structure: using network methods to map system architecture. **Harvard Business School Working Paper**, [S.l.], 2013.

BARTOLOMEI, J. E. **Qualitative Knowledge Construction for Engineering Systems: extending the design structure matrix methodology in scope and procedure**. 2007. 391p. Tese (Doutorado em Ciência da Computação) — MIT.

BARTOLOMEI, J. E. et al. Engineering Systems Multiple-Domain Matrix: an organizing framework for modeling large-scale complex systems. **Systems Engineering**, [S.l.], v.15, n.1, p.41–61, mar 2012.

BEETZ, K.; KOLBE, L. Towards Managing IT Complexity: an it governance framework to measure business-it responsibility sharing and structural it organization. **17th Americas Conference on Information Systems AMCIS 2011**, [S.l.], p.1–12, 2011.

BERNARD, S. A. **An Introduction to Enterprise Architecture**. [S.l.]: AuthorHouse, 2012.

BOELL, S. K.; CECEZ-KECMANOVIC, D. On being ‘systematic’ in literature reviews in IS. **Journal of Information Technology**, [S.l.], v.30, n.2, p.161–173, 2015.

BOER, F. de et al. Change impact analysis of enterprise architectures. In: IRI -2005 IEEE INTERNATIONAL CONFERENCE ON INFORMATION REUSE AND INTEGRATION, CONF, 2005. **Anais. . .** IEEE, 2005. p.177–181.

BONACICH, P. Power and centrality: a family of measures. **American journal of sociology**, [S.l.], p.1170–1182, 1987.

BORGATTI, S. P.; HALGIN, D. S. Analyzing affiliation networks. **The SAGE handbook of social network analysis**, [S.l.], p.417–433, 2011.

BOUCHARAS, V. et al. The Contribution of Enterprise Architecture to the Achievement of Organizational Goals: a review of the evidence. In: TEAR. **Anais. . .** [S.l.: s.n.], 2010. p.1–15.

BOUCHER, X. et al. Process clusters for information system diagnostics: an approach by organisational urbanism. **Production Planning & Control**, [S.l.], v.22, n.1, p.91–106, 2011.

BRADLEY, R. V. et al. The Role Of Enterprise Architecture In The Questforit Value. **MIS Quarterly Executive**, [S.l.], v.10, n.2, 2011.

BRAUN, V.; CLARKE, V. Using thematic analysis in psychology. **Qualitative research in psychology**, [S.l.], v.3, n.2, p.77–101, 2006.

BROWNING, T. R. Applying the design structure matrix to system decomposition and integration problems: a review and new directions. **Engineering Management, IEEE Transactions on**, [S.l.], v.48, n.3, p.292–306, 2001.

BROWNING, T. R. The many views of a process: toward a process architecture framework for product development processes. **Systems Engineering**, [S.l.], v.12, n.1, p.69–90, dec 2009.

BROWNING, T. R. On the alignment of the purposes and views of process models in project management. **JOURNAL OF OPERATIONS MANAGEMENT**, [S.l.], v.28, n.4, p.316–332, jul 2010.

BROWNING, T. R. Design structure matrix extensions and innovations: a survey and new opportunities. **IEEE Transactions on Engineering Management**, [S.l.], v.63, n.1, p.27–52, 2016.

BUCKL, S. et al. Towards Simulation-supported Enterprise Architecture Management. In: MOBIS. **Anais. . .** [S.l.: s.n.], 2008. p.131–145.

BUCKL, S. et al. A Pattern-based Approach to Quantitative Enterprise Architecture Analysis. In: FIFTEENTH AMERICAS CONFERENCE ON INFORMATION SYSTEMS. **Proceedings. . .** [S.l.: s.n.], 2009.

BUSCHLE, M. et al. A tool for enterprise architecture analysis using the PRM formalism. **Lecture Notes in Business Information Processing**, [S.l.], v.72 LNBIP, p.108–121, 2011.

BUSCHLE, M. et al. A Tool for Enterprise Architecture Analysis Using the PRM Formalism. In: INFORMATION SYSTEMS EVOLUTION. **Anais. . .** [S.l.: s.n.], 2011. p.108–121. (Lecture Notes in Business Information Processing, v.72).

BUSS, T. F.; SHILLABEER, A. IT and Enterprise Architecture in US Public Sector Reform: issues and recommendations. **Enterprise Architecture for Connected E-Government: Practices and Innovations, P. Saha (Ed.), Hershey, PA: Information Science Reference**, [S.l.], p.412–440, 2012.

BUUREN, R. V. et al. Composition of Relations in Enterprise. **Graph Transformations**, [S.l.], v.3256, p.39–53, 2004.

CAETANO, A.; SILVA, A. R.; TRIBOLET, J. A role-based enterprise architecture framework. **Proceedings of the 2009 ACM symposium on Applied Computing - SAC '09**, [S.l.], p.253, 2009.

CASEWISE. **Casewise modeler tool**. 2016.

CASTELLANOS, C.; CORREAL, D.; MURCIA, F. An Ontology-Matching Based Proposal to Detect Potential Redundancies on Enterprise Architectures. In: INTERNATIONAL CONFERENCE OF THE CHILEAN COMPUTER SCIENCE SOCIETY, 2011. **Anais...** [S.l.: s.n.], 2011. p.118–126.

CASTELLANOS, C.; CORREAL, D.; MURCIA, F. An ontology-matching based proposal to detect potential redundancies on enterprise architectures. **Proceedings - International Conference of the Chilean Computer Science Society, SCCC**, [S.l.], p.118–126, 2012.

CHEN, K.; CHIANG, M.; POOR, H. From Technological Networks to Social Networks. **IEEE Journal on Selected Areas in Communications**, [S.l.], v.31, n.9, p.548–572, 2013.

CHIRIAC, N. et al. Level of modularity and different levels of system granularity. **Journal of Mechanical Design**, [S.l.], v.133, n.10, p.101007, 2011.

CHIZZOTTI. **A Pesquisa em Ciências humanas e sociais**. São Paulo: Cortez, 1995.

CHRISTENSEN, C.; ALBERT, R. Using graph concepts to understand the organization of complex systems. **International Journal of Bifurcation and Chaos**, [S.l.], v.17, n.07, p.2201–2214, 2007.

CLEVEN, A.; GUBLER, P.; HÜNER, K. M. Design alternatives for the evaluation of design science research artifacts. **Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology**, [S.l.], n.MAY 2009, p.1–8, 2009.

CLOO, J. et al. Positioning Enterprise Architecture. **Enterprise Architecture**, [S.l.], p.25–47, 2009.

COLLIS, J.; HUSSEY, R. **Business Research**: a practical guide for undergraduate and postgraduate students. [S.l.]: Palgrave Macmillan, 2009.

COOK, K. S. et al. The distribution of power in exchange networks: theory and experimental results. **American journal of sociology**, [S.l.], p.275–305, 1983.

CRESWELL, J. **Research design : qualitative, quantitative, and mixed methods approaches**. Los Angeles: Sage, 2009.

CRESWELL, J. W. **Qualitative inquiry and research design**: choosing among five approaches. [S.l.]: Sage, 2013.

CRUZES, D. S.; DYBA, T. Recommended Steps for Thematic Synthesis in Software Engineering. **2011 International Symposium on Empirical Software Engineering and Measurement**, [S.l.], n.7491, p.275–284, 2011.

DAM, H. K.; LÊ, L.-S.; GHOSE, A. Managing changes in the enterprise architecture modelling context. **Enterprise Information Systems**, [S.l.], n.March, p.1–31, 2015.

DAVIS, R. **ARIS Design Platform: advanced process modelling and administration**. [S.l.]: Springer London, 2008.

DAVOUDI, M. R.; ALIEE, F. S.; MOHSENZADEH, M. A New Approach towards Enterprise Architecture Analysis. In: INTERNATIONAL CONFERENCE ON ENTERPRISE INFORMATION SYSTEMS AND WEB TECHNOLOGIES (EISWT-09). **Anais...** [S.l.: s.n.], 2009. p.44–51.

Degenne, A. & Forse, M. Introducing Social Networks. London. , [S.l.], 1999.

DREYFUS, D.; IYER, B. Enterprise Architecture: a social network perspective. **Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)**, [S.l.], v.8, n.C, p.1–10, 2006.

DREYFUS, D.; IYER, B. Managing architectural emergence: a conceptual model and simulation. **Decision Support Systems**, [S.l.], v.46, n.1, p.115–127, 2008.

DREYFUS, D.; WYNER, G. M. Digital Cement: software portfolio architecture, complexity, and flexibility. In: A RENAISSANCE OF INFORMATION TECHNOLOGY FOR SUSTAINABILITY AND GLOBAL COMPETITIVENESS. 17TH AMERICAS CONFERENCE ON INFORMATION SYSTEMS, AMCIS 2011, DETROIT, MICHIGAN, USA, AUGUST 4-8 2011. **Anais...** [S.l.: s.n.], 2011. v.1.

EASTERBROOK, S. et al. Selecting Empirical Methods for Software Engineering Research. **Guide to Advanced Empirical Software Engineering**, [S.l.], p.285–311, 2008.

EMERSON, R. M. Power-dependence relations. **American sociological review**, [S.l.], p.31–41, 1962.

ENGEL, R.; SCHUTT, R. Single-subject design. **The practice of research in social work**, [S.l.], p.206–247, 2009.

EROL, O.; SAUSER, B. J.; MANSOURI, M. A framework for investigation into extended enterprise resilience. **Enterprise Information Systems**, [S.l.], v.4, n.2, p.111–136, 2010.

FASANGHARI, M. et al. A novel credibility-based group decision making method for Enterprise Architecture scenario analysis using Data Envelopment Analysis. **Applied Soft Computing**, [S.l.], v.32, p.347–368, 2015.

FISCHER, C. The Information Systems Design Science Research Body Of Knowledge–A Citation Analysis In Recent Top-Journal Publications. **PACIS 2011 Proceedings**, [S.l.], 2011.

FISCHER, F.; MATTHES, F.; WITTENBURG, A. Improving IT management at the BMW Group by integrating existing IT management processes. In: EDOC ENTERPRISE COMPUTING CONFERENCE, 2005 NINTH IEEE INTERNATIONAL. **Anais...** [S.l.: s.n.], 2005. p.219–225.

FRAMEWORK, F. E. A. Washington, D. **: Chief Information Officers Council We**, [S.l.], 2013.

FREEMAN, L. C. Centrality in social networks conceptual clarification. **Social Networks**, [S.l.], v.1, n.3, p.215–239, jan 1978.

FUERSTENAU, D.; ROTHE, H. Shadow IT Systems: discerning the good and the evil. **ECIS 2014 Proceedings**, [S.l.], p.1–14, 2014.

FÜRSTENAU, D. Exploring Enterprise Transformation from a Path Dependence Perspective : a recycling case and conceptual model. , [S.l.], 2015.

FURTMEIER, F. A.; TOMMELEIN, I. D. Explorative Application of the Multi-Domain Matrix Methodology in Lean Design. **18th Annual Conference of the International Group for Lean Construction**, [S.l.], p.151–160, 2010.

GAALOUL, K.; GUERREIRO, S. A decision-oriented approach supporting enterprise architecture evolution. In: IEEE 24TH INTERNATIONAL CONFERENCE ON ENABLING TECHNOLOGIES: INFRASTRUCTURE FOR COLLABORATIVE ENTERPRISES, 2015. **Anais. . .** [S.l.: s.n.], 2015. p.116–121.

GAMMELGÅRD, M.; EKSTEDT, M.; NÄRMAN, P. Architecture Scenario Analysis – Estimating the Credibility of the Results. **INCOSE International Symposium**, [S.l.], v.17, n.1, p.1399–1413, 2007.

GIL, A. C. **Como elaborar projetos de pesquisa**. 5.ed. Sao Paulo: Atlas, 2010.

GIRVAN, M.; NEWMAN, M. E. Community structure in social and biological networks. **Proceedings of the national academy of sciences**, [S.l.], v.99, n.12, p.7821–7826, 2002.

GMATI, I.; RYCHKOVA, I.; NURCAN, S. On the way from research innovations to practical utility in enterprise architecture: the build-up process. **International Journal of Information System Modeling and Design**, [S.l.], v.1, n.3, p.20–44, 2010.

GREGOR, S.; HEVNER, A. R. POSITIONING AND PRESENTING DESIGN SCIENCE RESEARCH FOR MAXIMUM IMPACT. **MIS Quarterly**, [S.l.], v.37, n.2, p.337–355, 2013.

GRIGORIEV, L.; KUDRYAVTSEV B. D.A GRIGORIEV, L.; KUDRYAVTSEV, D. ORG-Master: combining classifications, matrices and diagrams in the enterprise architecture modeling tool. In: KNOWLEDGE ENGINEERING AND THE SEMANTIC WEB (KESW 2013). **Anais. . .** [S.l.: s.n.], 2013. p.250–257. (Communications in Computer and Information Science, v.394).

GROUP, C. A. W. et al. C4ISR architecture framework version 2.0. **US Department of Defense**, [S.l.], 1997.

GROUP, T. O. **ArchiMate® 3.0 Specification**. [S.l.]: van Haren Publishing, 2016.

GUNGOR, C.; OGUZTUZUN, H. GoalDAG - ArchiMate Integration. In: INFORMATION AND SOFTWARE TECHNOLOGIES, ICIST 2014. **Anais. . .** [S.l.: s.n.], 2014. p.194–210. (Communications in Computer and Information Science, v.465).

HANNEMAN, R. A.; RIDDLE, M. **Introduction to Social Network Methods**. Riverside: University of California, 2005.

HANSCHKE, I. **Strategic IT management - A toolkit for Enterprise Architecture Management**. [S.l.]: Springer Heidelberg Dordrecht London New York, 2010.

HANSCHKE, I. **Strategic IT Management: a toolkit for enterprise architecture management**. [S.l.]: Springer Publishing Company, Incorporated, 2014.

HEVNER, A.; CHATTERJEE, S. **Design Research in Information Systems. Integrated Series in Information Systems, vol. 22**. [S.l.]: New York, NY: Springer, 2010.

HEVNER, A. R. et al. Design science in information systems research. **MIS Quarterly**, [S.l.], v.28, n.1, p.75–105, mar 2004.

HEVNER, A. R.; MARCH, S. T.; RAM, S. DESIGN SCIENCE IN INFORMATION SYSTEMS RESEARCH. **Management Information Systems Quarterly**, [S.l.], v.32, n.4, p.725–730, 2008.

HILLIARD, R. Ieee-std-1471-2000 recommended practice for architectural description of software-intensive systems. **IEEE, http://standards. ieee. org**, [S.l.], v.12, p.16–20, 2000.

HOFER, S. Modeling the Transformation of Application Landscapes. In: IFIP WORKING CONFERENCE ON THE PRACTICE OF ENTERPRISE MODELING. **Anais...** [S.l.: s.n.], 2013. p.101–113.

HOLLAUER, C.; WILBERG, J.; OMER, M. A Matrix-based Framework to Support Dynamic Modeling of Sociotechnical Systems. In: INTERNATIONAL DSM CONFERENCE 2015, 17. **Proceedings...** [S.l.: s.n.], 2015. p.19–28.

HOMMES, Q. D. V. E. Comparison and application of metrics that define the components modularity in complex products. In: ASME 2008 INTERNATIONAL DESIGN ENGINEERING TECHNICAL CONFERENCES AND COMPUTERS AND INFORMATION IN ENGINEERING CONFERENCE. **Anais...** [S.l.: s.n.], 2008. p.287–296.

IYER, B.; DREYFUS, D.; GYLLSTROM, P. A Network-based View of Enterprise Architecture. **... Enterprise Systems Architecture ...**, [S.l.], p.1–25, 2007.

JAMES, G. A. et al. Gartner Enterprise Architecture Framework: evolution 2005. **Gartner, Inc. and/or its Affiliates**, [S.l.], 2005.

JENKINS, S.; KIRK, S. R. Software Architecture Graphs As Complex Networks: a novel partitioning scheme to measure stability and evolution. **Inf. Sci.**, New York, NY, USA, v.177, n.12, p.2587–2601, June 2007.

JOHNSON, P.; NORDSTRÖM, L.; LAGERSTRÖM, R. Formalizing analysis of enterprise architecture. **Enterprise Interoperability, New Challenges and Approaches**, [S.l.], p.1–10, 2007.

JOHNSON, P.; NORDSTROM, L.; LAGERSTROM, R. Formalizing analysis of enterprise architecture. In: ENTERPRISE INTEROPERABILITY: NEW CHALLENGES AND APPROACHES. **Anais...** [S.l.: s.n.], 2007. p.35–44.

KITCHENHAM, B. A. **Guidelines for performing Systematic Literature Reviews in Software Engineering**. Durham: University of Durham, 2007.

KITCHENHAM, B. A. Systematic Review in Software Engineering: where we are and where we should be going. In: ND INTERNATIONAL WORKSHOP ON EVIDENTIAL ASSESSMENT OF SOFTWARE TECHNOLOGIES, 2., New York, NY, USA. **Proceedings...** ACM, 2012. p.1–2. (EAST '12).

KITCHENHAM, B.; CHARTERS, S. Guideline for Performing Systematic Literature Reviews in Software Engineering. Version 2.3. **RN**, [S.l.], 2007.

KOTUSEV, S. The History of Enterprise Architecture: an evidence-based review. **Journal of Enterprise Architecture**, [S.l.], v.12, n.1, p.29–37, 2016.

KOTUSEV, S. et al. Consolidating Enterprise Architecture Management Research. In: HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES. **Anais...** IEEE, 2015. p.4069–4078.

KREIMEYER, M. **A Structural Measurement System for Engineering Design Processes**. 2009. Tese (Doutorado em Ciência da Computação) — Technische Universität München.

KREIMEYER, M.; GÜRTLER, M.; LINDEMANN, U. Structural metrics for decision points within Multiple-Domain Matrices representing design processes. **2008 IEEE International Conference on Industrial Engineering and Engineering Management, IEEM 2008**, [S.l.], p.435–439, 2008.

KRUCHTEN, P. The 4+1 View Model of Architecture. **IEEE Softw.**, Los Alamitos, CA, USA, v.12, n.6, p.42–50, 1995.

KUECHLER, W.; VAISHNAVI, V. A Framework for Theory Development in Design Science Research : multiple perspectives science research : multiple perspectives. **Journal of the Association for Information**, [S.l.], v.13, n.6, p.395–423, 2012.

LAGERSTRÖM, R. et al. Visualizing and Measuring Software Portfolio Architectures: a flexibility analysis. In: INTERNATIONAL DSM CONFERENCE:, 6., Paris. **Anais...** [S.l.: s.n.], 2014. p.65–74.

LAGERSTRÖM, R.; JOHNSON, P.; EKSTEDT, M. Architecture analysis of enterprise systems modifiability: a metamodel for software change cost estimation. **Software quality journal**, [S.l.], v.18, n.4, p.437–468, 2010.

LAGERSTRöM, R. et al. Visualizing and Measuring Enterprise Application Architecture: an exploratory telecom case. In: HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES, 2014. **Anais...** [S.l.: s.n.], 2014. p.3847–3856.

LANGERMEIER, M.; SAAD, C.; BAUER, B. A Unified Framework for Enterprise Architecture Analysis. **2014 IEEE 18th International Enterprise Distributed Object Computing Conference Workshops and Demonstrations**, [S.l.], p.227–236, 2014.

DIETZ, J.; PROPER, E.; JOSE´TRIBOLET (Ed.). **Enterprise Architecture at Work**. 3.ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. (The Enterprise Engineering Series, v.27).

LANKHORST, M. **Enterprise Architecture: key to successful business transformations**. 2016.

LANKHORST, M. M. Enterprise architecture modelling - The issue of integration. **Advanced Engineering Informatics**, [S.l.], v.18, n.4, p.205–216, 2004.

LEE, H. et al. Enterprise architecture content model applied to complexity management while delivering IT services. **Proceedings - 2014 IEEE International Conference on Services Computing, SCC 2014**, [S.l.], p.408–415, 2014.

LEHNERT, M. et al. ProcessPageRank – A Network-based Approach to Process to Prioritization Decisions. In: ECIS 2015 PROCEEDINGS. **Anais...** [S.l.: s.n.], 2015. v.4801, p.1–15.

LEIST, S.; ZELLNER, G. Situational Architecture Engineering (SAE) - Improving Strategic Change Through Architecture Methods. In: INTERNATIONAL CONFERENCE ON INFORMATION SYSTEMS (ICIS), 29. **Anais...** [S.l.: s.n.], 2008. p.1–17.

