



Pós-Graduação em Ciência da Computação

JOSÉ MÁRIO DE MENDONÇA LEMOS

**UM PROCESSO PARA PUBLICAÇÃO DE DADOS ABERTOS EM
INSTITUTOS FEDERAIS BASEADO EM BPM**



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

**RECIFE
2017**

José Mário de Mendonça Lemos

Um processo para publicação de Dados Abertos em Institutos Federais baseado em BPM

Este trabalho foi apresentado à Pós-Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Mestre Profissional em Ciência da Computação.

Orientadora: **Prof^a. Dr^a. Bernadette Farias Lóscio**

RECIFE
2017

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

L557p Lemos, José Mário de Mendonça
Um processo para publicação de dados abertos em institutos federais baseado em BPM / José Mário de Mendonça Lemos. – 2017.
147 f.: il., fig., tab.

Orientadora: Bernadette Farias Lóscio.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CIn, Ciência da Computação, Recife, 2017.
Inclui referências, apêndices e anexo.

1. Banco de dados. 2. Integração de dados. 3. Gerenciamento de processos de negócio. I. Lóscio, Bernadette Farias (orientadora). II. Título.

025.04 CDD (23. ed.) UFPE- MEI 2017-122

José Mário de Mendonça Lemos

**Um processo para publicação de Dados Abertos em Institutos Federais
baseado em BPM**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre Profissional em 20 de fevereiro de 2017.

Aprovado em: 20/02/2017.

BANCA EXAMINADORA

Prof. Carla Taciana Lima Lourenço Silva Schuenemann
Centro de Informática / UFPE

Prof. André Câmara Alves do Nascimento
Universidade Federal Rural de Pernambuco

Prof. Bernadette Farias Lóscio
Centro de Informática / UFPE
(Orientador)

Dedico este trabalho a minha família, por todo amor e esforço dedicado na minha educação, e a minha esposa que sempre está ao meu lado nos meus propósitos.

AGRADECIMENTOS

Primeiramente a Deus pela energia e luz iluminando meus caminhos.

A minha esposa Ângela, pelos momentos subtraídos de nossa convivência para serem investidos no mestrado. A família e amigos (Minha avó Tercília (in memoriam); Meu avô Amaro (in memoriam); Minha tia Célia (in memoriam); Meus tios Amaro, Bartolomeu e Antônio; Minha tia Eliane e Meu irmão Leandro), por todo apoio, educação e incentivo que me deram. Agradecimentos especiais aos meus pais Maria Faustina e José Mario.

A minha orientadora, Professora Bernadette Farias Lóscio, por seu apoio e amizade, além de sua dedicação, competência e especial atenção nas revisões e sugestões, fatores fundamentais para a conclusão deste trabalho.

Aos professores do Centro de Informática que destinaram parte de seu precioso tempo para repassarem seus conhecimentos e participarem desta pesquisa.

Ao Instituto Federal de Pernambuco, pelo total apoio, bem como aos funcionários pela competência no suporte de todo período dessa caminhada.

Sou muito grato aos meus colegas de trabalho do IFPE – Campus Ipojuca, pois me deram muita força e suporte quando eu precisei me ausentar para estar presente nas atividades relacionadas ao mestrado.

À Secretaria de Educação Profissional e Tecnológica – SETEC MEC, por ter patrocinado o referido mestrado em rede para os Institutos Federais.

Ao diretor e amigo do Instituto Federal de Pernambuco – Campus Ipojuca, Ênio Camilo de Lima pelo apoio nessa realização.

Ao meu primeiro coordenador e amigo Igor Negromonte, por ter me ajudado no início da minha carreira e acreditado no meu potencial.

E a todos aqueles que direta ou indiretamente contribuíram para a realização deste trabalho.