LEVINA, O.; HILLMANN, R. Network-Based Business Process Analysis. **2012 45th Hawaii International Conference on System Sciences**, [S.l.], p.4356–4365, 2012.

LEVIS, A. H.; WAGENHALS, L. W. C4ISR architectures: i. developing a process for c4isr architecture design. **Systems engineering**, [S.l.], v.3, n.4, p.225–247, 2000.

LI, J. et al. Distribute service components into multi-level architecture from legacy applications. **2007 International Conference on Convergence Information Technology, ICCIT 2007**, [S.l.], p.1134–1140, 2007.

LINDEMANN, U.; MAURER, M.; BRAUN, T. **Structural complexity management**: an approach for the field of product design. [S.l.]: Springer Science & Business Media, 2008.

MANZUR, L. et al. xArchiMate: enterprise architecture simulation, experimentation and analysis. **Simulation**, [S.l.], v.91, n.3, p.276–301, 2015.

MARCH, S. T.; SMITH, G. F. Design and natural science research on information technology. **Decision support systems**, [S.l.], v.15, n.4, p.251–266, 1995.

MATHEWS, Z.; KAESLIN, L.; RYTZ, B. Harnessing Multimodal Architectural Dependencies for IT-Business Alignment and Portfolio Optimization: a statistical approach. In: BENGHOZI, P. et al. (Ed.). **Digital Enterprise Design {&} Management**: proceedings of the second international conference on digital enterprise design and management ded{&}m 2014. Cham: Springer International Publishing, 2014. p.35–46.

MAYRING, P. Qualitative Content Analysis. **Analysis**, [S.l.], n.December, 2004.

MAZHELIS, O. et al. Defining complexity factors for the architecture evaluation framework. **Proceedings of the Annual Hawaii International Conference on System Sciences**, [S.l.], v.9, n.C, p.1–10, 2006.

MOCKER, M. What is complex about 273 applications? untangling application architecture complexity in a case of European investment banking. **Proceedings of the 42nd Annual Hawaii International Conference on System Sciences, HICSS**, [S.l.], p.1–14, 2009.

MORKEVIIUS, A.; GUDAS, S.; ILINGAS, D. Model-driven Quantitative Performance Analysis of Updm-based Enterprise Architecture. In: 29TH . **Anais...** [S.l.: s.n.], 2010.

NARANJO, D.; SANCHEZ, M.; VILLALOBOS, J. Visual Analysis of Enterprise Models. **2012 IEEE 16th International Enterprise Distributed Object Computing Conference Workshops**, [S.l.], p.19–28, 2012.

NARANJO, D.; SÁNCHEZ, M.; VILLALOBOS, J. PRIMROSe: a graph-based approach for enterprise architecture analysis. In: INTERNATIONAL CONFERENCE ON ENTERPRISE INFORMATION SYSTEMS. **Anais. . .** [S.l.: s.n.], 2014. p.434–452.

NÄRMAN, P. et al. Data accuracy assessment using enterprise architecture. **Enterprise Information Systems**, [S.l.], v.5, n.1, p.37–58, 2011.

NEWMAN, M. E. The structure and function of complex networks. **SIAM review**, [S.l.], v.45, n.2, p.167–256, 2003.

NIEMANN, K. D. **From enterprise architecture to IT governance**. [S.l.]: Springer, 2006. v.1.

NIGHTINGALE, D. J.; RHODES, D. H. Enterprise systems architecting: emerging art and science within engineering systems. In: ESD EXTERNAL SYMPOSIUM. **Proceedings. . .** [S.l.: s.n.], 2004.

ONWUEGBUZIE, A.; LEECH, N. &.; COLLINS, K. Qualitative Analysis Techniques for the Review of the Literature. **Qualitative Report**, [S.l.], v.17, n.56, p.1–28, 2012.

ÖSTERLIND, M.; LAGERSTRÖM, R.; ROSELL, P. Assessing modifiability in application services using enterprise architecture models–a case study. In: **Trends in Enterprise Architecture Research and Practice-Driven Research on Enterprise Transformation**. [S.l.]: Springer, 2012. p.162–181.

OSTROWSKI, L. Detailed Design Science Research and Its Impact on the Quality of Design Artefacts. In: HELFERT, M.; DONNELLAN, B. (Ed.). **Practical Aspects of Design Science**: european design science symposium, edss 2011, leixlip, ireland, october 14, 2011, revised selected papers. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. p.60–70.

OSTROWSKI, L.; HELFERT, M. Commonality in various design science methodologies. **2011 Federated Conference on Computer Science and Information Systems (FedCSIS)**, [S.l.], p.317–320, 2011.

PAN, W. Applying complex network theory to software structure analysis. **World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering**, [S.l.], v.5, n.12, p.1634–1640, 2011.

PEFFERS, K. E. N. et al. A Design Science Research Methodology for Information Systems Research. **Journal of Management Information Systems**, [S.l.], v.24, n.3, p.45–77, 2007.

PEFFERS, K. et al. A design science research methodology for information systems research. **Journal of management information systems**, [S.l.], v.24, n.3, p.45–77, 2007.

PETTICREW, M.; ROBERTS, H. **Systematic Reviews in the Social Sciences: a practical guide**. [S.l.: s.n.], 2006. 336p.

PLATANIOTIS, G.; De Kinderen, S.; PROPER, H. a. Relating decisions in enterprise architecture using decision design graphs. **Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOC**, [S.l.], p.139–146, 2013.

PLATANIOTIS, G. et al. A Conceptual Model for Compliance Checking Support of Enterprise Architecture Decisions. In: BUSINESS INFORMATICS (CBI), 2015 IEEE 17TH CONFERENCE ON. **Anais. . .** [S.l.: s.n.], 2015. v.1, p.191–198.

PONTUS, J. et al. **IT Management with Enterprise Architecture**. Stockolm: ePub, 2013.

POPKIN, J. 6.2. 1 “Systems Engineering: driving the evolution to actionable architecture”. In: INCOSE INTERNATIONAL SYMPOSIUM. **Anais. . .** [S.l.: s.n.], 2005. v.15, n.1, p.856–872.

POPOVA, V.; SHARPANSKYKH, A. A formal framework for modeling and analysis of organizations. In: **Situational Method Engineering**: fundamentals and experiences. [S.l.]: Springer, 2007. p.343–358.

POSTINA, M.; SECHYN, I.; STEFFENS, U. Gap analysis of application landscapes. In: ENTERPRISE DISTRIBUTED OBJECT COMPUTING CONFERENCE WORKSHOPS, 2009. **Anais. . .** IEEE, 2009. p.274–281.

POWER, D. J. Data science: supporting decision-making. **Journal of Decision Systems**, [S.l.], v.25, n.4, p.345–356, oct 2016.

PRAT, N.; COMYN-WATTIAU, I.; AKOKA, J. **Artifact Evaluation In Information Systems Design-Science Research – A Holistic View**. 2014.

CORDEIRO, J. et al. (Ed.). **PRIMROSe**: a graph-based approach for enterprise architecture analysis. Cham: Springer International Publishing, 2015. p.434–452.

QUATRANI, T.; PALISTRANT, J. **Visual Modeling with IBM Rational Software Architect and UML**. [S.l.]: IBM Press, 2006. (The developerWorks series).

RAMOS, A. et al. Automated Enterprise-Level Analysis of ArchiMate Models. In: BIDER, I. et al. (Ed.). **Enterprise, Business-Process and Information Systems Modeling SE - 30**. [S.l.]: Springer Berlin Heidelberg, 2014. p.439–453. (Lecture Notes in Business Information Processing, v.175).

RAYMOND, K. Reference model of open distributed processing (RM-ODP): introduction. In: **Open distributed processing**. [S.l.]: Springer, 1995. p.3–14.

RAZAVI, M.; ALIEE, F. S.; TAFRESH, A. E. S. A fuzzy AHP based approach towards enterprise architecture evaluation. In: EUROPEAN CONFERENCE ON INFORMATION MANAGEMENT & EVALUATION. **Proceedings. . .** [S.l.: s.n.], 2009. p.408–421.

REMENYI, D.; WILLIAMS, B. **Doing research in business and management**: an introduction to process and method. [S.l.]: Sage, 1998.

RICO, D. F. A framework for measuring ROI of enterprise architecture. **Journal of Organizational and End User Computing**, [S.l.], v.18, n.2, p.I, 2006.

ROSS, J. W.; WEILL, P.; ROBERTSON, D. **Enterprise architecture as strategy**: creating a foundation for business execution. [S.l.]: Harvard Business Press, 2006.

ROSSI, M.; SEIN, M. K. Design research workshop: a proactive research approach. **Presentation delivered at IRIS**, [S.l.], v.26, p.9–12, 2003.

RUNESON, P.; HÖST, M. Guidelines for Conducting and Reporting Case Study Research in Software Engineering. **Empirical Softw. Engg.**, Hingham, MA, USA, v.14, n.2, p.131–164, 2009.

S., S.; S., R.; V., K. Using intentional and system dynamics modeling to address whys in enterprise architecture. **ICSOFT 2013 - Proceedings of the 8th International Joint Conference on Software Technologies**, [S.l.], n.January, p.24–31, 2013.

SAAT, J.; AIER, S.; GLEICHAUF, B. Assessing the Complexity of Dynamics in Enterprise Architecture Planning–Lessons from Chaos Theory. In: AMERICAS CONFERENCE ON INFORMATION SYSTEMS (AMCIS 2009), SAN FRANCISCO, 15. **Proceedings. . .** [S.l.: s.n.], 2009.

SANTANA, A. . et al. Empirical investigation of enterprise architecture analysis based on network measures and expert knowledge: a case from the automotive industry. In: INTERNATIONAL DSM CONFERENCE 2016, 18., São Paulo. **Anais. . .** [S.l.: s.n.], 2016b.

SANTANA, A. et al. Combining Network Measures and Expert Knowledge to Analyze Enterprise Architecture at the Component Level. In: ENTERPRISE DISTRIBUTED OBJECT COMPUTING CONFERENCE (EDOC), 2016 IEEE 20TH INTERNATIONAL. **Anais. . .** [S.l.: s.n.], 2016c. p.1–10.

Santana A.; Fischbach K.; Moura H. Enterprise Architecture Analysis and Network Thinking : a literature review. In: HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES (HICSS 2016), 49., Kauai. **Anais. . .** [S.l.: s.n.], 2016a.

ŠAŠA, A.; KRISPER, M. Enterprise architecture patterns for business process support analysis. **Journal of Systems and Software**, [S.l.], v.84, n.9, p.1480–1506, 2011.

SCHEER, A. W. Industrie 4.0. **Germany Trade and Invest**, [S.l.], 2013.

SCHMIDT, C. **Measuring and Managing Enterprise Architecture Complexity**. 2013.

SCHMIDT, C.; BUXMANN, P. Outcomes and success factors of enterprise IT architecture management: empirical insight from the international financial services industry. **European Journal of Information Systems**, [S.l.], v.20, n.2, p.168–185, 2011.

SCHNEIDER, A. W. et al. Empirical Results for Application Landscape Complexity. , [S.l.], 2015.

SCHNEIDER, A. W.; ZEC, M.; MATTHES, F. Adopting Notions of Complexity for Enterprise Architecture Management. **Twentieth Americas Conference on Information Systems**, [S.l.], p.1–10, 2014.

SCHOENHERR, T. Logistics and supply chain management applications within a global context: an overview. **Journal of business logistics**, [S.l.], v.30, n.2, p.1–25, 2009.

SCHÜTZ, A.; WIDJAJA, T.; GREGORY, R. Escape from Winchester Mansion–Toward a Set of Design Principles to Master Complexity in IT Architectures. , [S.l.], 2013.

SCOTT, J. **Social Network Analysis**. Newbury Park:CA: Sage, 1992.

SCOTT, J. **Social network analysis**. [S.l.]: Sage, 2012.

SCRIVEN, M. The new science of evaluation. **Scandinavian Journal of Social Welfare**, [S.l.], v.7, n.2, p.79–86, apr 1998.

SEAMAN, C. Qualitative methods in empirical studies of software engineering. **IEEE Transactions on Software Engineering**, [S.l.], v.25, n.4, p.557–572, 1999.

SESSIONS, R. **A Comparison of the Top Four Enterprise-Architecture Methodologies**. 2007.

SILVER, M. S.; MARKUS, M. L.; BEATH, C. M. The information technology interaction model: a foundation for the mba core course. **MIS quarterly**, [S.l.], p.361–390, 1995.

SIMON, D.; FISCHBACH, K. IT Landscape Management Using Network Analysis. In: **Lecture Notes in Business Information Processing**. [S.l.: s.n.], 2013. v.139, p.18–34.

SIMON, D.; FISCHBACH, K.; SCHODER, D. An exploration of enterprise architecture research. **Communications of the Association for Information Systems**, [S.l.], v.32, n.1, p.1–72, 2013.

SIMON, H. A. The sciences of the artificial (3rd ed.). , [S.l.], nov 1996.

SIMONIN, J.; Le Traon, Y.; JÉZÉQUEL, J. M. An enterprise architecture alignment measure for telecom service development. **Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOC**, [S.l.], p.476–483, 2007.

SINGH, H.; SINGH, R. On Formal Models and Deriving Metrics for Service-Oriented Architecture. **Journal of Software**, [S.l.], v.5, n.8, p.866–871, aug 2010.

SNIDERMAN, B.; MONIKA, M.; COTTELEER, M. J. **Industry 4.0 and manufacturing ecosystems**. [S.l.: s.n.], 2016.

SOLINGEN, R. van; BERGHOUT, E. The Goal/Question/Metric Method: a practical guide for quality improvement of software development. **McGraw Hill, ISBN**, [S.l.], v.7, p.7, 1999.

SOSA, M. E.; EPPINGER, S. D.; ROWLES, C. M. A network approach to define modularity of components in complex products. **TRANSACTIONS-AMERICAN SOCIETY OF MECHANICAL ENGINEERS JOURNAL OF MECHANICAL DESIGN**, [S.l.], v.129, n.11, p.1118, 2007.

STECHER, P. Building business and application systems with the Retail Application Architecture. **IBM Systems Journal**, [S.l.], v.32, n.2, p.278–306, 1993.

STEEN, M. W. A. et al. Supporting Viewpoint-Oriented Enterprise Architecture. , [S.l.], n.Edoc, 2004.

STELZER, D. Enterprise architecture principles: literature review and research directions. In: INTERNATIONAL CONFERENCE ON SERVICE-ORIENTED COMPUTING. **Anais...** Springer-Verlag, 2009. p.12–21.

SUNKLE, S. et al. Incorporating Directives into Enterprise TO-BE Architecture. In: EDOC WORKSHOPS. **Anais...** [S.l.: s.n.], 2014. p.57–66.

SUNKLE, S. et al. Incorporating Directives into Enterprise TO-BE Architecture. In: IEEE 18TH INTERNATIONAL ENTERPRISE DISTRIBUTED OBJECT COMPUTING CONFERENCE WORKSHOPS AND DEMONSTRATIONS, 2014. **Anais...** IEEE, 2014. p.57–66.

TAMM, T. et al. How does enterprise architecture add value to organisations? , [S.l.], 2011.

TARENSKEEN, D.; BAKKER, R.; JOOSTEN, S. Applying the V Model and Axiomatic Design in the Domain of {IT} Architecture Practice. **Procedia {CIRP}**, [S.l.], v.34, p.263–268, 2015.

The DoDAF Architecture Framework Version 2.02. [S.l.]: US. Department of Defense, 2010.

The Open Group. **TOGAF Version 9.1**. [S.l.]: Van Haren Publishing, 2011.

The Open Group. **ArchiMate® 2.1 Specification**. [S.l.: s.n.], 2013.

VAISHNAVI, V.; KUECHLER, B. Design Research in Information Systems. **Association for Information Systems**, [S.l.], 2005.

VAKKURI, E. **Developing Enterprise Architecture With The Design Structure Matrix**. 2013. 97p. Tese (Doutorado em Ciência da Computação) — Tampere University ff Technology.

VAN DER RAADT, B.; SLOT, R.; VAN VLIET, H. Experience report: assessing a global financial services company on its enterprise architecture effectiveness using naomi. In: SYSTEM SCIENCES, 2007. HICSS 2007. 40TH ANNUAL HAWAII INTERNATIONAL CONFERENCE ON. **Anais...** [S.l.: s.n.], 2007. p.218b–218b.

VASCONCELOS, A. et al. Open Issues on Information System Architecture Research Domain: the vision. In: ICEIS (3). **Anais...** [S.l.: s.n.], 2004. p.273–282.

VENEBERG, R. et al. Enterprise Architecture Intelligence: combining enterprise architecture and operational data. In: IEEE 18TH INTERNATIONAL ENTERPRISE DISTRIBUTED OBJECT COMPUTING CONFERENCE, 2014. **Anais...** IEEE, 2014. p.22–31.

VRIES, M. de; GERBER, A.; MERWE, A. van der. The Enterprise Engineering Domain. In: ADVANCES IN ENTERPRISE ENGINEERING IX. **Anais...** [S.l.: s.n.], 2015. p.47–63. (Lecture Notes in Business Information Processing, v.211).

WALDMAN, F.; SANGAL, N. Applying DSM to enterprise architectures. In: DSM 2007: PROCEEDINGS OF THE 9TH INTERNATIONAL DSM CONFERENCE, MUNICH, GERMANY, 16.-18.10. 2007. **Anais...** [S.l.: s.n.], 2007.

WAN, H.; CARLSSON, S. Towards an Understanding of Enterprise Architecture Analysis Activities. In: EUROPEAN CONFERENCE ON INFORMATION MANAGEMENT AND EVALUATION. **Anais...** [S.l.: s.n.], 2012. p.334.

WANG, S. et al. Features of enterprise information systems integration: a systemic analysis. In: SMC. **Anais...** [S.l.: s.n.], 2011. p.333–339.

WANG, Y. On the Cognitive Complexity of Software and its Quantification and Formal Measurement. **International Journal of Software Science and Computational Intelligence**, [S.l.], v.1, n.June, p.31–53, 2009.

WASSERMAN, S.; FAUST, K. Social network analysis: methods and applications. **Cambridge University Press**, [S.l.], v.1, p.116, 1994.

WATTS, D. J.; STROGATZ, S. H. Collective dynamics of 'small-world'networks. **nature**, [S.l.], v.393, n.6684, p.440–442, 1998.

WENDT, T.; BRIGL, B.; WINTER, A. Assessing the integration of information system components. In: INTEROPERABILITY OF HETEROGENEOUS INFORMATION SYSTEMS. **Proceedings...** [S.l.: s.n.], 2005. p.55–62.

WIERINGA, R. **Design science methodology**. [S.l.: s.n.], 2010. 493p.

WIJEGUNARATNE, I.; FERNANDEZ, G.; EVANS-GREENWOOD, P. **Enterprise Architecture for Business Success.** [S.l.]: Bentham Science Publishers, 2014. 234p.

WOHLIN, C.; AURUM, A. Towards a decision-making structure for selecting a research design in empirical software engineering. **Empirical Software Engineering**, [S.l.], v.20, n.6, p.1427–1455, 2015.

WOOD, J. et al. A framework for capturing the hidden stakeholder system. **SYSTEMS ENGINEERING**, [S.l.], v.16, n.3, p.251–266, 2013.

YIN, R. K. Case study research design and methods third edition. **Applied social research methods series**, [S.l.], v.5, 2003.

YU, E.; STROHMAIER, M.; DENG, X. Exploring Intentional Modeling and Analysis for Enterprise Architecture. In: IEEE INTERNATIONAL ENTERPRISE DISTRIBUTED OBJECT COMPUTING CONFERENCE WORKSHOPS (EDOCW'06), 2006. **Anais...** [S.l.: s.n.], 2006. p.32–32.

ZACARIAS, M. et al. Adding a Human Perspective to Enterprise Architectures. In: DATABASE AND EXPERT SYSTEMS APPLICATIONS, 2007. DEXA '07. 18TH INTERNATIONAL WORKSHOP ON. **Anais...** [S.l.: s.n.], 2007. p.840–844.

ZACHMAN, J. a. A framework for information systems architecture. **IBM Systems Journal**, [S.l.], v.26, n.3, p.276–292, 1987.

ZACHMAN, J. **The Zachman Framework Evolution**. 2011.

Appendix

A

NETWORK ANALYSIS BASIC CONCEPTS

In this appendix, we will describe some concepts from network analysis field which will be used during the work.

On the context of social network analysis, the vertexes of a graph are called **nodes** and their edges are called **connections or relationships**. A graph is called sociogram. Sociograms were developed to illustrate the structure of informal relationships inside group's studies by Hawthorne(SCOTT,2010), who analyzed behavior aspects on groups like: game´s involvement, job's changes/turnover, mutual help, friendship and antagonisms. In a sociogram, a circle represents relationships among actors or persons. These relationships can be **weighted** or not, can have bi-directional or uni-directional connection and may have a positive meaning (friendship) or a negative one. The Figure A.1 presents a sociogram with 5 (five) nodes and their relations. In this case, the relations are not mutual (directed sociogram).

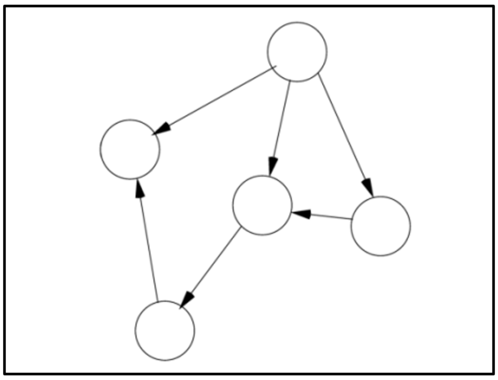

Figure A.1: Elements of a simple sociogram

A sequence of adjacent and not repeat nodes is called **path**. The shortest path between two nodes is called **geodesic path**. The biggest geodesic path of a network or sociogram is called **network diameter**. We can visualize the network diameter between the nodes A and B on the Figure A.2.

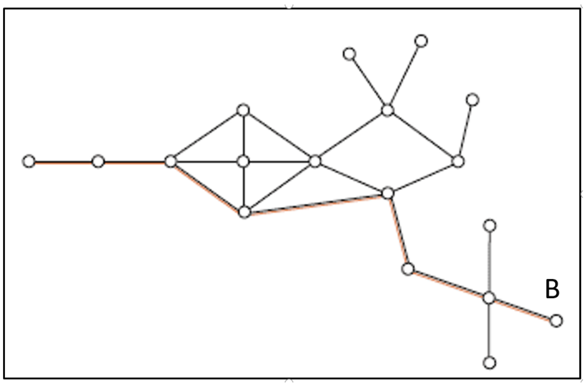

Figure A.2: Network diameter illustration

DENSITY

Scott (2000) defines Density as a global measure of a network structure, that is, it is not applied on a single node. Measuring the density of a network gives us a ready index of the degree of dyadic connection in a population. For binary data, density is simply the ratio of the number of adjacencies that are present divided by the number of possible connections. This would be the Density ratio for a Graph G:

$$Density(G) = \frac{\text{amount of existent edges or connections}}{\text{total of possible connections}}$$

Or formally:

$$Density(G) = \frac{2 \times e}{n \times (n-1)}$$

Where n is the number of nodes and e is the number of edges in the Graph G. Visually, we can observe the difference between densities of these two networks on the Figure A.3 (the right one in more dense).

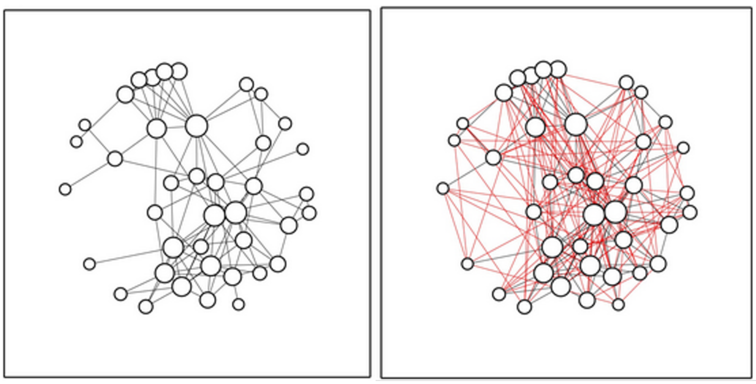

Figure A.3: Difference on network density

High levels of density on the IT landscape can be a warning about the high level of dependence among the applications, or that some design improvements must be implemented.

Centrality

Centrality is one of the most used concepts of the network analysis. The fundamental idea here is: if a node hold a central position in its network , it will have more access to resources or will be more requested or accessed from other nodes. Basically, the centrality measures quantify how close the nodes are connected, on a direct or indirect way. Finding out which is the most central node is important, once it could help disseminate information in the network faster, stopping epidemics, protect a network from breaking at all or find some risky/weak points which need to be watched carefully, for example.

But, who is central to a social network? It might not be that obvious. It depends on the centrality measure we are using. The main network metrics about centrality are degree centrality, closeness centrality and betweenness centrality. We have them explained in the following sections.

Degree centrality

The simples measure of centrality. It consists in counting the number of direct connections of a node divided by the total number of nodes in the network(normalized measure).The degree of connectivity might be sub-classified in the Degree-in or Degree-out (C_{DIn} and C_{Dout}), if the network or graph is directed.

A node is considered disconnected if it does not have any connections. One might think degree centrality as a local centrality, once it measures only direct connections among one node and the others, not considering indirect connections (just the immediate ones). In this sense, the degree of a vertex in a network is the number of edges attached to it. In mathematical terms, the degree centrality C_D of a vertex i is:

$$C_D = \frac{\sum_{j=1}^{n} A_i j}{n-1}$$

Where $A_i j = 1$, if there is an edge between the nodes i and j.

In terms of information systems networks, we could think on an information system which is accessed very often by many other systems of the IT Landscape. According to the level of granularity adopted in the analysis, the node “ERP” in the Figure A.4 could represent an ERP system, for example, due to its central hole on the hypothetical network:

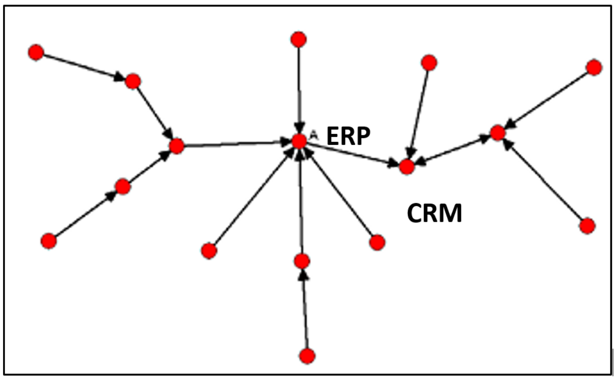

Figure A.4: Node ERP is the most central system on the network

On the perspective of IT Landscape yet, a considerable C_{DIn} may also point to a certain IT value: examples may be business intelligence and premium collection applications that collect, aggregate, and process data from many different sources to support activities of great strategic or monetary relevance (SIMON; FISCHBACH, 2013). Centrality of Degree also can be high for a premium collection of programs and policy applications which are very often requested, for example (high C_{DIn}).

A VARIATION OF DEGREE CENTRALITY: BONACICH'S APPROACH

The original degree centrality approach argues that actors who have more connections are more likely to be powerful because they can directly affect more other actors (HANNEMAN; RIDDLE, 2005). This makes sense, but having the same degree does not necessarily make actors equally important. For example, on the graph illustrated in Figure A.5, clearly the nodes A and B have different values for the network although they present the same centrality degree value.

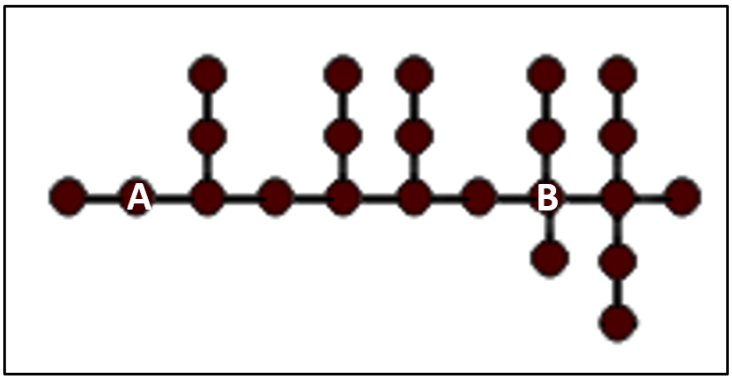

Figure A.5: Bonanich's centrality demonstration

BONACICH (1987) argued that one's centrality is a function of how many connections one has, and how many the connections the actors in the neighborhood had. In simple terms, this measure takes into account the amount and type of connections pursued by the neighbors of the vertex of interest. The formula for Bonacich's centrality can be seeing below:

$$c_i(\alpha,\beta) = \sum_j (\alpha - \beta c_j) R_{\text{i,j}}$$

The C_i, centrality or power of node *i*, generates a value for each node, which values depend on the structure of the network and the value of β. The β parameter might be thought of as characterizing the climate or culture of the network. A positive value for β $(+k)$ means the anyone who knows powerful others is made more powerful as a result, while a negative value $(-k)$ makes those connected to powerful nodes weaker. R is an adjacency matrix, and α is a scaling factor. Solutions are suggested for values of β in the range $(-k, +k)$ where $k = \frac{1}{\upsilon}$ and υ is the largest Eigenvalue in the solution of $\lambda x = Rx$. With β=0, the measure is equivalent to degree centrality; at β = k, it is equivalent to eigenvector centrality. Setting $\beta = -k$ yields values consistent with Cook, Emerson et al.'s prediction regarding power (COOK et al., 1983; EMERSON, 1962).

CLOSENESS CENTRALITY

Degree centrality measures might be criticized because they only take into account the immediate ties that an actor has, or the ties of the actor's neighbors, rather than indirect ties to all others (HANNEMAN; RIDDLE, 2005). One actor might be tied to a large number of others, but those others might be rather disconnected from the network as a whole. In a case like this, the actor could be quite central, but only in a local neighborhood.

FREEMAN (1978) expanded the concept of centrality , when proposed the term global centrality or closeness, which takes into account the distance between the node N to all of its neighbors(besides the ones immediately connected). With this measure, the closeness centrality can defined as the average of geodesic distances from a node *N* to all other nodes of the network. The node *N* with the best closeness is the one which has the less average distance to others, meaning that it has the best access to all other nodes, in average. In another hand, a node with bad closeness centrality might indicate one which is located in the periphery of the network. The closeness centrality – C_c - can be expressed formally as:

$$C_c(i) = [\sum_{j=1}^{N} d(i,j)]^{-}1$$

Where $d(i,j)$ is the geodesic distance between the nodes *i* and *j*. Normalizing C_c $(0 < C_c < 1)$ with the number of nodes except *i*:

$$C'_c(i) = \frac{C_c(i)}{N-1}$$

An entity with a high closeness centrality generally:

- Has quick access to other entities in a network.

- Has a short path to other entities.
- Is close to other entities.
- Has high visibility as to what is happening in the network.

Thinking about information systems and closeness, systems with high values for this measure could have their information spread through another systems more often, on a direct or indirect way. Assuming that cross-functional applications (e.g., enterprise resource planning) are part of the landscape's shared core and thus are connected to many other applications either directly or at least indirectly, CC seems a reasonable measure for indicating the initiative's cross functional character as it can capture the mean distance of an application to all others. Depending on the number of applications with a $CCvalue$ that exceeds a pre-defined threshold, the decision about the EA involvement could be made (SIMON; FISCHBACH, 2013).

EIGENVECTOR CENTRALITY

The eigenvector approach is an effort to find the most central actors (i.e. those with the smallest farness from others) in terms of the "global" or "overall" structure of the network, and to pay less attention to patterns that are more "local" (HANNEMAN; RIDDLE, 2005).

The eigenvector centrality defined in this way accords each vertex a centrality that depends both on the number and the quality of its connections: having a large number of connections still counts for something, but a vertex with a smaller number of high-quality contacts may outrank one with a larger number of mediocre contacts.

An eigenvector (an algebra concept) can be multiplied by the adjacency matrix representing a graph and return itself multiplied by a scalar. It is a characteristic of this particular adjacency matrix. The elements of this vector are the Eigenvector centralities of the vertices of the graph and give the relative accessibilities of the vertices.

$$Mx = \lambda x$$

Where M is the adjacency matrix of the Graph, x is the eigenvector (containing the centralities values) and λ is the eigenvalue (a scalar value.)

A variant of eigenvector centrality is employed by the well-known Web search engine Google to rank Web pages. In the context of information systems , systems with high values for this kind of centrality may represent the core modules or control points in the IT landscape.

BETWEENNESS CENTRALITY

Another measure defined by Freeman was the **betweenness centrality**. This measure calculates the degree of intermediation of a node between two structures, subgraphs or subnetworks. In the Figure A.6 bellow, we can see that despite the node A has lower local centrality than node B, for example, node A has a greater value for betweenness centrality, once it connects two important subgraphs.

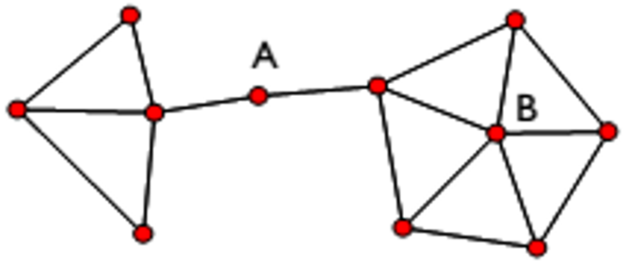

Figure A.6: Differences between high local centrality (node B) and high betweenness centrality (node A)

Mathematically, the betweenness centrality of vertex v is the fraction of geodesic paths between other vertices that v falls on. That is, we find the shortest path between every pair of vertices, and ask on what fraction of those paths vertex v lays. A formal definition can be:

$$C_B(v) = \frac{\sigma_{\mathrm{st}}(v)}{\sigma_{\mathrm{st}}}, s \neq v \neq t$$

Where σ_{st} is the total number of shortest paths from all nodes s to all nodes t and $\sigma_{\mathrm{st}}(v)$ is the number of those paths that pass through v. The node which more appears in the geodesic paths of each pair of nodes , has the highest betweenness.

Betweenness is a crude measure of the control v exerts over the flow of information (or any other commodity) between others. If we imagine information flowing between individuals in the network and always taking the shortest possible path, then betweenness centrality measures the fraction of that information that will flow through i on its way to wherever it is going (NEWMAN, 2003).

An entity with a high betweenness centrality generally (SENTINEL, 2014):

- Holds a favored or powerful position in the network.
- Represents a single point of failure—take the single betweenness spanner out of a network and you sever ties between cliques.
- Has a greater amount of influence over what happens in a network.

It is in most cases only an approximation to assume that information flows along geodesic paths; normally it will not, and variations of betweenness centrality such as “flow betweenness” and “random walk betweenness” have been proposed to allow for this. In many practical cases however, the simple (geodesic path) betweenness centrality gives quite informative answers (NEWMAN, 2003).

In IT landscape, C_B can indicate, for example, a sales support application to be significant that did not have a high C_D at all, like the Node A in the Figure A.6. This application might be important due to automatic forwarding of application data, which plays a crucial role in achieving high quantities of new policies daily, for example (SIMON; FISCHBACH, 2013).

GROUPS AND SUB-STRUCTURES

Divisions of actors into groups, communities, sub-structures or sub-graphs can be a very important aspect of social structure. The problem of community detection requires the partition of a network into communities of densely connected nodes, with the nodes belonging to dierent communities being only sparsely connected (Bondel et al., 2009). Putting in a different way, looking at the whole network, we can think of sub-structures as areas of the graph that seem to be locally dense, but separated to some degree, from the rest of the graph. In a sense, this more macro lens is looking for "holes" or "vulnerabilities" or "weak spots" in the overall structure or solidarity of the network.

Main questions about a graph, in terms of its sub-structures, may arise (HANNEMAN; RIDDLE, 2005):

- How separate are the sub-graphs? Do they overlap and share members, or do they divide or factionalize the network?
- How large are the connected sub-graphs? Are there a few big groups, or a larger number of small groups?
- Are there particular actors that appear to play network roles? For example, act as nodes that connect the graph, or who are isolated from groups?

Past work on methods for discovering groups in networks divides into two main lines of research. The first, which goes by the name of graph partitioning, has been pursued particularly in computer science and related fields. The second, identified by names such as block modeling, hierarchical clustering or community structure detection, has been pursued by sociologists and more recently by physicists and applied mathematicians, with applications especially to social and biological networks (NEWMAN, 2003) SCOTT, 2000).

Some common techniques/algorithms used to detect community structures are listed on the work of Lancichinetti and Fortunato(2009):

- Spectral bi-section algorithm
- Kernighan-Lin algorithm
- Hierarchical clustering (single linkage and complete linkage) and its similarity measure(Euclidean, structural equivalence, K-independent paths)
- Methods based on Edge removal:

- GIRVAN; NEWMAN (2002) algorithm and the edge betweenness
- Tyler et. Al (2002)

The proper explanation of these techniques is beyond the scope of this work. Once these techniques are applied, we are able to find several types of structures as: components, blocks/cutpoints, K-cores, Lambda sets and bridges, factions, and f-groups (HANNEMAN; RIDDLE, 2005) each one with different meanings. Some of them are presented below (SCOTT, 2000).

COMPONENTS

Formally, a Component is a maximal connected sub-graph. Components of a graph are sub-graphs that are connected within, but disconnected between sub-graphs. A sub-graph, like a graph, is connected when **all of its points are linked to one another through one or more paths** and they have no connections outside the sub-graph. A basic step in structural description of a network is to identify the number and size of its components.

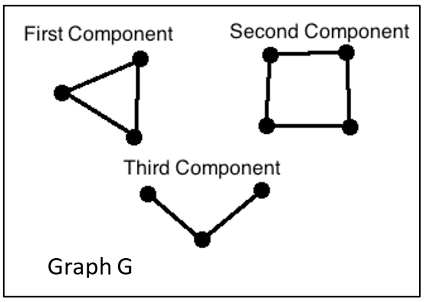

Figure A.7: Graph G and its four components

The result of a component analysis is a view of the graph as composed of one or more components, hangers, bridgers, and a number of isolated points. Directed components can either be classified as strong or weak, according to the flow of interest.

MODULARITY

Now that we know about some techniques which could help us to find some structures or communities in the networks, one pertinent question comes up: How can we evaluate the quality of the communities or structures detected? Some techniques to evaluate the partitions are **modularity measures**. The definition and application of the modularity is independent of the particular community structure algorithm used, and it a therefore also be applied to any other algorithm.

Modularity is one measure of the structure of networks graphs which calculates the strength of division of a network into modules (also called groups, clusters or communities).

Clusters or groups with high modularity have dense connections between the nodes within modules but sparse connections between nodes in different modules. On the Figure A.8 is showed the variation of densities over the network. In this case, it is clearly the existence of the two specific structures.

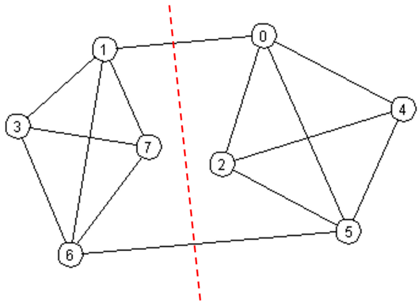

Figure A.8: Partition of Graph in two structures according to their internal densities.

It may be imperative to identify the communities in networks since the communities may have quite different properties such as node degree, clustering coefficient, betweenness, and centrality, etc., from that of the average network. Modularity is one such measure, which when maximized, leads to the appearance of communities in a given network. One modularity measure is presented below:

$$Q = \frac{1}{2m}\sum_{i,j}[A_{\mathrm{i,j}} - \frac{k_i k_j}{2m}]\sigma(c_i, c_j)$$

Where:

A_{ij}: represents the weight of the edge between i and j,

k_i: is the sum of the weights of the edges attached to vertex i,

c_i: is the community u or v to which vertex i is assigned,

$\sigma(u,v)$: is 1 when $u = v$, otherwise 0,

m: sum of all edge's weight/2

Thus, the modularity of a partition is a scalar value between -1 and 1that measures the density of links inside communities as compared to links between communities.

So far, we have seen general and broad analysis techniques which might tell us about the macro-structure of the graph. Now, we might want to take a look on the micro-structures as well. On the following sections, botton-up approaches to detect communities/groups.

CLIQUES

According to HANNEMAN; RIDDLE (2005), a clique is the maximum number of actors who have all possible ties present among themselves. A "Maximal complete sub-graph" is such a grouping, expanded to include as many actors as possible.