*“Havendo olhos em quantidade suficiente,
todos os erros são triviais”.*

(Eric Raymond)

RESUMO

Dados são criados a cada instante pelo Governo, inclusive pelas Instituições Educacionais. No entanto, o aproveitamento desses dados pela sociedade é comprometido, uma vez que, os dados não estão disponíveis de forma pública, além disso, quando estão disponíveis, é muito comum encontrar dados que não estão devidamente estruturados ou são de difícil compreensão e manipulação. No âmbito governamental brasileiro, existe um movimento para ampliação da oferta de dados na Web, ora motivado por temáticas como transparência, democracia, crescimento econômico, inovação e valor social aos cidadãos, ora imposto por instrumentos regulatórios. Com o intuito de auxiliar no processo de abertura dos dados, diversas iniciativas têm sido desenvolvidas, incluindo a elaboração de metodologias, guias, padrões, leis, regulamentos e boas práticas. No entanto, cada iniciativa faz de um jeito diferente, não apresentando uma linguagem padronizada. Porém, a imaturidade dos órgãos, principalmente do nicho foco desta pesquisa (Rede Federal), muitas vezes os responsáveis pela abertura de dados, acabam perdidos nesse turbilhão de possibilidades, devido à falta de padronização das mesmas. Além do mais, nenhuma das opções existentes apresenta uma linguagem apropriada para descrição de processos. Com isso, surge a necessidade de uma nova abordagem padronizada para guiar as referidas instituições no processo de publicação de Dados Abertos. Por um lado, esta nova abordagem deve propor alguns passos essenciais, por outro lado não deve limitar a criatividade profissional. Além disso, deve se configurar como um instrumento que determine um planejamento metódico para um processo de abertura de dados. Dessa forma, espera-se harmonizar as áreas envolvidas, facilitando e acelerando a implantação do processo, além de contribuir para a criação de um guia único, padronizado e de referência. Neste contexto, este trabalho apresenta como principal contribuição, uma proposta de processo para Publicação de Dados Abertos na Rede Federal de Educação Profissional, Científica e Tecnológica, baseado nos conceitos de Gerenciamento de Processos de Negócios (BPM, do inglês *Business Process Management*) com a utilização da notação BPMN (*Business Process Model and Notation*).

Palavras-chaves: Dados Abertos. Dados Governamentais. Processos. Gerenciamento de Processos de Negócios. BPM.

ABSTRACT

Data are created at any moment by the Government, including educational institutions. However, the use of these data by society is compromised because the data are not publically available. Besides, when they are available, they are commonly not properly structured or they are difficult to understand and manipulate. In the Brazilian government sphere, there is a movement to increase the supply of data on the Web, sometimes motivated by themes such as transparency, democracy, economic growth, innovation and social value to citizens and sometimes imposed by regulatory instruments. In order to assist in the process to make the data public, several initiatives have been developed, such as the development of methodologies, guides, standards, laws, regulations and good practices. However, each initiative does in a different way, not presenting a standardized language. But the immaturity of the institutions, especially the ones in the Federal sphere (focus of this work), in charge of making the data public, can get confused in this whirlwind of possibilities, because there is no standardization. Moreover, none of the existing options present an appropriate language for describing processes. Thus, there is a need for a new approach to guide these institutions to publicize data. On one hand, this new approach must propose some essential steps. On the other hand, it should not limit professional creativity. In addition, it must be configured as an instrument that determines a methodical planning for a process of making data public. This way, it is hoped the approach can harmonize involved areas, facilitating and accelerating the implementation of the process and contributing to the creation of a single and standard guide. In this context, this work presents, as main contribution, a proposal of process for publication of public data in the Federal Network of Vocational, Scientific and Technological Education based on the concepts of Business Process Management (BPM) using the notation BPMN (Business Process Model and Notation).

Keywords: Open Data. Government Data. Process. Business Process Management. BPM.