The smallest "cliques" are composed of two actors: the dyad. But dyads can be "extended" to become more and more inclusive - forming strong or closely connected regions in graphs. A number of approaches to finding groups in graphs can be developed by extending the close-coupling of dyads to larger structures. The Figure A.9 bellow shows some possibilities of cliques with 3,4 and 5 nodes

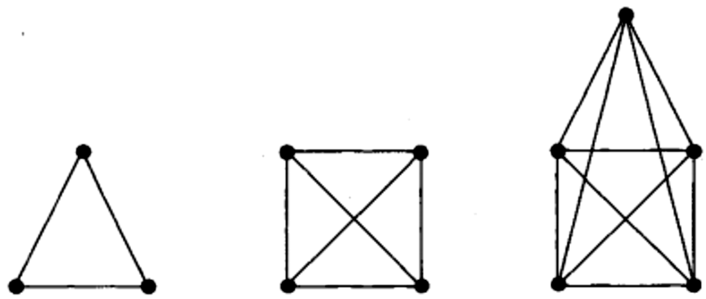

Figure A.9: Cliques of 3, 4 and 5 nodes

With this detection approach the analyst might reveal the existence of subgroups whose characteristics can be analyzed further in their particularities. Some analysis possibilities with UCINET (SCOTT, 2000) about cliques are:

- Clique and actor-by-clique analysis of reciprocity: Allow us to analyze how "adjacent" each actor is to each clique. In this sense, we could investigate if a specific software component has dependencies with other modules, for example.
- Actor-by-actor analysis of reciprocity: This indicates how many cliques each pair of actors are both members of.
- Clique-by-clique analysis of reciprocity

We agree with HANNEMAN; RIDDLE (2005) that the definition of a clique is very strong, especially in case of real networks. Thus, others structures more flexible can be of interest for analysis as well. N-cliques, Clans, K-plexes and K-cores are some of them. They are explained as follows.

K-CORES

A k-core is a maximal group of actors, all of whom are connected to some number (*k*) of other members of the group. All the nodes in a k-Core have a degree greater than or equal to *k*.

The k-core approach is more relaxed, allowing actors to join the group if they are connected to k members, regardless of how many other members they may not be connected to. By varying the value of k (that is, how many members of the group do you have to be connected to), different pictures can emerge. In order to form a 2k-core, for example, all nodes with degree 1 are ignored and the structure of connection of the remaining node is examined.

K-cores can be (and usually are) more inclusive than k-plexes. And, as *k* becomes smaller, group sizes will increase. The idea with k-cores is to study areas of high and low cohesion.

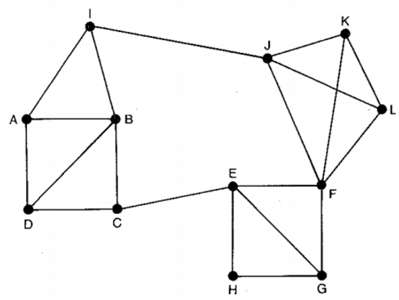

Figure A.10: Example of a 3k-core (Scott, 2010)

B

The EANA meta-model and library's evaluation survey

Introductory document sent to the respondents

In order to produce an initial conceptual alignment among the respondents, the following document was sent by email to the respondents together with a link for the web survey hosted at Surveymonkey platform[1].

B.1 Introduction

In this survey, you are kindly invited to evaluate two artifacts designed based on our previous literature review outputs (Santana A.; Fischbach K.; Moura H., 2016a): the meta-model and the goal-question-metric based library for EA network analysis, which are detailed in the following.

B.2 The EA network analysis meta-model (EANA-MM)

The EANA meta-model was constructed using the concepts taken from EA body of knowledge and Graph theory. We present the meta-model in Figure B.1 and discuss its conceptual elements in the following.

We summarize the several conceptual elements involved in the EA network analysis introduced above in Table B.1:

[1] https://www.surveymonkey.com

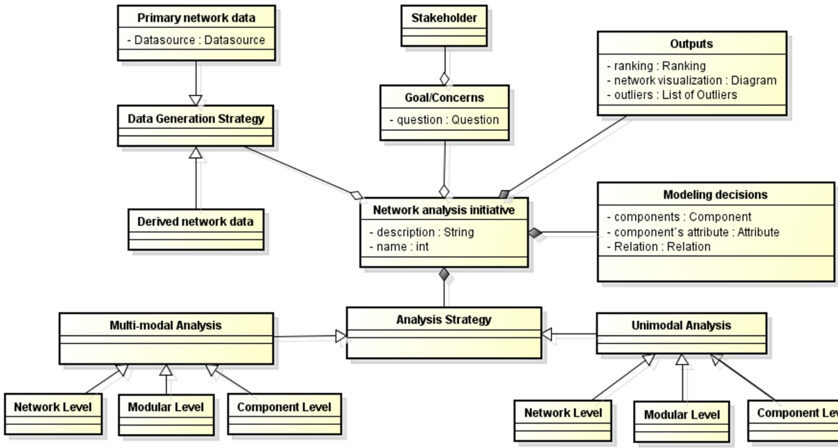

Figure B.1: EANA-MM

EANA Concept	Description
Network analysis initiative(NAI)	Each attempt to analyze a particular concern in EA modeled as a network. This can be a simple, well-known network measure, more sophisticated techniques or methods or analysis inspired by network models
Analysis concern	An aspect of the EA components or the EA as whole that is investigated/measured by the experts in order to help EA evolution, to communicate with stakeholders and/or to ensure business-IT alignment.
Information requirements	The information needed to build the model the EA such as components, edges´ semantics, edges´ weight, and EA layer(s) used.
Network level of analysis	The chosen structural level to analyze the network: component, module or group and network level (overall network)
Unimodal analysis	Analysis strategy which uses a network composed of just of type of component(e.g. application)
Bi-modal analysis	Analysis strategy which uses a network composed of two types of components(e.g. application and business process)
Multi-modal	Analysis strategy which uses a network composed of more than two types of components(e.g. technology, application and business process)
Data generation strategy – Primary data	It is related to approaches used to generate EA models from scratch, using documents, interviews or automated collection of EA data
Data generation strategy – Derived data	This approach uses primary data and some derivation mechanisms to create artificial edges between components and build derived networks of EA data.
Stakeholder	The user or stakeholder that might benefit from the results generated by the network analysis initiatives.

Table B.1: Summary of the EANA conceptual elements

B.3 THE GQM LIBRARY

We structured the NAIs found in a Goal-question-metric based model. The idea is to help practitioners to operationalize the use of those NAIs to help them to perform EA structural analysis.

For the NAI library we consider the following concepts:

- A goal/concern An aspect of the EA components or of the EA as whole that is investigated/measured by the experts. An example of concern is: to be aware of the key components which occupy important structural positions in the network.
- Question For each NAI, we identified the analysis goal proposed by the researcher in their primary studies and then, we describe it as a question. Such questions are designed based on our interpretation about the analysis goal claimed by the researchers in the papers. An example of a question is: What are the key structural components of the application layer?
- Network analysis initiative (metric) Each attempt to analyze a concern in EA modeled as a network. This can be a simple well-known network measure, more sophisticated techniques, or analysis inspired by network models. For example, to operationalize the modularity calculation of an EA layer, one can apply modularity measures. In this document, we instantiate our EANA meta-model to classify the "Coupling" metric proposed in DREYFUS; WYNER (2011). The information extracted regarding that metric is then organized and showed in Table B.2.

Finally, in Figure B.2, we present graphically the decision support mechanism implicit in the NAI GQM based library proposed for that identified NAI:

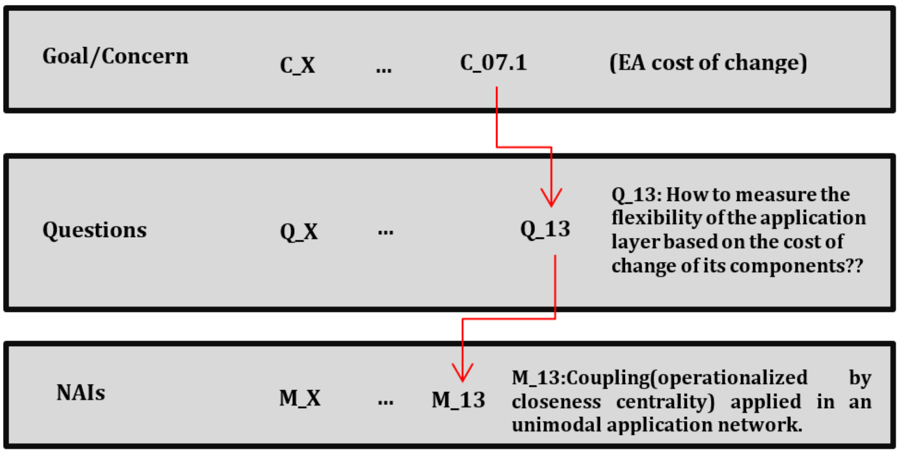

Figure B.2: GQM-based EANA library – a graphical representation for the "Coupling" NAI

We did the same procedure described in Table B.2 for each of the 74 NAIs which compose the EANA library. We believe that such description, together with the decision schema

1 - NAI (network analysis initiative)			
ID	Name(label)		
M_13	Coupling(operationalized by closeness centrality)		
2- Analysis strategy			
1- EA Layer(s):Value, Business Information, Application or Technology. It can be one or more layers	1- EA components(s):	1 - Component´s attribute(s)	1- Level of analysis (Network level, group level, component level)
Application	Application	None	Component level
3- Analysis Concern			
ID_Concern	3- EA concern-goal	ID_Question	3- EANA Question
C_07.1	EA cost of change	Q_13	How to measure the flexibility of the application layer based on the cost of change of its components?
4- Data strategy			
4- EA Data generation strategy (Primary data; Derived data; Simulated data; No data)	4 - Unimodal or Multi-modal?		4- Data source (Documents, Models, Diagrams, interviews, tools?)
Primary data	Unimodal		Documents, interviews
5 - Modeling decisions			
Relation meaning	Directed or undirected data	Transitive relations	Weighted?
Application depends on Application	directed	Yes	no
6 - Output			
List of applications			
7 - Stakeholder			
Application architect, CO			

Table B.2: EANA meta-model instantiation

depicted in Figure B.2, may help researchers and practitioners to map **which EA concerns** to analyze in terms of structure, how to do it (**which NAI** to choose), which **level of analysis, data strategy** to apply; **identify stakeholders** who can benefit and all other **information** needed to run those metrics or methods.

With the present survey, we aim to evaluate the efficacy and other criteria of the GQM-based EANA library and EANA-MM.

B.4 The survey instrument

Tables B.3, B.4 and B.5 reproduce the web survey implement and answered by the respondents.

1 - Respondent background	
Years of experience:	
In industry	
In Academy	

Table B.3: Respondent's background form

2- EANA meta-model - assessment
Efficacy
Description: With this meta-model, it is possible to classify existent network analysis initiatives. The meta-model also helps in understanding which elements one needs to develop new analysis initiatives.
Score: ( ) 2 "I Strongly disagree" ( ) 1 "I disagree" ( ) 0 " I am not sure" ( ) 1 " I agree" ( ) 2 "I strongly agree"
General comments for the criteria
Generality
Description: With this meta-model, it is possible to classify a broad variety of network analysis initiatives.
Score: ( ) 2 "I Strongly disagree" ( ) 1 "I disagree" ( ) 0 " I am not sure" ( ) 1 " I agree" ( ) 2 " I strongly agree"
General comments for the criteria
Utility
Description: I could use this meta-model to frame any method developed for EA network analysis research or document new ones.
Score: ( ) 2 "I Strongly disagree" ( ) 1 "I disagree" ( ) 0 " I am not sure" ( ) 1 " I agree" ( ) 2 " I strongly agree"
General comments for the criteria
Understandability
Description: I can clearly understand the concepts which integrate the model and the relations among them.
Score: ( ) 2 "I Strongly disagree" ( ) 1 "I disagree" ( ) 0 " I am not sure" ( ) 1 " I agree" ( ) 2 " I strongly agree"
General comments for the criteria
Continued on next page

Table B.4: EANA meta-model - assessment

Table B.4 – continued from previous page

2- EANA meta-model - assessment
Consistency
Description: This meta-model is consistent, that is, its parts are harmonic to one another and as a role.
Score: ( ) 2 “I Strongly disagree” ( ) 1 “I disagree” ( ) 0 “ I am not sure” ( ) 1 “ I agree” ( ) 2 “ I strongly agree”
General comments for the criteria
Level of details
Description: The level of details of the meta-model is suitable to its goal, that is, to classify or design network analysis initiatives for the EA context.
Score: ( ) 2 “I Strongly disagree” ( ) 1 “I disagree” ( ) 0 “ I am not sure” ( ) 1 “ I agree” ( ) 2 “ I strongly agree”
General comments for the criteria

Table B.4: EANA meta-model - assessment

3- EANA library - assessment
Criteria of assessment
Efficacy
Description: Organizing the existent network analysis initiatives for EA as a GQM library helps the expert to focus on a specific concern and capture the relevant data for the analysis.
Score: ( ) 2 “I Strongly disagree” ( ) 1 “I disagree” ( ) 0 “ I am not sure” ( ) 1 “ I agree” ( ) 2 “ I strongly agree”S
General comments for the criteria
Utility
Description: Once each analysis initiative is implemented and available, this library can be used as a toolbox for experts to analyze several EA structural concerns.
Score: ( ) 2 “I Strongly disagree” ( ) 1 “I disagree” ( ) 0 “ I am not sure” ( ) 1 “ I agree” ( ) 2 “ I strongly agree”
General comments for the criteria
Level of details
Description: The level of details of each network analysis initiative description is suitable for one to know what is needed execute it.
Continued on next page

Table B.5: EANA library evaluation form

Table B.5 – continued from previous page

3- EANA library - assessment
Score: ( ) 2 “I Strongly disagree” ( ) 1 “I disagree” ( ) 0 “ I am not sure” ( ) 1 “ I agree” ( ) 2 “ I strongly agree”
General comments for the criteria
Applicability
Description: Together with the meta-model, such a catalog of measures helps practitioners in applying proper structural metrics and in collecting proper data in real cases.
Score: ( ) 2 “I Strongly disagree” ( ) 1 “I disagree” ( ) 0 “ I am not sure” ( ) 1 “ I agree” ( ) 2 “ I strongly agree”
General comments for the criteria

Table B.5: EANA library evaluation form

C

EA ANALYSIS SLR - LIST OF PRIMARY STUDIES

ID-3	Reference
1	PLATANIOTIS, G. et al. A Conceptual Model for Compliance Checking Support of Enterprise Architecture Decisions. Business Informatics (CBI), 2015 IEEE 17th Conference on. Anais...2015
2	ID3_2 – LANGERMEIER, M.; SAAD, C.; BAUER, B. Adaptive Approach for Impact Analysis in Enterprise Architectures. In: Lecture Notes in Business Information Processing. [s.l: s.n.]. v. 220p. 22–42.
3	Wood, J., Sarkani, S., Mazzuchi, T., & Eveleigh, T. (2013). A framework for capturing the hidden stakeholder system. Systems Engineering, 16(3), 251-266.
4	JAVANBAKHT, M.; POURKAMALI, M.; FEIZI, M. A new method for enterprise architecture assessment and decision-making about improvement or redesign. 4th International Multi-Conference on Computing in the Global Information Technology, ICCGI 2009, p. 69–76, 2009.
6	FASANGHARI, M. et al. A novel credibility-based group decision making method for Enterprise Architecture scenario analysis using Data Envelopment Analysis. Applied Soft Computing Journal, v. 32, p. 347–368, 2015.
8	ZACARIAS, M. et al. Adding a human perspective to enterprise architectures. Proceedings - International Workshop on Database and Expert Systems Applications, DEXA, n. 2, p. 840–844, 2007.
10	TARENSKEEN, D.; BAKKER, R.; JOOSTEN, S. Applying the V Model and Axiomatic Design in the Domain of {IT} Architecture Practice. Procedia {CIRP}, v. 34, p. 263–268, 2015.
Continued on next page	

Table C.1: EA Analysis SLR primary studies

Table C.1 – continued from previous page

ID-3	Reference
11	CHEN, H. M.; KAZMAN, R.; GARG, A. BITAM: An engineering-principled method for managing misalignments between business and IT architectures. Science of Computer Programming, v. 57, n. 1, p. 5–26, 2005.
12	IMACHE, R.; IZZA, S.; AHMED-NACER, M. Clustering-based urbanisation to improve enterprise information systems agility. Enterprise Information Systems, v. 9, n. 8, p. 861–877, 2015.
14	KRUIZE, J. W. et al. Improving arable farm enterprise integration – Review of existing technologies and practices from a farmer's perspective. Computers and Electronics in Agriculture, v. 96, p. 75–89, 2013.
15	GARG, A.; KAZMAN, R.; CHEN, H. M. Interface descriptions for enterprise architecture. Science of Computer Programming, v. 61, n. 1, p. 4–15, 2006.
18	DAVOUDI, M. R.; ALIEE, F. S. Characterization of enterprise architecture quality attributes. Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOC, p. 131–137, 2009.
19	GRANDRY, E.; FELTUS, C.; DUBOIS, E. Conceptual Integration of Enterprise Architecture Management and Security Risk Management. Enterprise Distributed Object Computing Conference Workshops (EDOCW), 2013 17th IEEE International. Anais...2013
21	FRANKE, U. Enterprise Architecture Analysis with Production Functions. Enterprise Distributed Object Computing Conference (EDOC), 2014 IEEE 18th International. Anais...2014
22	ADDICKS, J. S. Enterprise architecture dependent application evaluations. Digital Ecosystems and Technologies, 2009. DEST '09. 3rd IEEE International Conference on. Anais...2009
23	ARCHITECTURE, C. E.; DATA, O. Enterprise Architecture Intelligence. 2014.
24	ERNST, A. M. Enterprise architecture management patterns. In Proceedings of the 15th Conference on Pattern Languages of Programs , v. 93, p. 20, 2008.
25	SAAT, J. et al. Enterprise Architecture Meta Models for IT/Business Alignment Situations. Enterprise Distributed Object Computing Conference (EDOC), 2010 14th IEEE International. Anais...2010

Continued on next page

Table C.1: EA Analysis SLR primary studies

Table C.1 – continued from previous page

ID-3	Reference
26	ANTUNES, C.; CAETANO, A.; BORBINHA, J. Enterprise Architecture Model Analysis Using Description Logics. 2014 IEEE 18th International Enterprise Distributed Object Computing Conference Workshops and Demonstrations. Anais...IEEE, set. 2014
27	FLOREZ, H.; SANCHEZ, M.; VILLALOBOS, J. Extensible Model-Based Approach for Supporting Automatic Enterprise Analysis. Proceedings . IEEE 18th international Enterprise Distributed object computing conference, v. 2014–December, n. December, p. 32–41, 2014.
28	AIER, S. How Clustering Enterprise Architectures helps to Design Service Oriented Architectures. Services Computing, 2006. SCC '06. IEEE International Conference on, p. 269–272, 2006.
30	SUNKLE, S. et al. Incorporating Directives into Enterprise TO-BE Architecture. Enterprise Distributed Object Computing Conference Workshops and Demonstrations (EDOCW), 2014 IEEE 18th International. Anais...2014
31	MAROSIN, D.; GHANAVATI, S. Measuring and managing the design restriction of enterprise architecture (EA) principles on EA models. Requirements Engineering and Law (RELAW), 2015 IEEE Eighth International Workshop on. Anais...2015
32	PAPADACCI-STEPHANOPOLI, E.; SALINESI, C.; ROLLAND, C. NENO process: Information systems arbitration process in Enterprise Architecture Project. Information and Communication Technologies, 2006. ICTTA '06. 2nd. Anais...2006
33	GOIKOETXEA, A. a Mathematical Framework for Enterprise Architecture Representation and Design. International Journal of Information Technology and decision making, v. 3, n. 1, p. 5–32, 2004.
36	GAALOUL, K.; GUERREIRO, S. A Decision-Oriented Approach Supporting Enterprise Architecture Evolution. Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), 2015 IEEE 24th International Conference on. Anais...2015GAALOUL, K.; GUERREIRO, S. A Decision-Oriented Approach Supporting Enterprise Architecture Evolution. Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), 2015 IEEE 24th International Conference on. Anais...2015
37	POPOVA, V.; SHARPANSKYKH, A. A Formal Framework for Modeling and Analysis of Organizations. Knowledge Creation Diffusion Utilization, v. 244, n. June, p. 343–358, 2007.

Continued on next page

Table C.1: EA Analysis SLR primary studies

Table C.1 – continued from previous page

ID-3	Reference
38	RICO, D. F.; ENGINEER, S. A Framework for Measuring ROI. [s.d.].
39	J., U. et al. A framework for service interoperability analysis using enterprise architecture models. Services Computing, 2008. SCC '08. IEEE International Conference on. Anais...2008
40	RAZAVI, M.; ALIEE, F. S.; TAFRESHI, A. E. S. A Fuzzy AHP Based Approach Towards Enterprise Architecture Evaluation. PROCEEDINGS OF THE 3RD EUROPEAN CONFERENCE ON INFORMATION MANAGEMENT AND EVALUATION. Anais...2009
42	PEPIN, J. et al. A Method for Business-IT Alignment of Legacy Systems. p. 229–237, 2015.
43	HORST, J. TER. A Method For Enterprise Architecture. Thesis, v. 120 LNBIP, p. 48–76, 2012.
44	M.A, L.-V. et al. A model for enterprise architecture scenario analysis based on fuzzy cognitive maps and OWA operators. CONIELECOMP 2014 - 24th International Conference on Electronics, Communications and Computers, p. 243–247, 2014b.
47	BUSCHLE, M. et al. A Tool for Enterprise Architecture Analysis Using the PRM Formalism. (E. Soffer, P and Proper, Ed.)INFORMATION SYSTEMS EVOLUTION. Anais...: Lecture Notes in Business Information Processing.2011
48	BUCKL, S. et al. A Pattern-based Approach to Quantitative Enterprise Architecture Analysis. Proceedings of the Fifteenth Americas Conference on Information Systems. Anais...
51	LANGERMEIER, M.; SAAD, C.; BAUER, B. Adaptive Approach for Impact Analysis in Enterprise Architectures. In: Lecture Notes in Business Information Processing. [s.l: s.n.]. v. 220p. 22–42.
56	CAPIROSSI, J. A; RABIER, P. B. An enterprise architecture and data quality framework. Advances in Intelligent Systems and Computing, v. 205 {AISC}, p. 67–79, 2013.
57	HÖÖK, D. et al. An Enterprise Architecture Based Method Enabling Quantified Analysis of IT Support System ' s Impact on Maintenance Management. p. 3176–3189, 2009.

Continued on next page

Table C.1: EA Analysis SLR primary studies

Table C.1 – continued from previous page

ID-3	Reference
58	SUBRAMANIAN, N.; CHUNG, L.; SONG, Y. T. An NFR-based framework for establishing traceability between enterprise architectures and system architectures. Proc. - Seventh ACIS Int. Conf. on Software Eng., Artific. Intelligence, Netw., and Parallel/Distributed Comput., SNPD 2006, including Second ACIS Int. Worshop on SAWN 2006, v. 2006, p. 21–28, 2006.
59	CASTELLANOS, C.; CORREAL, D.; MURCIA, F. An ontology-matching based proposal to detect potential redundancies on enterprise architectures. Proceedings - International Conference of the Chilean Computer Science Society, SCCC, p. 118–126, 2012.
61	SUNKLE, S.; KULKARNI, V.; ROYCHOUDHURY, S. Analyzing Enterprise Models Using Enterprise Architecture-Based Ontology. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). [s.l: s.n.]. v. 8107p. 622–638.
62	LAGERSTRÖM, R.; JOHNSON, P.; EKSTEDT, M. Architecture analysis of enterprise systems modifiability: A metamodel for software change cost estimation. Software Quality Journal, v. 18, n. 4, p. 437–468, 2010.
63	ALVES, J.; VASCONCELOS, A.; SOUSA, P. Architecture Principles Compliance Analysis. Proceedings of the 16th International Conference on Enterprise Information Systems, p. 328–334, 2014.
64	GAMMELGÅRD, M.; EKSTEDT, M.; NÄRMAN, P. Architecture Scenario Analysis – Estimating the Credibility of the Results. INCOSE International Symposium, v. 17, n. 1, p. 1399–1413, 2007.
65	OSTERLIND, M.; LAGERSTROM, R.; ROSELL, P. Assessing Modifiability in Application Services Using Enterprise Architecture Models - A Case Study. (J. Aier, S and Ekstedt, M and Matthes, F and Proper, E and Sanz, Ed.)TRENDS IN ENTERPRISE ARCHITECTURE RESEARCH AND PRACTICE-DRIVEN RESEARCH ON ENTERPRISE TRANSFORMATION. Anais...: Lecture Notes in Business Information Processing.2012
66	RADERIUS, J.; NARMAN, P.; EKSTEDT, M. Assessing System Availability Using an Enterprise Architecture Analysis Approach. Service-Oriented Computing - Icsoc 2008 Workshops, v. 5472, p. 351–362, 2009.
67	GINGNELL, L. Availability of enterprise IT systems – an expert- based Bayesian model. 2012.

Continued on next page

Table C.1: EA Analysis SLR primary studies

Table C.1 – continued from previous page

ID-3	Reference
70	MANZUR, L.; SÁNCHEZ, M.; VILLALOBOS, J. Calculating indicators for simulation of IT architecture models. 2012 7th Colombian Computing Congress, CCC 2012 - Conference Proceedings, p. 1–6, 2012.
71	R.A, M. et al. Calculating risk of integration relations in application landscapes. Electro/Information Technology, 2007 IEEE International Conference on. Anais...2007
73	PLATANIOTIS, G.; DE KINDEREN, S.; PROPER, H. A. Capturing Decision Making Strategies in Enterprise Architecture - A Viewpoint. (I. Nurcan, S and Proper, HA and Soffer, P and Krogstie, J and Schmidt, R and Halpin, T and Bider, Ed.)ENTERPRISE, BUSINESS-PROCESS AND INFORMATION SYSTEMS MODELING, BPMDS 2013. Anais...: Lecture Notes in Business Information Processing.2013
75	DE BOER, F. S. et al. Change impact analysis of enterprise architectures. Information Reuse and Integration, Conf, 2005. IRI-2005 IEEE International Conference on., p. 177–181, 2005.
76	SOMMESTAD, T.; EKSTEDT, M.; JOHNSON, P. Combining defense graphs and enterprise architecture models for security analysis. Proceedings - 12th IEEE International Enterprise Distributed Object Computing Conference, EDOC 2008, p. 349–355, 2008.
78	SCHÜTZ, A.; WIDJAJA, T.; KAISER, J. Complexity in enterprise architectures: conceptualization and introduction of a measure from a system theoretic perspective. 21st European Conference on Information Systems, p. 12, 2013.
80	NARMAN, P. et al. Enterprise Architecture Analysis for Data Accuracy Assessments. 2009 IEEE International Enterprise Distributed Object Computing Conference, p. 24–33, 2009.
82	AIER, S. et al. Deriving SOA evaluation metrics in an enterprise architecture context. International Conference on Service-Oriented Computing. Anais...2007
83	FRANCALANCI, C.; PIURI, V. Designing information technology architectures: A cost-oriented methodology. Journal of Information Technology, 1999.
85	S.A, G. et al. Disentangling the Value of Information and Analytics through Componentized Business Architecture. Proceedings of the Annual Hawaii International Conference on System Sciences. Anais...jan. 2011

Continued on next page

Table C.1: EA Analysis SLR primary studies

Table C.1 – continued from previous page

ID-3	Reference
86	PLATANIOTIS, G.; DE KINDEREN, S.; PROPER, H. A. EA Anamnesis. International Journal of Information System Modeling and Design, v. 5, n. 3, p. 75–95, 2014.
89	NÄRMAN, P. et al. Data accuracy assessment using enterprise architecture. Enterprise Information Systems, v. 5, n. 1, p. 37–58, 2011.
90	Franke, U., Johnson, P., Ericsson, E., Flores, W. R., & Zhu, K. (2009, June). Enterprise architecture analysis using fault trees and MODAF. In Proceedings of CAiSE Forum (Vol. 453, pp. 61-66).
93	NÄRMAN, P. et al. Enterprise architecture availability analysis using fault trees and stakeholder interviews. Enterprise Information Systems, v. 8, n. 1, p. 1–25, 2014.
94	LEE, H. et al. Enterprise Architecture Content Model Applied to Complexity Management While Delivering IT Services. Services Computing (SCC), 2014 IEEE International Conference on. Anais...2014
96	ŠAŠA, A.; KRISPER, M. Enterprise architecture patterns for business process support analysis. Journal of Systems and Software, v. 84, n. 9, p. 1480–1506, 2011.
98	GUSTAFSSON, P., Johnson, P., & Nordstrom, L. (2009, June). Enterprise architecture: A framework supporting organizational performance analysis. In Electricity Distribution-Part 1, 2009. CIRED 2009. 20th International Conference and Exhibition on (pp. 1-4). IET.
99	NARMAN, P.; JOHNSON, P.; NORDSTROM, L. Enterprise Architecture: A Framework Supporting System Quality Analysis. Enterprise Distributed Object Computing Conference, 2007. EDOC 2007. 11th IEEE International, p. 130, 2007.
100	DREYFUS, D.; IYER, B. Enterprise Architecture: A Social Network Perspective. HICSS 2006 Proceedings, v. 0, n. C, p. 1–10, 2006.
104	KUMAR, A. et al. Enterprise Interaction Ontology for Change Impact Analysis of Complex Systems. 2008 IEEE Asia-Pacific Services Computing Conference, p. 303–309, 2008.
105	VAN DER RAADT, B.; SLOT, R.; VAN VLIET, H. Experience report: Assessing a global financial services company on its enterprise architecture effectiveness using NAOMI. Proceedings of the Annual Hawaii International Conference on System Sciences, p. 1–10, 2007.

Continued on next page

Table C.1: EA Analysis SLR primary studies

Table C.1 – continued from previous page

ID-3	Reference
106	YU, E.; STROHMAIER, M.; DENG, X. Exploring Intentional Modeling and Analysis for Enterprise Architecture. Enterprise Distributed Object Computing Conference Workshops, 2006. EDOCW ’06. 10th IEEE International. Anais...2006
108	JOHNSON, P. et al. Extended Influence Diagrams for Enterprise Architecture Analysis. Enterprise Distributed Object Computing Conference, 2006. EDOC ’06. 10th IEEE International. Anais...2006
109	ENGELSMAN, W. et al. Extending enterprise architecture modelling with business goals and requirements. ENTERPRISE INFORMATION SYSTEMS, v. 5, n. 1, p. 9–36, 2011.
110	JOHNSON, P.; NORDSTRÖM, L.; LAGERSTRÖM, R. Formalizing analysis of enterprise architecture. Enterprise Interoperability, New Challenges and Approaches, p. 1–10, 2007.
111	ZEE, M. VAN et al. Formalizing Enterprise Architecture Decision Models Using Integrity Constraints. 2014 IEEE 16th Conference on Business Informatics, p. 143–150, 2014.
112	DIEFENTHALER, P.; BAUER, B. From Gaps to Transformation Paths in Enterprise Architecture Planning. (J. Hammoudi, S and Cordeiro, J and Maciaszek, LA and Filipe, Ed.)ENTERPRISE INFORMATION SYSTEMS, ICEIS 2013. Anais...: Lecture Notes in Business Information Processing.2014
116	GUNGOR, C.; OGUZTUZUN, H. GoalDAG - ArchiMate Integration. Information and Software Technologies, Icist 2014, v. 465, p. 194–210, 2014.
117	SAFARI, H.; FARAJI, Z.; MAJIDIAN, S. Identifying and evaluating enterprise architecture risks using FMEA and fuzzy VIKOR. Journal of Intelligent Manufacturing, v. 27, n. 2, p. 475–486, 2016.
118	BINZ, T. et al. Improving the manageability of enterprise topologies through segmentation, graph transformation, and analysis strategies. Proceedings of the 2012 IEEE 16th International Enterprise Distributed Object Computing Conference, EDOC 2012, n. i, p. 61–70, 2012.
124	DAM, H. K.; LÊ, L.-S.; GHOSE, A. Managing changes in the enterprise architecture modelling context. Enterprise Information Systems, n. March, p. 1–31, 2015.

Continued on next page

Table C.1: EA Analysis SLR primary studies

Table C.1 – continued from previous page

ID-3	Reference
125	D., O. Misalignment symptom analysis based on enterprise architecture model assessment. Proceedings of the International Conferences on ICT, Society and Human Beings 2014, Web Based Communities and Social Media 2014, e-Commerce 2014, Information Systems Post-Implementation and Change Management 2014 and e-Health 2014 - Part of the Multi Conf. Anais...2014D
126	AIER, S.; SCHÖNHERR, M. Model Driven Service Domain Analysis 1 Challenges of Managing Complex Enterprise Architectures. p. 190–200, 2007.
127	MORKEVICIUS, A.; GUDAS, S.; SILINGAS, D. MODEL-DRIVEN QUANTITATIVE PERFORMANCE ANALYSIS OF UPDM-BASED ENTERPRISE ARCHITECTURE. (R. Targamadze, A and Butleris, R and Butkiene, Ed.)INFORMATION TECHNOLOGIES' 2010. Anais...: Information Technologies Conference Proceedings.2010
128	AZEVEDO, C. L. B. et al. Modeling resources and capabilities in enterprise architecture: A well-founded ontology-based proposal for ArchiMate. Information Systems, v. 54, p. 235–262, 2015.
129	GUSTAFSSON, P. et al. Modeling the IT Impact on Organizational Structure. 2009 IEEE International Enterprise Distributed Object Computing Conference, p. 14–23, 2009.
130	MAYER, R.; CAETANO, A. Ontology-based enterprise architecture model analysis Ontology-Based Enterprise Architecture Model Analysis. n. March, 2014.
131	VASCONCELOS, A. et al. Open issues on information system architecture research domain: The vision. ICEIS 2004 - Proceedings of the Sixth International Conference on Enterprise Information Systems, n. February 2016, p. 273–282, 2004.
133	NARANJO, D.; SÁNCHEZ, M.; VILLALOBOS, J. PRIMROSe - A Tool for Enterprise Architecture Analysis and Diagnosis. 16th International Conference on Enterprise Information Systems, n. January, p. 201–213, 2014.
134	BOUCHER, X. et al. Process clusters for information system diagnostics: an approach by Organisational Urbanism. Production Planning & Control, v. 22, n. 1, p. 91–106, jan. 2011.

Continued on next page

Table C.1: EA Analysis SLR primary studies

Table C.1 – continued from previous page

ID-3	Reference
135	VAN ECK, P.; BLANKEN, H.; WIERINGA, R. Project graal: Towards operational architecture alignment. INTERNATIONAL JOURNAL OF COOPERATIVE INFORMATION SYSTEMS, v. 13, n. 3, p. 235–255, 2004.
136	IACOB, M. E.; JONKERS, H. Quantitative analysis of enterprise architectures. (N. Konstantas, D and Bourrieres, JP and Leonard, M and Boudjlida, Ed.)Interoperability of Enterprise Software and Applications. Anais...2006
139	CHEN, W. et al. Semantic Enterprise Architecture Management. 15th International Conference on Enterprise Information Systems (ICEIS), p. 318–325, 2013.
142	FÜRSTENAU, D.; ROTHE, H. Shadow IT systems: Discening the good and the evil. Twenty Second European Conference on Information Systems, p. 1–14, 2014.
143	KREIMEYER, M.; GÜRTLER, M.; LINDEMANN, U. Structural metrics for decision points within Multiple-Domain Matrices representing design processes. 2008 IEEE International Conference on Industrial Engineering and Engineering Management, IEEM 2008, p. 435–439, 2008.
145	JABLOUN, M. et al. Supporting Enterprise Changes Using Actor Performance Assessment. (S. Toumani, F and Pernici, B and Grigori, D and Benslimane, D and Mendling, J and HadjAlouane, NB and Blake, B and Perrin, O and Saleh, I and Bhiri, Ed.)SERVICE-ORIENTED COMPUTING - ICSOC 2014 WORKSHOPS. Anais...: Lecture Notes in Computer Science.2015
146	STEINHORST, M. et al. Supporting Enterprise Transformation Using A Universal Model Analysis Approach. Ecis, n. 2012, p. 1–13, 2012.
147	PESSI, K. et al. Sustainable Alignment in Enterprise Architecture: A Case Study of Architectural Principles. (A. Abramowicz, W and Kokkinaki, Ed.)BUSINESS INFORMATION SYSTEMS WORKSHOPS (BIS 2014). Anais...: Lecture Notes in Business Information Processing.2014
153	M., B.; T., R.; F., M. Tool support for analyzing the evolution of Enterprise Architecture metrics. ICEIS 2015 - 17th International Conference on Enterprise Information Systems, Proceedings. Anais...2015Disponível em: ://www.scopus.com/inward/record.url?eid=2-s2.0-84939543169&partnerID=40&md5=ad47afcf11e4f751b5944ee14ce8e23b>

Continued on next page

Table C.1: EA Analysis SLR primary studies

Table C.1 – continued from previous page

ID-3	Reference
155	CHAARI, S. et al. Towards a service-oriented enterprise based on business components identification. In: Enterprise Interoperability II. London: Springer London, [s.d.]. p. 495–506.
157	BEETZ, K.; KOLBE, L. Towards Managing IT Complexity: An IT Governance Framework to Measure Business-IT Responsibility Sharing and Structural IT Organization. 17th Americas Conference on Information Systems AMCIS 2011, p. 1–12, 2011.
158	BUCKL, S. et al. Towards simulation-supported enterprise architecture management. [s.d.].
159	S.H., J.; A.M., S.; M., B. Traffic Flow Management (TFM) operational analysis through enterprise architecture (EA) activity modeling (OV-5) with Business Process Modeling Notation (BPMN). 12th AIAA Aviation Technology, Integration and Operations (ATIO) Conference and 14th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference. Anais...2012
161	JOHNSON, P. Using architectural models to predict the maintainability of enterprise systems. Software Maintenance and Reengineering, 2008. CSMR 2008. 12th European Conference on. Anais...2008
162	NARMAN, P. et al. Using enterprise architecture analysis and interview data to estimate service response time. JOURNAL OF STRATEGIC INFORMATION SYSTEMS, v. 22, n. 1, p. 70–85, mar. 2013.
163	NÄRMAN, P. et al. Using enterprise architecture and technology adoption models to predict application usage. Journal of Systems and Software, v. 85, n. 8, p. 1953–1967, 2012.
164	HOLSCHKE, O. et al. Using Enterprise Architecture Models and Bayesian Belief Networks for Failure Impact Analysis. (W. Feuerlicht, G and Lamersdorf, Ed.)SERVICE-ORIENTED COMPUTING - ICSOC 2008 WORKSHOPS. Anais...: Lecture Notes in Computer Science.2009
165	NÄRMAN, P. et al. Using Enterprise Architecture models for system quality analysis. Enterprise Distibuted Object Computing Conference. 12th International IEEE, n. October, p. 14–23, 2008.
166	S., S.; S., R.; V., K. Using intentional and system dynamics modeling to address whys in enterprise architecture. ICSOFT 2013 - Proceedings of the 8th International Joint Conference on Software Technologies. Anais...201

Continued on next page

Table C.1: EA Analysis SLR primary studies

Table C.1 – continued from previous page

ID-3	Reference
169	PÁSCOA, C.; BELO, N.; TRIBOLET, J. Value model for enterprise and process architectures alignment verificationCommunications in Computer and Information Science, 2010.
171	LAGERSTRÖM, R. et al. Visualizing and Measuring Enterprise Application Architecture: An Exploratory Telecom Case The Harvard community has made this article openly available . Please share how this access benefits you . Your story matters . Citation Visualizing and Measuring. p. 20, 2013.
172	LAGERSTRÖM, R. et al. Visualizing and measuring enterprise architecture: an exploratory biopharma case. IFIP Working Conference on The Practice of Enterprise Modeling. Anais...2013
173	SIMONIN, J.; NURCAN, S.; GOURMELEN, J. Weighted Alignment Measures of Enterprise Architecture Viewpoints. (J. Bajec, M and Eder, Ed.)ADVANCED INFORMATION SYSTEMS ENGINEERING WORKSHOPS, CAISE 2012. Anais...: Lecture Notes in Business Information Processing.2012
175	B., F.; Y., P. A visual approach to business IT alignment between business model and enterprise architecture. International Journal of Information System Modeling and Design, v. 6, n. 1, p. 1–23, 2015.
176	QUARTEL, D.; STEEN, M. W. A.; LANKHORST, M. M. Application and project portfolio valuation using enterprise architecture and business requirements modelling. ENTERPRISE INFORMATION SYSTEMS, v. 6, n. 2, p. 189–213, 2012.
177	MORIMOTO, S. Encouragement of Defining Moderate Semantics for Artifacts of Enterprise Architecture. (Gaol, FL and Nguyen, QV, Ed.)PROCEEDINGS OF THE 2011 2ND INTERNATIONAL CONGRESS ON COMPUTER APPLICATIONS AND COMPUTATIONAL SCIENCE, VOL 1. Anais...: Advances in Intelligent and Soft Computing.2012
178	I., G.; I., R.; S., N. On the way from research innovations to practical utility in enterprise architecture: The build-up process. International Journal of Information System Modeling and Design, v. 1, n. 3, p. 20–44, 2010.
179	BAKHSHADEH, M. et al. Ontology transformation of enterprise architecture models. Doctoral Conference on Computing, Electrical and Industrial Systems. Anais...2014LAGERSTROM, R.;
180	D.V., T. Policy mapping: Relating enterprise architecture to policy goals. [s.l: s.n.].