LISTA DE FIGURAS

Figura 1.1 - Índice de envolvimento em Dados Abertos pelos órgãos federais.	18
Figura 1.2 - Índice de envolvimento em Dados Abertos pelos Institutos Federais.	19
Figura 1.3 - Fluxograma do Kit de Dados Abertos.	22
Figura 1.4 - Ciclo de vida BPM típico para processos com comportamento previsível.	24
Figura 2.1 - Evolução da Nuvem de Dados Conectados.	30
Figura 2.2 - Esquema das 5 estrelas para Dados Abertos.	31
Figura 2.3 - Ranking do Open Data Barometer.	33
Figura 2.4 - Elementos básicos de um processo.	36
Figura 2.5 - Processos <i>As-is</i> e <i>To Be</i>	38
Figura 2.6 - Processo de catalogação no portal dados.gov.br representado com BPMN.	40
Figura 2.7 - Breve descrição dos elementos básicos da notação BPMN.	40
Figura 2.8 - BPMS Bizagi versão Modeler.	41
Figura 3.1 - Instituições da Rede Federal.	43
Figura 3.2 - Quantitativo de pedidos de acesso à informação de todos os órgãos públicos no e-SIC.	44
Figura 3.3 - Quantitativo de Pedidos de acesso à informação de todos Institutos Federais no e-SIC.	45
Figura 3.4 - Forma de coleta na análise da transparência ativa.	46
Figura 3.5 - Processo de pedido de informação via e-SIC.	48
Figura 3.6 - Transparência Ativa da Rede Federal por formato.	50
Figura 3.7 - Transparência Ativa da Rede Federal de acordo com esquema de 5 estrelas.	50
Figura 3.8 - Reatividade das instituições em relação ao espaço temporal para respostas.	52
Figura 3.9 - Transparência Passiva em relação ao conteúdo.	53
Figura 3.10 - Transparência Passiva em relação ao formato.	54
Figura 4.1 - Relação dos processos analisados	64
Figura 4.2 - Sumarização dos resultados das recomendações classificadas como obrigatórias.	86
Figura 4.3 - Sumarização dos resultados das recomendações classificadas como desejáveis.	86
Figura 5.1 - Levantamento das iniciativas em Mapeamento de Processos nos Institutos Federais.	89
Figura 5.2 - Modelo de processo inovação-decisão	92
Figura 5.3 - Ciclo de Vida dos Dados na <i>Web</i>	95
Figura 5.4 - Alocação dos artefatos no modelo.	97
Figura 5.5 - Planilha de Abertura dos Dados.	98
Figura 5.6 - Visão Global do Processo ODSP	99
Figura 6.1 - Processo de planejamento da coleta de dados e evidências do estudo de caso.	108
Figura 6.2 - Processo de análise das evidências coletadas em estudos de casos.	109
Figura 6.3 - Fases do Ciclo de vida BPM típico para processos utilizadas no Estudo de Caso.	111
Figura 6.4 - Categorias do Vocabulário Controlado do Governo Eletrônico (VCGE).	115
Figura 6.5 - Termos de uso para o Portal de Dados Abertos do IFPE – Campus Ipojuca.	118
Figura 6.6 - Entendimento da equipe a respeito da necessidade de um PDA na organização.	119
Figura 6.7 - Conjunto de dados em seu formato inicial (XLSX).	119
Figura 6.8 - Captura da tela referente ao arquivo CSV do conjunto de dados.	120
Figura 6.9 - Conjunto de dados do IFPE - Campus Ipojuca catalogado no portal dados.gov.br.	121
Figura 6.10 - Conjunto de dados do IFPE - Campus Ipojuca visualizado no portal dados.gov.br.	121
Figura 6.11 - Índice de conhecimento da equipe em relação a temática de Dados Abertos.	122

LISTA DE TABELAS

Tabela 2.1 - Quantitativos de conjuntos de dados dos portais de Dados Abertos nacionais.	34
Tabela 4.1 - Quantitativo de recomendações para "Preparar Partes Interessadas" nos processos de publicação de dados abertos analisados.	66
Tabela 4.2 - Quantitativo de recomendações para "Selecionar Conjuntos de Dados" nos processos de publicação de dados abertos analisados.	69
Tabela 4.3 - Quantitativo de recomendações para "Modelar os dados" nos processos de publicação de dados abertos analisados.	72
Tabela 4.4 - Quantitativo de recomendações para "Especificar uma licença apropriada" nos processos de publicação de dados abertos analisados.	74
Tabela 4.5 - Quantitativo de recomendações para "Estabelecer bons identificadores universais (URIs)" nos processos de publicação de dados abertos analisados.	75
Tabela 4.6 - Quantitativo de recomendações para "Utilização de vocabulário padrão" nos processos de publicação de dados abertos analisados.	77
Tabela 4.7 - Quantitativo de recomendações para "Converter e enriquecer dados" nos processos de publicação de dados abertos analisados.	79
Tabela 4.8 - Quantitativo de recomendações para "Prover acesso automatizado aos dados" nos processos de publicação de dados abertos analisados.	81
Tabela 4.9 - Quantitativo de recomendações para "Anunciar os novos conjuntos de dados para o público" nos processos de publicação de dados abertos analisados.	82
Tabela 4.10 - Quantitativo de recomendações para "Estabelecer um contrato social para os dados publicados" nos processos de publicação de dados abertos analisados.	84
Tabela 6.1 - Resultados da análise para “Impacto” e “Dificuldade” dos conjuntos de dados.	116