Continued on next page

Table C.1: EA Analysis SLR primary studies

Table C.1 – continued from previous page

ID-3	Reference
181	CANSADO, F.; VASCONCELOS, A.; SANTOS, G. Using Multi-criteria Analysis to Evaluate Enterprise Architecture Scenarios. ICEIS 2012. 14th International Conference on Enterprise Information Systems. Proceedings, p. INSTICC, 2012.
182	MATHEWS, Z.; KAESLIN, L.; RYTZ, B. Harnessing Multimodal Architectural Dependencies for IT-Business Alignment and Portfolio Optimization: A Statistical Approach. In: Digital Enterprise Design & Management. [s.l.] Springer, 2014. p. 35–46.
183	MANZUR, L. et al. xArchiMate: Enterprise Architecture simulation, experimentation and analysis. Simulation-Transactions of the Society for Modeling and Simulation International, v. 91, n. 3, p. 276–301, 2015.

Table C.1: EA Analysis SLR primary studies

D

EA ANALYSIS SLR - PRIMARY STUDIES GROUPED BY CONCERN

Category	Concern	Some evidences of the concern in the paper	Papers
Actors aspects	Human aspects in EA	“this kind of usage requires the inclusion of a human perspective in current enterprise architectures” (ID3_8, p.840)"This enables reasonable estimations of what subjects (individual or collective) do at given moments i.e. which actions they are performing and which resources they are using." (ID3_8, p.843)	[ID3_8]
	EA actor’s competence impact	There is a lack of change management methods that consider the actors’impact when making decision. Thus, we aim to elaborate a support change process that focuses on actors’impact. We notice that actors’behaviour in a changing situation is an important factor that contributes greatly to ensure alignment between the various IS’s components and so ensures the success of the change management in the enterprise. (ID3_145, p. 161)	[ID3_145]
Application Portfolio Analysis	Application Portfolio Analysis	Using this framework means collecting qualitative data on the application landscape as well as quantitative data from the application’s users so as to be able to make an analysis of why certain applications are well-liked and widely used and – more importantly – why others are not. (ID3_163, p.2)	[ID3_163]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
EA standard conformance	EA standard conformance	The EAM Pattern Catalog can e. g. be used to look for typical concerns, which occur in other companies. Additionally, the EAM Pattern Catalog can suggest visualizations that can be found in academia and practice, which may be helpful in an already selected EA management approach." (ID3_24, p.2)"The article also presented the EAM Pattern Catalog, a collection of best practices for EAM building on the pattern based approach." (ID3_24, p.9)	[ID3_24]
Cost analysis	Cost allocation	"The case concerns the problem of cost allocation, which arises because the company needs to make a decision."(ID3_23, p.24)"To establish a cost allocation per business unit, the two-stage cost allocation process is used [9], which in our case comprises two steps, namely (1) calculating product cost based on server costs and (2) reallocating product cost to multiple divisions."(ID3_23_,p.26)	[ID3_23]
	EA ROI	"... simple metrics and models for return on investment were introduced." (ID3_38, p. 11)	[ID3_38]
	IT architecture cost	This paper proposes a design methodology of information technology architectures tying organizational requirements to technical choices and costs. [ID3_83, p.1]	[ID3_83]
	IT portfolio analysis	The value of some IT project or IT artefact, such as a software application, is a combination of its benefits, costs and risks. (ID3_176, p.192)	[ID3_176]
	Cost of change of components	"The proposed metamodel for modifiability analysis, cf. Fig. 2, focuses on the software systems and the surrounding environment involved in or affected by the modifications implemented in a change project... " (ID3_62, p.11)"Centrality degree indicates the cost associated with a given application." (ID3_122, p.26)	[ID3_62] [ID3_122]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
EA Alignment	Alignment between layers	“As the rates of business and technological changes accelerate, misalignments between business and IT architectures are inevitable.This paper introduces the BITAM (Business IT Alignment Method) which is a process that describes a set of twelve steps for managing, detecting and correcting misalignment.” (ID3_11, p. 5) "We develop a reference model to be used to create an architectural description of the object system (arable farm), describing the relations between the business layer and advanced tools that (should) support these processes. We test the validity of the reference model by case studies to create detailed descriptions of bottlenecks in arable farm enterprise integration." (ID3_14, p. 2)"Our objective is to establish and maintain a tight link between the applications of the legacy system and the business models of the enterprise. This link makes them aligned and the mismatches can be revealed (as-is) and avoided in the future state of the system (to-be)." (ID3_42, p.1)	[ID3_2] [ID3_11] [ID3_12] [ID3_14] [ID3_42] [ID3_43] [ID3_134] [ID3_139] [ID3_175] [ID3_178]
	EA redundancy	Our proposal main objectives are: i) Semi-automate the process of identifying redundant entities in the context of an EA, exploiting ontology alignment techniques. (ID3_59, p.1)	[ID3_59]
	Support coverage	Landscape maps can be used both visually and non-visually. Their most general use is in finding mutual dependence of three different kinds of entities in EA" (ID3_61, p.12)	[ID3_61]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
	Application Redundancy	"Another point about the network presented in Fig. 3 is that many of these applications only provide a thin wrapper around existing functionality. For example, C5 provides an EJB interface to C10 and C7. C1 provides the BizCo.com wrapper to this functionality. C4 provides Web Service and EJB wrappers around other tracking applications (not shown here). Their justification for existence as separate applications in their own right is tenuous. Rather, they are acting as integration glue and could be targets of architectural simplification."(ID3_15, p.12)	[ID3_15]
	Overall alignment	"This paper proposes to decompose the IT/business alignment problem into tangible qualities for business, IT systems, and IT governance." (ID3_25, p.14)."An explorative study among 162 EA professionals is used to distinguish four different IT/business alignment situations. These situations each represent the current state described by different qualities and also influenced by the priorities for future development (TO-BE states)" (ID3_25, p.21)."This framework is used as a foundation for case study research to find alignment patterns used in practice. Our first results indicate that our approach might yield an operationalization of a strategic architecture alignment model" (ID3_135, p.16)	[ID3_25] [ID3_135]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
EA Change	EA change impact	“In its application to EA, adaptation represents how the enterprise and architecture in place respond to various forms of changes (see below), which make the new EA more suitable (Modifications, deletions,additions of components)” (ID3_2, p.397)“The case organisations were all in the process of reorganizing the core business, the care process, due to changes in governmental regulations, changing vision on health care and a more central role of the patient in the interaction with medical personnel. They have an IT infrastructure in operation that functions sufficiently for the current requirements. The board of directors in all three cases wanted to re-evaluate the IT systems in the light of new requirements.” (ID3_10, p.265)."Impact analysis, which determines the effects of changes or failures on other architectural elements, can therefore provide valuable information for enterprise architects." (ID3_51, p.22) "In this paper we address the problem of mastering the ripple effects of a proposed change. This allows architects to assess the consequences of a particular change to the enterprise, in order to identify potential impacts of a change before it actually takes place." (ID3_75, p.1)	[ID3_2] [ID3_10] [ID3_51] [ID3_61] [ID3_65] [ID3_75] [ID3_100] [ID3_104] [ID3_118] [ID3_122] [ID3_124] [ID3_129] [ID3_158] [ID3_182]
	EA gap analysis	We have shown how it is possible to get from identified gaps to transformation paths by creating a transformation model, detailing a target architecture and using an action repository to create possible sequences of actions for transformation paths. (ID3_112, p.15)	[ID3_112] [ID3_176]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
	Scenario analysis	"Architects and stakeholders need a precise method to compare EA scenarios[...]to know the impact of their decisions." (ID3_6, p.348)"This paper describes a method for making architecture analysis that presents not only goodness or the quality of a scenario given explicit and wide-ranging theory but also an estimate on the credibility of this analysis." (ID3_64, p.2)	[ID3_6] [ID3_37] [ID3_38] [ID3_39] [ID3_40] [ID3_44] [ID3_64] [ID3_122]
EA Decisions	Decision rationales tracking	"Our approach helps an enterprise architect (probably not the actual designer) to reconstruct the decision making process and understand how his predecessor made an EA decision" (ID3_73, p.1)"In particular, we present an integration of two viewpoints presented in earlier work: (1) an ex-ante decision making viewpoint for rationalizing EA during decision making, which for example captures a decision and its anticipated consequences, and (2) an ex-post decision making viewpoint, which for example captures the unanticipated decision consequences, and possible adjustments in criteria." (ID3_86, p.1)"In this paper we introduced a logic-based framework for capturing relationships between Enterprise Architecture decisions. We demonstrated how these constraints can be used to check a decision graph for consistency." (ID3_111, p.150)	[ID3_73] [ID3_86] [ID3_111]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
	EA decisions about of EA portfolio projects	"This paper describes the process proposed in the NENO method to guide decisional stakeholders and experts in managing a portfolio of EA projects."(ID3_21, p.298)."To determine which architectural project support is most appropriate (e.g., regular compliance reviews, ongoing participation) [1], three of our sample organizations conduct a check of architectural relevance. Among other criteria, they check the complexity of the project architecture, which our sample bank gauges by the number of interfaces of the components to be implemented and the resulting number of affected components" (ID3_122, p. 28)	[ID3_31] [ID3_122] [ID3_134]
	Prioritization of EA project initiatives	'In this paper, multifactor systems are used to provide a practical method for assessment of any given organization and making accurate decisions on improvement or redesign of its architecture based on missions, goals and restrictions of the organization. . . [...] possibility of making decision about planning and scheduling of activities for development of enterprises more accurately" (ID3_4, p. 70)	[ID3_4] [ID3_131]
	Stakeholder analysis	"is to allow for right-sized stakeholder involvement, promote effective use of resources, and increase the probability of overall program success with the assurance of lasting stakeholder commitment." (ID3_p1)	[ID3_3]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
EA Governance	EA quality	" To better understand the above definition, we present detailed definition of EA maintainability in different layers of EA: Business layer: EA maintainability in this layer is the ease with which strategic components and business functions and processes can be modified to correct faults, improve performance or other attributes, or adapt to a changed environment"(ID3_18, p.135)."The purpose of this paper is to present an enterprise architecture metamodel for maintainability analysis."' (ID3_161, p.1)	[ID3_18] [ID3_40] [ID3_161]
	EA data quality	"For Data governance purpose, we design an indicator [5] which could be calculated periodically and support a review by management. It is made of architecture maturity indicators which are viewed as levers for management and operations indicators(ID3_56,p.9)"This article uses the PRM formalism and the work presented in Närman et al. (2009) to describe a method for analysing data accuracy changes in business processes." (ID3_80, p. 2)"This paper describes how the Bayesian approach can be used for architectural analyses of data quality. The paper focus primarily on the accuracy of data and how it deteriorates in a business process involving multiple automated and manual processing steps." (ID3_89, p.1)	[ID3_56] [ID3_80] [ID3_89]
	EA documentation	starting with the documentation of the current state, two of our reference practices thus check the extent to which the enterprise architect needs to be involved in that endeavor.(ID3_122, p.23)	[ID3_122]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
	EA effectiveness	"Our approach assesses an organization on its architecture effectiveness from three essential perspectives: (1) architecture awareness, (2) architecture alignment, and (3) architecture maturity."(ID3_105, p. 1)	(ID3_105)
	Metrics monitoring	In this paper, we first propose a model-based approach to capture the temporal aspects of EA metrics and then extend a domain specific language to compute EA metrics at any point of time in the past. This allows visualizing the evolution of EA metrics and as a consequence the evolution of the EA (ID3_153., p.1)	[ID3_153]
Model consistency	Information dependence of an application	"We defined a criterion Business Information Object -BIO- criticality and a metric to determine appropriate values. We require information on which BIO is managed by which applications, which application has the data sovereignty over which BIO and which BIOs are being transferred via defined interfaces and data transfers."(ID3_22,p. 598)	[ID3_22]
	Model Correctness checking during EA evolution	"In this paper, we discuss the design of EA transformational change. Our goal is to define an approach supporting design decision during EA evolution. [...] we define an approach using Event-ConditionAction (ECA) rules, as a means of implementing a reactive mechanism to alert the enterprise architects about a suspicious evolution."(ID3_36,p .119)	[ID3_36] [ID3_37]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
	EA model consistency	"Then, we present a tool-independent strategy to support the requirements, which can be adopted in any enterprise modeling tool. This strategy is based on two central ideas: firstly, that it should be possible to easily extend metamodels with the specific information that selected analysis methods requires; and secondly, that avoiding over-complication of the modeling process requires the relaxation of conformance requirements." [ID3_27, p2)"The resulting ontological representation is therefore computable, allowing for the analysis of the consistency and completeness of the enterprise architecture models. " (ID3_179, p. 55)	[ID3_27] [ID3_124] [ID3_177] [ID3_179]
Performance	Application reliability	"C1 is dependent upon its immediate interfaces to 5 applications. However the nature of the interfaces is such that C1's ability to provide correct service is dependent upon the interface dependencies of those systems. Therefore, C1's reliability (its ability to offer correct and timely service) is dependent on not only the 5 applications listed above, but also on C6, C7, C8, C9 and C10." (ID3_15, p.3)	[ID3_15]
	Business performance	"Since the scope of this paper is on analysis of the performance of maintenance processes the business layer of Archimate´s metamodel will be used." (ID3_57, p.3180)"This paper has presented the Modelling and Analysis Framework for Organizational Performance, supporting organizational performance analysis. The benefit of this metamodel is that it contains all information relevant to make analysis of maintenance performance, resulting in a quantitative performance score as well as a model that is easy to understand." (ID3_98,p.4)	[ID3_57] [ID3_98]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
	EA component performance	"we propose a module to observe simulation experiments and gather information about the responses of the architecture model to certain stimuli. This information is then processed and presented as performance indicators." (ID3_70, p.1)"In summary, this paper aims to design and evaluate and an analytical method for modeling and estimating service response time based on enterprise architecture " (ID3_162, p.2)	[ID3_70] [ID3_127] [ID3_136] [ID3_162]
	System quality	"This paper has presented the PERDAF enterprise architecture metamodel supporting enterprise system quality analysis. The metamodel consists of classes with accompanying attributes that can be used to create enterprise architecture models from which it is possible to extract precisely the information that is needed for quantitative system quality analysis"(ID3_99, p.10)"This paper describes how this Bayesian analysis approach is used for estimating availability, accuracy, confidentiality and integrity of two SOA platforms" (ID3_165, p.1)	[ID3_99] [ID3_155] [ID3_165]
Risk	Risk analysis	risks can be indicated by CD, given the number of related applications it represents (ID3_122, p. 27)	[ID3_122]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
	Risk failure	In the context of enterprise architectures, failure impact analysis can be used to simulate the effect of the non-availability of an architectural element. Examples for failures are the shutdown of a server or inaccessible applications. Potential effects could be a defunct business process or a non-available product."[ID3_51, p.32]. "Based on this research, we present a quantification of the failure impact analysis using availability measures."[ID3_51, p.34]Obviously, when applications occupy a central position in relation to their context of related applications, they can influence their environment (maybe even the network as a whole) negatively (ID3_122, p.24)"We propose a Decision Support System (DSS) for failure impact analysis for Enterprise Architectures based precisely on the Bayesian Belief Networks (BBN)" (ID3_164, p.343)	[ID3_15] [ID3_51] [ID3_122] [ID3_164]
	Information security risk	"Our proposal aims at using it[EA] in conjunction with concepts of information security risks analysis." (ID3_19, p.115)"In this case study the cyber security of an electric substation was the concern." (ID3_47, p.3)"the approach makes it possible to calculate the probability that attacks succeed based on an enterprise architecture model." (ID3_76, p. 1)	[ID3_19] [ID3_47] [ID3_76]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
	Availability	"This paper describes an EID for availability assessment, how it was used together with associated metamodels for analysis of a set of system properties according to the ISO 9126 standard [6] [7]." (ID3_66, p.2)"The application of Bayesian networks for information system quality analysis is proposed and applied in [12].[...]. The present paper is similar in method, but focuses on availability rather than modifiability" (ID3_67, p. 2)"This article presents an analysis framework integrating EA modelling with fault tree analysis (FTA) of availability (Stamatelatos et al. 2002)." (ID3_93, p.2)	[ID3_21] [ID3_26] [ID3_34] [ID3_39] [ID3_48] [ID3_66] [ID3_67] [ID3_93]
	EA implementation risk	Enterprise architecture is a strategic initiative that integrates strategy, structure, processes, and applications. But there are some potential risks that threat effective deployment of this initiative (ID3_115, p.9)	[ID3_115]
	Integration risk	This paper presents an approach for calculating risk of integration relations between applications in an application landscape (ID3_71, p.1)	[ID3_71]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
Strategy compliance	Compliance of EA decisions with requirements, principles, directives and capabilities	"...we can represent the extent to which EA design decisions, which define the EA design, comply with given goals, principles and requirements" (ID3_1, p.1)" We present a model-based solution that enables a) modeling directives at various levels of detail on top of extended enterprise architecture-based models of enterprise, b) analyzing the models for compliance, and c) ensuring operationalization of directives " (ID3_30, p. 57)"Our extension is intended to model the organization's core capabilities and key resources with a strategic management focus. With our extension, the enterprise can consider the required capabilities and resources to achieve a desired state without actually having to pursue a complete and extended view on the business processes and tasks that are necessary to realize that state. This means that at the strategic management level enterprise architects are able to focus on the proper level of abstraction, avoiding unnecessary commitments with lower level details.""(ID3_128, p. 25)"	[ID3_1] [ID3_144] [ID3_30] [ID3_128]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
	Goal compliance	“This introduces a new kind of analysis to enterprise architecture that allows for investigating the degree of goal achievement of a set of competing EA configurations in light of stakeholders’ goals.” (ID3_106, p.6)"it will be extremely useful to the organization to have an understanding of the extent to which the SA satisfies the EA – this traceability will ensure that the chosen system architecture meets the goals of the enterprise architecture as well." (ID3_58, p.1)"Our method to design a set of metrics to evaluate SOA focuses on the different goals SOA may pursue and takes the individual situation of the respective organization into account. " (ID3_85, p.3)“EA elements can be traced back to the goals and requirements that motivated their introduction. Reversely, RE elements can be traced forward to the services, processes and applications that implement these elements. This traceability is needed to successfully analyse and manage the impact of changes to an enterprise.(ID3_109, p. 16)	[ID3_32] [ID3_58] [ID3_82] [ID3_85] [ID3_106] [ID3_109] [ID3_116] [ID3_169] [ID3_180]
	Compliance of EA principles, by EA models	“We aim to create a formal framework for measuring and managing this impact manifested by the EA principles on the EA models.” [ID3_31, p. 37]”"Hereupon, our vision pretends to formalize architecture principles, based on ArchiMate to enable their EA compliance analysis. This formalization enables to analyse an enterprise architecture description (EAD) through the detection of architecture structures that represent the principle expected impact, and consequently identify their compliant elements." (ID3_63, p.1)	[ID3_31] [ID3_61] [ID3_63] [ID3_147]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
Structural aspects	EA domain analysis	With a documented state of the overall landscape, one can also check whether the clusters that may have emerged in the actual network of applications are consistent with the pre-defined structure of domains, which are most often derived based on means of functional decomposition (ID3_122, p.23)	[ID3_122] [ID3_126]
	EA complexity	Candidate variables in the Applications Architectural View are: A1: Number of Applications that manage data; A2: Number of Applications that support operational business systems; A3: Number of Applications that support strategic business systems; A4: Number of APIs to common services; A5: Number of lines of CUSTOM code (in thousands); A6: Number of lines of COTS code (in thousands) (ID3_33,p.9)"Against this background the contribution of this paper is a) the introduction of a system theoretic conceptualization of complexity in enterprise architectures and b) a measure to quantify complexity in enterprise architectures. We define complexity as the number and the heterogeneity of the components and relations of an enterprise architecture" (ID3_78, p.1)"This research provides a conceptual framework to measure IT organization and business-IT responsibility sharing (IT O&R) as well as IT architecture complexity for large enterprises." (ID3_157, p.11)	[ID3_33] [ID3_78] [ID3_90] [ID3_134] [ID3_153] [ID3_157]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
	Positional value	“In our understanding an architecture can be defined as an abstract and holistic concept of structures and patterns considering planning aspects.” (ID3_28, p.1)"betweenness is an adequate evaluation metric given that applications can be considered especially valuable if they occupy a central position within one (or several) business process(es) and may, for example, somewhat control the data exchange therein." (ID3_122, p.26)" With the use of the BPMN representation, the OV-5 models enabled the evaluation of information architectures and infrastructure use in each activity for identification of potential weaknesses, deficiencies, or inefficiencies, providing a method of validation for the identified shortfalls. In addition to validation of identified shortfalls, the structure and detail of the models provided other avenues of analysis, such as identification of additional shortfalls, examination of stakeholder data needs, and identification of data and system consolidation opportunities" (ID3_159, p. 5)"Intentional analysis of vulnerability (Yu, 2009) may be performed by classifying dependencies as open (failure does not affect depender), committed (depender is significantly affected), and critical (depender’s goals may fail) when dependee does not deliver. Critical dependencies may result in severe vulnerability on depender’s part. We begin to investigate vulnerabilities of these actors with intentional model of the problem" (ID3_166, p.5)	[ID3_28] [ID3_110] [ID3_122] [ID3_159] [ID3_143] [ID3_166] [ID3_171] [ID3_172]

Continued on next page

Table D.1: EA Analysis SLR Primary Studies classified by concern

Table D.1 – continued from previous page

Category	Concern	Some evidences of the concern in the paper	Papers
	Impact of Shadow IT systems	We were surprised by the extent to which shadow IT systems embed in Recycle Inc.'s IS architecture. Among the twenty most central IT systems, we identified three shadow systems with high degree centrality and four systems with high betweenness. (ID3_142, p.8)	[ID3_142]
Traceability	Traceability	This article looks at EA validation from a completely different perspective and defines it as the use of objective evidence to confirm that a model meets the intrinsic goals defined by their requirements by querying the artifacts in a resource description framework (ID3_168, p.2)	[ID3_146]

Table D.1: EA Analysis SLR Primary Studies classified by concern

E

EA ANALYSIS SLR - PRIMARY STUDIES GROUPED BY ANALYSIS APPROACH

Category	Specific approach	Concerns analyzed	Papers
Agent simulation based	Agent simulation based	Change impact	[ID3_158]
Analytic Hierarchy Process (AHP)	AHP	Scenario analysis	[ID3_40]
Architecture Theory Diagrams (ATDs) based	ATDs for modifiability	Impact analysis	[ID3_110]
Axiomatic Design	Axiomatic Design	Business-IT ALignment	[ID3_10]
Best practice conformance	EA analysis pattern catalog	Best practices analysis	[ID3_24]
BI	Total cost of server [ID3_23]	EA Alignment, EA Change, Cost allocation	[ID3_02] [ID3_23]
BITAM	BITAM	Business-IT alignment	[ID3_11]
Compliance analysis	Analysis of the contribution of EA components to goal achievement; Goal evaluation algorithm, Expert reasoning about influence links and compliance to the principles	Alignment of changes with EA capabilities, Goal compliance, Compliance of EA decisions with Goals, requirements, principles, directives and capabilities	[ID3_31] [ID3_106] [ID3_128] [ID3_116]
Cost benefit analysis	Cost benefit analysis	Prioritization of projects for EA change	[ID3_04]

Continued on next page

Table E.1: EA Analysis SLR primary studies classified by analysis approach

Table E.1 – continued from previous page

Category	**Specific approach**	**Concerns analyzed**	**Papers**
Design Structured Matrix	Design Structured Matrix	Human aspects, Business-IT Alignment, EA Complexity	[ID3_08] [ID3_10] [ID3_171] [ID3_172]
EA Anamnesis	EA Anamnesis with ex-ante and ex-post analysis	Decision rationales tracking	[ID3_01] [ID3_73] [ID3_86]
EA executable models/Simulation	EA executable models/Simulation	EA component performance	[ID3_183]
FCM (Functional Cognitive Maps) with OWA (Ordered Weighted Averaging) operators	FCM (Functional Cognitive Maps) with OWA (Ordered Weighted Averaging) operators	Scenario analysis	[ID3_44]
	Event-ConditionAction (ECA) rules	Model Correcteness checking during EA evolution	[ID3_36]
	AgreementMaker - algorithm of ontology comparison and match	Data redundancies	[ID3_59]
	Indicator basd on domain specific language	Business performance	[ID3_70]

Continued on next page

Table E.1: EA Analysis SLR primary studies classified by analysis approach

Table E.1 – continued from previous page

Category	Specific approach	Concerns analyzed	Papers
	Queries based on Model-based Expression Language (DSL) including timeseries analysis; Queries with formal language; ontologies matching for consistency analysis; Queries with OWL-DL; Model checking with Alloy language and rules; Formal Reasoning with OWL; Constraints rules for Decision Design Graphs; Model checking	EA complexity, EA alignment, Traceability, EA change impact, EA gap analysis, Decision rationales tracking	[ID3_43] [ID3_146] [ID3_153] [ID3_139] [ID3_130] [ID3_124] [ID3_112] [ID3_111]
	Propagation rules for change in ontology models	Change impact	[ID3_104]
Formalism -based	Description Logic Checking	EA model consistency; Traceability; Change Impact, Compliance of EA decisions with Goals, requirements, principles, directives and capabilities; Model checking during EA evolution	[ID3_179] [ID3_61] [ID3_26] [ID3_27] [ID3_26] [ID3_30] [ID3_37]
	Traceability based on set theory	EA alignemnt	[ID3_177]
Functional-Business Object Matrix (FBOM) e ROCK algorithm	Functional-Business Object Matrix (FBOM) e ROCK algorithm	System quality	[ID3_155]
Fuzzy based	Data envelopment Group analysis based on fuzzy credibility constrained programming and p-robustness	Scenario analysis	[ID3_06]
	Failure mode and effect analysis (FMEA) and VIKOR	EA implementation Risks	[ID3_117]

Continued on next page

Table E.1: EA Analysis SLR primary studies classified by analysis approach

Table E.1 – continued from previous page

Category	Specific approach	Concerns analyzed	Papers
Mathematical functions	Linear regression to estimate actor performance based on EA changes	EA change; EA actor´s competence impact	[ID3_145]
Metrics based	A set of metrics for measure IT fitness (alignment)	EA alignment	[ID3_178]
	Production function	Availability	[ID3_21]
	IT importance, IT cost and IT effectviness analysis	EA cost; Application portfolio valuation; EA gap analysis	[ID3_176]
	Return of investment metric	Return of Investment	[ID3_38]
	the workload, the processing time, the utilisation	EA component performance	[ID3_136]
	Entropy measure	EA Complexity	[ID3_78]
	Several mathematical formulas for EA alignment between layers	EA alignment between layers	[ID3_131]
	Several mathematical formulas for EA project prioritization	Prioritization of EA project initiatives	[ID3_131]
	Complexity metrics	Complexity	[ID3_157]
	Time, Resource use anaysis	EA Component performance	[ID3_127]
	EA Viewpoint Alignment Measure	Overall alignment	[ID3_173]
	Business information object (BIO) criticality	Information depedence of an application	[ID3_22]
	Effectiveness measure	EA effectiveness	[ID3_105]
	Metric rules with OCL through models composition	EA Alignment	[ID3_42]
	Enterprise quality data indicator	EA Quality	[ID3_56]
	Overall Risk Score:	Integraton risk (risk)	[ID3_71]
	total cost of server (analysis based onon BI paradigm)	Cost allocation	[ID3_23]

Continued on next page

Table E.1: EA Analysis SLR primary studies classified by analysis approach

Table E.1 – continued from previous page

Category	Specific approach	Concerns analyzed	Papers
Multi-criteria analysis	Multiple attribute decision analysis (MADA)Cost analysis based on technology layer attributesPropagatory framework	scenario analysisCost analysisGoal conformanceTraceability	[ID3_32][ID3_181][ID3
Prescriptive models	Prescriptive model for EA Alignment evolution	Overall alignment	[ID3_25]
Probabilistic -based	Bayesian Networks	Data quality (accuracy)EA quality (maintanability)System quality	[ID3_89][ID3_172][ID3
	EID + GQMEID rules	AvailabilityRisk analysis-Goal complianceBusiness performanceSystem qualityEA change impactScenario analysis Information security risks	[ID3_57][ID3_66][ID3_
	PRM	Modifiability analysis (Change cost)AvailabilityData accuracyEA Component PerformanceApplication Portifolio AnalysisCost of change of componentsInformation security risk	[ID3_47][ID3_62][ID3_
	Fault Tree Analysis	Dependency Analysis	[ID3_90]
	The Predictive, Probabilistic Architecture Modeling Framework (P2AMF)	Structural analysis - Modifiability analysis (Change cost)	[ID3_65]
	leaky Noisy-OR model (Bayesian Network based)	Availability	[ID3_67]
	Bayesian Belief Networks	Risk failure	[ID3_164]
	Probabilistic rules for impact analysis	Change Impact;Risk failure	[ID3_51]

Continued on next page

Table E.1: EA Analysis SLR primary studies classified by analysis approach

Table E.1 – continued from previous page

Category	Specific approach	Concerns analyzed	Papers
Structural analysis	Stakeholder Crosswalk; Clustering analysis; Structural complexity metric; Network analysis(degree in and out); Quality of clusters maps (such decoupling, cohesion, number of clusters); Clustering algorithms + Maturity and strategic weight; Clustering algorithms + Visual analysis; Social Network analysis; Network analysis (eigenvector and degree centrality); Network analysis (degree centrality); Network analysis (out-degree and eigenvector); Network analysis (closeness and clustering, average degree centrality); Clustering techniques (Girvan/Neumann algorithm); Markov-Chain-MonteCarlo Gibbs	Stakeholder analysis; Business-IT alignment; EA complexity; EA decisions about EA portfolio projects; EA alignment; Impact analysis; Change impact; EA domain analysis; EA scenario analysis; Risk analysis; EA cost of change; Risk failure; Resource allocation for EA documentation	[ID3_03] [ID3_12] [ID3_28] [ID3_33] [ID3_134] [ID3_142] [ID3_143] [ID3_100] [ID3_126] [ID3_122] [ID3_182]

Continued on next page

Table E.1: EA Analysis SLR primary studies classified by analysis approach

Table E.1 – continued from previous page

Category	Specific approach	Concerns analyzed	Papers
Visual analysis	Visual analysis + interviews; Clustering algorithms + Visual analysis; Visual analysis + ordering algorithms; Visual analysis + Business Process analysis patterns; Visual analysis + qualitative questions; Graph Visual dependency analysis; Traceability analysis of goals changes and its impact on low level components (process, applications etc); Visual Semantic analysis	EA goal compliance; Application reliability, redundancy and risk failure; Risk; Alignment between layers;Strategy compliance; Impact analysis; EA Complexity; Structural aspects; Overall alignment; Positional value; EA change impact; Stakeholder analysis	[ID3_14] [ID3_180] [ID3_15] [ID3_19] [ID3_63] [ID3_175] [ID3_135] [ID3_85] [ID3_133] [ID3_134] [ID3_94] [ID3_96] [ID3_159] [ID3_125] [ID3_166] [ID3_169] [ID3_118] [ID3_109]

Table E.1: EA Analysis SLR primary studies classified by analysis approach

F

EA ANALYSIS SLR - PRIMARY STUDIES GROUPED BY MODELS

Model-Category	Model	Papers
Archimate-based	Archimate	[ID3_24] [ID3_75] [ID3_86] [ID3_96] [ID3_125] [ID3_181]
	Own (Archimate-extended)	[ID3_19] [ID3_23] [ID3_27] [ID3_56] [ID3_63] [ID3_109] [ID3_128] [ID3_176] [ID3_183]
	Own (Archimate-based)	[ID3_01] [ID3_44] [ID3_51] [ID3_71] [ID3_135] [ID3_136] [ID3_147] [ID3_165] [ID3_179]
Probabilistic networks-based	Bayesian Belief Networks	[ID3_164]
	EID	[ID3_82] [ID3_108] [ID3_129]
	PRM	[ID3_62]
	PRM based	[ID3_47]
	P2AMF	[ID3_65]
Combined models	Combined models	[ID3_26] [ID3_30] [ID3_39] [ID3_40] [ID3_42] [ID3_48] [ID3_57] [ID3_66] [ID3_76] [ID3_80] [ID3_89] [ID3_90] [ID3_93] [ID3_139] [ID3_159] [ID3_163] [ID3_164] [ID3_175] [ID3_178]
Graphs based	DSM-based model	[ID3_169] [ID3_134] [ID3_143]

Continued on next page

Table F.1: EA Analysis SLR primary studies classified by models

Table F.1 – continued from previous page

Model-Category	**Model**	**Papers**
	Graphs	[ID3_03] [ID3_15] [ID3_28] [ID3_58] [ID3_94] [ID3_100] [ID3_111] [ID3_116] [ID3_118] [ID3_122] [ID3_126] [ID3_132] [ID3_142] [ID3_166] [ID3_171] [ID3_172] [ID3_180] [ID3_182]
Intentional Modeling	GRL	[ID3_31]
	Intentional Modeling	[ID3_106]
	User Requirements Notation (URN)	[ID3_02]
Formal specification -based	Description language based	[ID3_124] [ID3_153]
	Ontology	[ID3_59] [ID3_61] [ID3_112] [ID3_130] [ID3_139] [ID3_146] [ID3_179]
	Set theory	[ID3_177]
Own	Own	[ID3_08] [ID3_10] [ID3_11] [ID3_12] [ID3_14] [ID3_22] [ID3_25] [ID3_32] [ID3_36] [ID3_37] [ID3_43] [ID3_65] [ID3_70] [ID3_85] [ID3_98] [ID3_99] [ID3_104] [ID3_131] [ID3_155] [ID3_158] [ID3_161] [ID3_162] [ID3_173]
UML based	UML based	[ID3_127]

Table F.1: EA Analysis SLR primary studies classified by models

G

EANA NAIs

We describe here a subset of the data extracted regarding the 74 NAIs identified in our EANA SLR described in Chapter 5. An even more detailed version of the dataset can be found online at www.eanaresearch.org.

NAI_ ID	Name of NAI	Analysis concern	Concern question	EA layer(s) targeted	Network measure/ method	Network level of analysis	Data source	Primary study
1	*Degree of informational dependency*	Information dependence of an application	*What are the most dependent applications in terms of information?*	*Application, Information*	*M_01*	Component level	Primary data	1S_ID_0047
2	*Degree of functional dependence*	Information dependence of an application	*What are the applications which depend on information from different enterprise domains?*	*Application, Business*	*M_02*	*Component level*	*No data*	*1S_ID_0047*
3	*Degree of heterogeneity of the communication paths in the application layer*	Reuse of an application	*How is the degree of reuse of the application?*	*Application, Information*	*M_03*	*Component level*	*No data*	1S_ID_0047
4	*Business Information criticality of an application*	Business criticality of an application in terms of information	*How critical is an application in terms of information consume and production?*	*Application, Information*	*M_04*	*Component level*	*No data*	*1S_ID_0094*
5	*Static alignment measure*	EA application-information alignment	*What is the static alignment between the information layer (Business Object) and the application layer?*	*Application, Information*	*M_05*	*component level*	*Primary data*	*1S_ID_0029*
6	*Centrality degree, closeness, betweenness*	*Key structural components*	*What are the key structural components of the EA?*	*Application*	*M_06*	*component level*	*Primary data*	*1S_ID_0062*
7	Whitney Index*	EA Modularity	How well the application/system is modularized?	Application	M_07	Network	Primary data	1S_ID_0062
8	*Change Cost*	EA Change	*How to measure the propagation effect of a change in EA?*	*Application*	*M_08*	*Network*	*Primary data*	*1S_ID_0062*

9	*Stakeholder Crosswalk (clustering algorithms, network measures not described)*	*Stakeholder management*	*How to identify stakeholder groups according to EA Domains or along business processes phases?*	*Value, Process*	*M_09*	*Network/Modules/Component Level*	*Primary data*	*1S_ID_0005*
10	*degree of connectedness*	*EA Complexity*	*How to evaluate the network behavior of the EA in terms of change over time?*	*Application*	*M_10*	*Network level*	*No data*	*1S_ID_0072*
11	Time and intensity between connections	EA complexity	How intensive the application interactions are?	Application	M_11	component level	No data	1S_ID_0072
12	*List of Complex Adaptive Systems features*	*EA complexity*	*How evaluate EA complexity?*	*Application*	*M_12*	*Network level*	*No data*	*1S_ID_0072*
13	Coupling(operationalized by closeness centrality)	EA cost of change	How to measure the flexibility of the application layer based on the cost of change of its components?	Application	M_13	component level	Primary data	1S_ID_0079
14	Application Clustering based on CSWPF (Connectivity Strength with Weight Parameter and Frequency (CSWPF))	EA Domain Analysis	How to cluster application base on amount and frequency of data transferred?	Application	M_14	component level	Primary data	1S_ID_0081
15	STA(Service time actual)	EA improvement	What are the components with high complexity in terms of time execution?	Business, Information, Application, Technology	M_15	*Component level*	Simulated data	1S_ID_0092
16	Degree centrality (CD)	Application usage	What is the level of connectivity of the application? / What are the applications that are more costly in terms of change?	Application	M_16	Component level	Primary data	1S_ID_0130
17	*Closeness centrality (CC)*	EA redesign/ Cross functional applications	*How to identify applications which belong to several EA domains?*	*Application*	*M_17*	*Component level*	Primary data	*1S_ID_0130*
18	Betweenness centrality (CB)	C_05/ Risk failure management	What are the key structural applications that might have to be monitored to minimize failure impact?	Application	M_18	Component level	Primary data	1S_ID_0130
19	*Eigenvector*	Risk failure management	*What are the most well connected components in the EA?*	*Application*	*M_19*	*Component level*	Primary data	*1S_ID_0130*
20	*modularity*	EA Modularity	*How to identify architectural domains in terms of structure? How structural domains are aligned with functional domains?*	*Application*	*M_20*	*Group level*	No data	*1S_ID_0130*
21	*Architectural relevance checking*	Architectural relevance check for implementation	*What are the components which can generate more impact after implementation of changes?*	*Application*	*M_21*	*Component level*	No data	*1S_ID_0130*

22	*AIS[s], Absolute Importance of the Service*	*Local dependence*	*What the importance of an application in terms of the number of applications which depend on it?*	*Application*	*M_22*	*Component level*	No data	*IS_ID_0151*
23	*ADS[s], Absolute Dependence of the Service*	*Local dependence*	*What is the absolute dependence of an application in terms of the number of other applications this application depends on?*	*Application*	*M_23*	*Component level*	No data	*IS_ID_0151*
24	ACS[s], Absolute Criticality of the Service	*Application monitoring / application criticality*	How to evaluate the component criticality in terms of its dependencies (in and out)?	Application	M_24	Component level	No data	*IS_ID_0151*
25	*Change cost*	*EA change*	*What are the applications that impact most on the network in case of changes?*	*Application*	*M_25*	*Component level*	*No data*	*IS_ID_0151*
26	*McCabe's Cyclomatic Number*	EA complexity	*How to detect weak spots in the business process layer?*	*Business*	*M_26*	*Network Level*	*simulated data*	*IS_ID_0172*
27	*Activity*	EA improvement	*What the importance the process in terms of the number of process that depend on it?*	*Business*	*M_27*	*Component level*	*simulated data*	*IS_ID_0172*
28	*Passivity*	EA improvement	*How much the process is impacted is the changes in other processes?*	*Business*	*M_28*	*Component level*	*simulated data*	
29	*IT landscape size*	*Business-IT alignment*	*What is the IT landscape coverage in terms of the amount of business process supported?*	*Business, Application*	*M_29*	*network level*	*No data*	*IS_ID_0200*
30	*Number of interfaces between applications*	*EA complexity*	*What is the number of interconnections in the application layer?*	*Application*	*M_30*	*network level*	*No data*	*IS_ID_0200*
31	*Purity of the Domain*	EA improvement	*What is the purity of the EA domains?*	*Application*	*M_31*	*network level*	*Primary data*	*2S_01_ID_0009*
32	*Business support*	IT-Business alignment	*What is the importance of a given application component for the business layer?*	*Business, Application*	*M_32*	*component Level*	*Primary data*	*2S_01_ID_0009*
33	*Discriminant function for Core process*	*Process improvement or optimization/Business process analysis*	*How to identify core business processes in terms of information flow in order to facilitate Business Process Management?*	*Business*	*M_33*	*Network Level*	*Primary data*	*FS_02*
34	*Discriminant function for automatable process*	*Process improvement or optimization/Business process analysis*	*How to identify potential automatable business processes considering the information flow in order to facilitate Business Process Management?*	*Business*	*M_34*	*Network Level*	*Primary data*	*FS_02*
35	*Discriminant function for information intensive process*	*Process improvement or optimization/Business process analysis*	*How to measure the information intensiveness of business processes considering the information flow and structural position?*	*Business*	*M_35*	*Network Level*	*Primary data*	*FS_02*

36	*Discriminant function for distributed process*	*Process improvement or optimization/Business process analysis*	*How to identify potential distributed business processes considering the information flow and structural position in order to facilitate Business Process Management?*	*Business*	*M_36*	*Network Level*	*Primary data*	*FS_02*
37	*Discriminant function for flexible process*	*Process improvement or optimization/Business process analysis*	*How to measure the flexibility of business processes considering the information flow and structural position in order to facilitate Business Process Management?*	*Business*	*M_37*	*Network Level*	*Primary data*	*FS_02*
38	*Number of Applications*	*EA Complexity*	*What is the complexity of the EA domain in terms of number of applications?*	*Application, Value*	*M_29*	*Network level, Group level*	*primary data*	*FS_04*
39	*Number of business functions*	*Business Complexity*	*What is the complexity of the business layer in terms of the number of business functions?*	*Business*	*M_38*	*Network level, Group level*	*primary data*	*FS_04*
40	*Number of Information Flows*	*Application Complexity*	*What is the complexity of an application in terms of the number applications interfaces?*	*Application*	*M_39*	*Component Level, group level*	*primary data*	*FS_04*
41	*Number of Infrastructure Elements*	Infrastructure Complexity	*What is the complexity of a component, domain or entire EA considering the number of technology components accessed by them?*	*Value, Application, Infrastructure*	*M_40*	*Network level, Component Level, Group level*	*primary data*	*FS_04*
42	*Functional Scope*	*Application Landscape Complexity*	*What is the complexity of an application considering the number of business function supported by it?*	*Application, Business*	*M_41*	*Component Level*	*primary data*	*FS_04*
43	Functional Redundancy	Application Landscape Complexity	How evaluate EA application complexity based on the number of redundant applications? / What are the redundant applications?	Application	M_42	Component Level	primary data	FS_04
44	Coupled domain complexity	Application landscape Complexity	How to evaluate the EA domain complexity of a domain D based on the number of external connections of applications from D to other domains?	Application	M_43	Group level	primary data	FS_04
45	Hidden structure method	EA complexity	How to classify the application architecture in groups such as "core", "periphery"?	Application	M_44	Network level	primary data	FS_04

46	Guo and Gershenson metric - MG&G	EA modularity	How well the application/system architecture is modularized?	Application	M_45	Group level	primary data	1S_ID_0132
47	*Guo and Gershenson metric - MG&G variation (inter)*	EA modularity	*How well the application/system architecture is modularized?*	*Application*	*M_46*	Group level	primary data	*1S_ID_0132*
48	Preservation of influence/Architectural control (computed by KeyPlayer metric KPP-POS AS_IS and TO_BE)	EA evolution	How to find architectural control points and measure their preservation during the EA evolution?	Application	M_47	Network Level	*primary data*	*1S_ID_0135*
49	*Network topology preservation (computed by Small World Quotient (SWQ) 'delta)*	*EA evolution*	*How to measure the preservation of the application network topology over the time?*	*Application*	*M_48*	*Network Level*	*primary data*	*1S_ID_0135*
50	*Degree Calculator*	*EA Business complexity, EA application complexity*	*How evaluate EA business complexity?* *How to evaluate EA application complexity based on the number of in and out connections?*	*Business , Application*	*M_49*	*Component level*	simulated data	*1S_ID_0155*
51	Impact Analysis	*EA local dependence*	What is the local neighborhood of dependence of an EA component?	Business , Application	M_50	Network Level	simulated data	1S_ID_0155
52	*Interdependency*	*EA cost of change*	*How can we identify the application candidates with the highest cost of change?*	*Application, Business*	*M_51*	*component level*	*primary data*	*1S_ID_0218*
53	Process Page Rank Modified Method (PPRM method)	*Process improvement or optimization/Business process analysis*	How to establish a prioritization order for process improvement based on structural analysis?	Business	M_52	Component level	simulated data	FS_01
54	EA evolution network model based on path dependence	EA evolution	How can we monitor the attachment behavior of new applications to existent applications?	Application	M_53	*network level*	primary data	*FS_03*
55	Weighted degree	Key structural components	What are the most critical applications considering qualitative aspects of the integration among applications?	Application	M_54	*component level*	*primary data*	FS_05
56	Weighted betweenness centrality	Key structural components	What are the most critical applications considering qualitative aspects of the integration among applications?	Application	M_55	*component level*	*primary data*	FS_05
57	*Weighted Closeness centrality (CC)*	Key structural components	*What are the most critical applications considering qualitative aspects of the integration among applications?*	*Application*	*M_56*	*component level*	*primary data*	*FS_05*
58	Consistency analysis of EA domains	*EA alignment*	How are aligned stakeholder, application , information and process dependencies?	Value, Business, Information, Application	M_57	*group level*	*no data*	*EBS_01*

59	Gap analysis between Business and IT layers	*EA business it alignment*	Are the applications supporting business process?	Busines, Application	M_58	*group level*	*no data*	*EBS_01*
60	analysis of 'maturity' and 'strategic weight' applied to the process clusters	*EA Change management decision making*	How to generate an order of prioritization for EA changes implementation based on process clustering?	Value, Business, Information, Application	M_59	*Group level*	*primary data*	*EBS_01*
61	Analysis of the quality of the cluster maps over time	*EA evolution*	How EA domain structural indicators are evolving over time?	Value, Business, Information, Application	M_60	group level	no data	EBS_01
62	Analysis of transformation scenarios	Analysis of change impact in scenarios	What is the extension of the impact for a particular change in EA?	Business, Information, Application	M_61	network	no data	EBS_01
63	Business-Application Alignment	Business-IT alignment	Is each application component supporting at least one business component?	Application, Business	M_41	component level	no data	1S_ID_0048
64	betweenness centrality	Risk Failure management	Which applications can impact the IT landscape in case of failure?	Value, Business, Application	M_06	component level	primary data	1S_ID_0087
65	degree centrality	Risk Failure management	Which applications can impact the IT landscape in case of failure?	Value, Business, Application, Technology	M_06	component level	primary data	1S_ID_0087
66	time series analysis of the role of stakeholder	EA stakeholder management	How is stakeholder behavior over the time according to betweenness centrality?	Value, Business, Application< Technology	M_62	component level	primary data	1S_ID_0087
67	average path over time	EA complexity	What is the average path of the EA?	Value, Business, Application< Technology	M_63	network level	primary data	1S_ID_0087
68	clustering techniques	EA domain analysis/ EA modularity	Which domains or modules can be identified in EA?	Process, Application	M_64	group level	primary data	1S_ID_0087
69	clustering coefficient over time	EA complexity	How is the cohesion evolution of the EA network?	Value, Business, Application< Technology	M_65	network level	primary data	1S_ID_0087
70	The hidden structure method	EA cost of change	What is the propagation cost of a change in EA?	Application	M_66	network level	primary data	1S_ID_0216
71	The hidden structure method	EA complexity	What is the architecture flow? (meaning that more than % of the applications are either in, depend on, or are dependent on the Core)	Application	M_66	network level	primary data	1S_ID_0216
72	The hidden structure method	Cost of architectural change	What is the propagation cost of a change in EA?	Application	M_66	network level	primary data	1S_ID_0215
73	The hidden structure method	EA complexity	What is the architecture flow? (meaning that more than % of the applications are either in, depend on, or are dependent on the Core)	Application	M_66	network level	primary data	1S_ID_0215

74	Modularity Q and Girvan/Newman clustering algorithm	*EA domain analysis*	How to find and define appropriate clusters in the EA domain?	application; Business	M_67	*group level*	*Simulate data*	*IS_ID_0142*

Figure G.1: A subset of the EANA containing the main fields of all 74 NAIs.

H

EANA LIBRARY WEB PORTAL

In this appendix, we describe the interfaces used by a researcher, practitioneer or the administrator, to add a new NAI to the library. In addition, we provide architecture details about the web portal.

ADDING A NEW NAI TO THE EANA LIBRARY - USE CASE

Figure H.1 depicts the interface which is used for adding new NAIs. The information contained in the EANA-MM is distributed in four blocks as illustrated in Figure H.1.

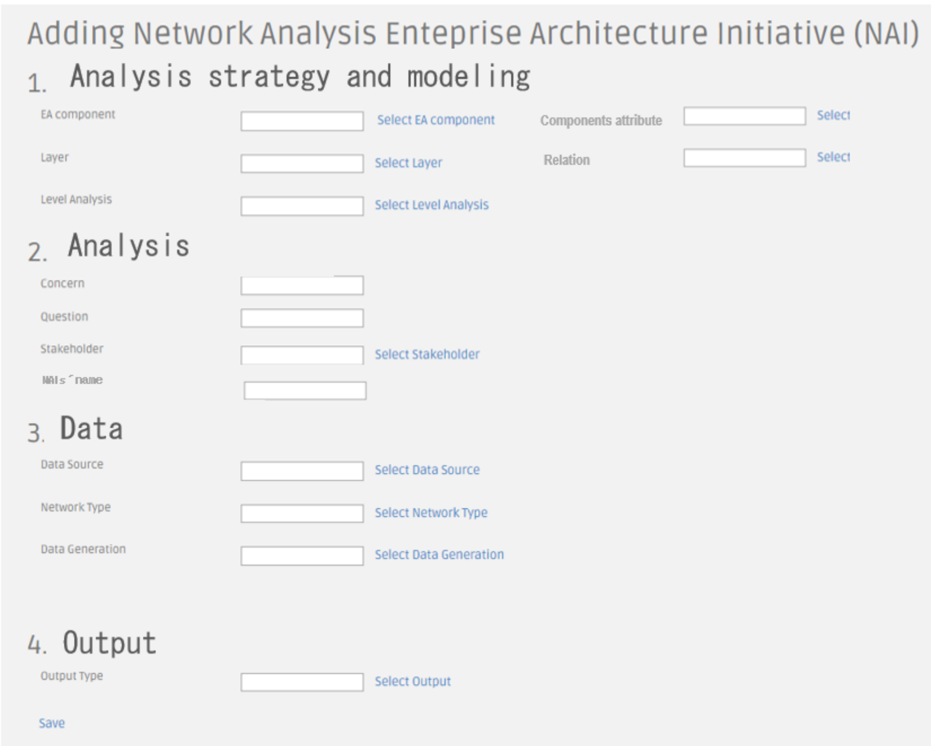

Figure H.1: Interface for adding a new NAI

In the first block ("Analysis strategy and modeling decisions"), the user must inform one or more EA components, their attributes when used, their respective layers and the level of the network analysis that will be applied. The semantic of the relation (e.g. "a business process is supported by an application") must be informed as well. The relation is also characterized in terms of being directed and/or transitive.

In the second block ("Analysis Initiative"), the user defines the targeted concern, the operational question the NAI's name (a label), the stakeholder that might benefit from the NAI and finally, a label to describe the NAI.

In the third block ("Data strategy"), the user indicates the data sources (e.g. Archimate-based tools, documents, interviews and so on) used to collect the primary data and the method used to derive data (derived data), if it is the case. The data can also result from simulation models (simulated data). In "network type, the user must inform if the network is a multi-modal or unimodal one.

Finally, in the fourth block, the type of generated output is informed, e.g., a graph, ranking list, or numerical values for each component.

In all previous information fields, the user can choose a particular pre-defined value (extracted from the primary studies) for the field in question or add a new value to it as well, in case it is not specified in the data grid. In this case, the user can click the "Select" link which directs the use case flow to a second interface. For instance, in Figure H.2, the user can select one or more pre-defined values for the stakeholder field. After that, the flow is redirected back to the main interface ("Adding NAI").

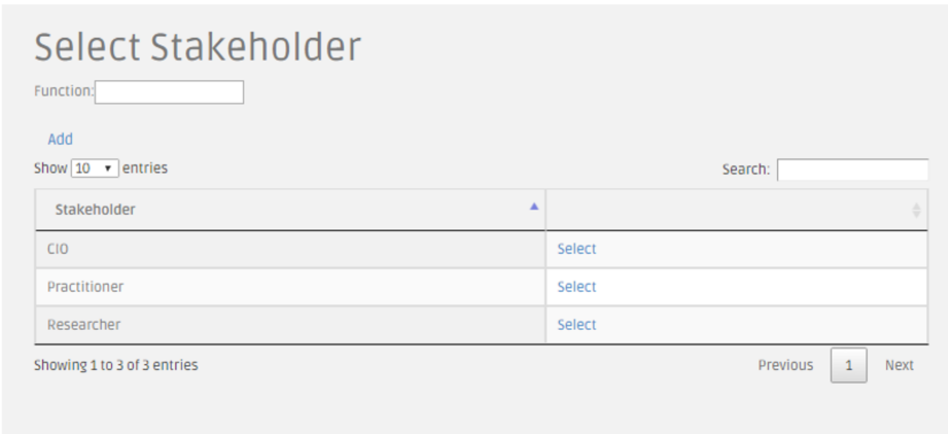

Figure H.2: Interface for selecting stakeholder to be associated to a new NAI

Each added NAI should be validated and eventually aggregated to the EANA library. This activity is performed by users granted with this special permission. This specific procedure and the administration module (related to the user management functionalities), despite already implemented are not described in this document.

TECHNICAL ARCHITECTURE

The web portal was modeled aiming to facilitate the implementation of future requirements in the platform. We adopted an architecture divided into four layers depicted in Figure

H.3:

- Business Classes (Entities): It is composed of the business information entities of the system and their relationships. They are manipulated by components in other layers. One example is the class “br.com.anStrategy.LevelofAnalysis” which represents the network level of analysis adopted in the NAI.
- Persistence Layer: This layer implements the necessary components responsible for the data recovery and persistence from/to the database.
- Business Layer: In this layer, business operations such “Adding an NAI” , “Search for an NAI”, “Add Researcher” are implemented using both previous layers.
- View: It contains the views or system interfaces and related resources (e.g. images). The view layer access the business layer to perform the transactions needed.

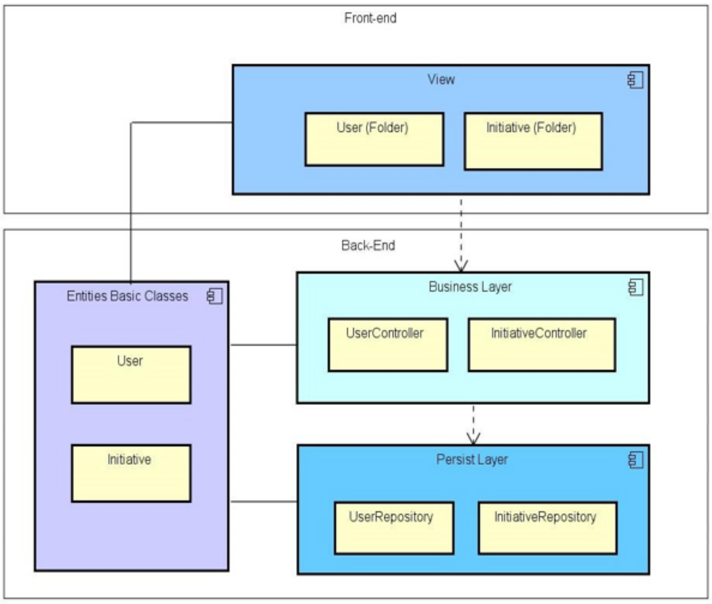

Figure H.3: Technical layers of the EANA library web portal

In addition to the general architectural diagram presented in Figure H.3, other architectural views can be showed to provide a better understanding of the system. In this direction, the platform’s classes diagram is depicted Figure H.4.

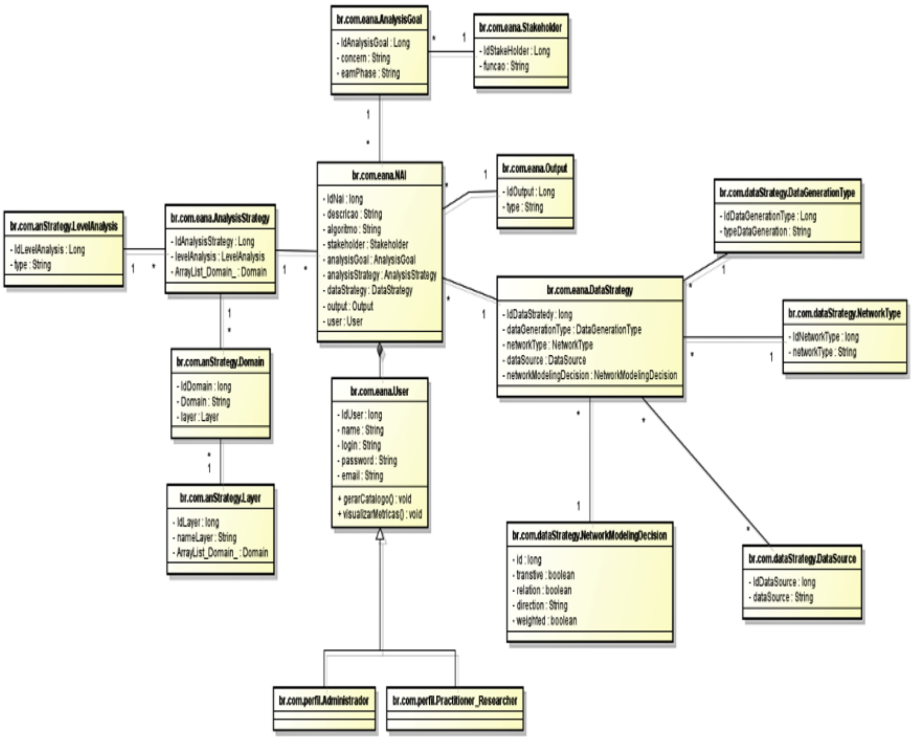

Figure H.4: Class diagram for the EANA library web portal

The development of the EANA web portal was based on JSP[1] (Java server pages) technology and MySQL[2] for database support. They were chosen due to the usefulness of features and to their wide adoption in commercial and academic projects. The research community is invited to adopt the portal and collaborate with its maintenance and development.

[1]http://www.oracle.com/technetwork/java/javaee/jsp/index.html
[2]https://www.mysql.com/

I

ACADEMIC PRODUCTION

Publication outlets

- Santana, A., Simon, D., Fischbach, K., de Moura, H. (2016, September). Combining Network Measures and Expert Knowledge to Analyze Enterprise Architecture at the Component Level. In Enterprise Distributed Object Computing Conference (EDOC), 2016 IEEE 20th International (pp. 1-10). IEEE.
- Santana, A., Kreimeyer, M. , Clo, P. , Moura, H. P. and Fischbach, K. 2016. “Empirical investigation of enterprise architecture analysis based on network measures and expert knowledge: a case from the automotive industry,” in 18th International DSM Conference 2016, São Paulo.
- Santana A., Fischbach K., and Moura H. 2016. “Enterprise Architecture Analysis and Network Thinking : a Literature Review,” in 49th Hawaii International Conference on System Sciences (HICSS 2016), Kauai.
- Júnior, J. C. P., Prochazka, F., Marques, A. C. R., da Silva Araújo, J. I., de Barros Costa, V., Santana, A. T. F. (2013). Análise de redes sociais em projetos: um mapeamento sistemático com foco na estrutura de times em projetos de software. Revista Brasileira de Administração Científica, 4(2), 18-33.
- Carlos, E., Ferraz, G., Costa, V. D. B., Thiago, A., and Santana, F. 2014. “Um mapeamento sistemático sobre a Análise de Redes Sociais aplicada à colaboração em projetos de software,” in Proc. XXXIV Congresso da Sociedade Brasileira de Computação.

Thesis co-supervised

- Pacheco Júnior, J. C. 2014. “Análise de Redes Sociais em Projetos: Um Mapeamento Sistemático,” Thesis (Master) UFPE. Recife, 2014.
- Costa, V. de B. 2014. “Análise de redes sociais aplicada a projetos de software open source”, Thesis (Master), UFPE.

Graduate final projects supervised

- De Souza, A. D.D. Análise de Arquitetura Corporativa baseada no paradigma de redes: Um modelo conceitual e um catálogo de iniciativas. Thesis (Bachelor), UFRPE, 2016.
- Barbosa, A. O., Enterprise architecture: what did we analyze so far? Thesis (Bachelor), UFRPE, 2016.