LISTA DE QUADROS

Quadro 4.1- Processos de publicação de Dados Abertos analisados.	56
Quadro 4.2 - Recomendações para publicação de Dados Abertos Governamentais.	61
Quadro 5.1 - Classificação das recomendações no processo.	93
Quadro 5.2 - Sequenciamento de fases, subprocessos, tarefas e artefatos do processo.	100
Quadro 5.3 - Quadro comparativo dos processos de publicação de Dados Abertos Governamentais nacionais.....	106
Quadro 6.1 - Perfil dos Servidores do IFPE - Campus Ipojuca participantes do estudo de caso.....	109
Quadro 6.2 - Observações pontuadas e ações tomadas para correção do processo ODSP.	111

LISTA DE ABREVIATURAS E SIGLAS

API	Application Programming Interface
BPLD	Best Practices for Publishing Linked Data
BPM	Business Process Management
BPMI	Business Process Management Initiative
BPMN	Business Process Model and Notation
BPMS	Business Process Management System
BPM CBOK	Business Process Management Common Book of Knowledge
CSV	Comma-separated Values
EPC	Event-driven Process Chain
ePING	Padrões de Interoperabilidade de Governo Eletrônico
e-PMG	Padrão de Metadados do Governo Eletrônico
ETL	Extract Transform Load
e-Gov	Governo Eletrônico
e-SIC	Sistema Eletrônico do Serviço de Informação ao Cidadão
FTP	File Transfer Protocol
HTTP	Hypertext Transfer Protocol
IFC	Instituto Federal Catarinense
IFES	Instituições Federais de Ensino Superior
IFPE	Instituto Federal de Pernambuco
INDA	Infraestrutura Nacional de Dados Abertos
JSON	JavaScript Object Notation
LAI	Lei de Acesso a Informação
LOD	Linked Open Data
LOGD	Linked Open Government Data
KML	Keyhole Markup Language

NIC.br	Núcleo de Informação e Coordenação do Ponto BR
ODS	OpenDocument Spreadsheet
ODSP	Open Data Starter Process
ODT	OpenDocument Text
OGP	Open Government Partnership
OMG	Object Management Group
PDA	Plano de Dados Abertos
PDF	Portable Document Format
PDTI	Plano Diretor de Tecnologia da Informação
PMBok	Project Management Body of Knowledge
RDF	Resource Description Framework
SIC	Serviço de Informação ao Cidadão
SPARQL	SPARQL Protocol and RDF Query Language
SEGES	Secretaria de Gestão do Ministério do Planejamento
SDPS	Society for Design and Process Science
SISP	Sistema de Administração dos Recursos de Informação e Informática
SLTI	Secretaria de Logística e Tecnologia da Informação
TIC	Tecnologias da Informação e Comunicação
UML	Unified Modeling Language
URI	Uniform Resource Identifier
VCGE	Vocabulário Controlado do Governo Eletrônico
WWW	World Wide Web
W3C	World Wide Web Consortium
XML	eXtensible Markup Language

SUMÁRIO

1	Introdução	17
1.1	Motivação	17
1.2	Problemática	18
1.3	Escopo.....	20
1.4	Proposta	22
1.5	Objetivos.....	23
1.6	Metodologia.....	23
1.7	Organização da dissertação.....	25
2	Fundamentação Teórica.....	27
2.1	Dados	27
2.2	Dados Abertos	28
2.2.1	<i>Dados Abertos Conectados.....</i>	<i>29</i>
2.2.2	<i>Esquema das 5 estrelas para Dados Abertos</i>	<i>31</i>
2.2.3	<i>Dados Abertos Governamentais</i>	<i>32</i>
2.2.4	<i>Legislação para Dados Abertos no Brasil.....</i>	<i>35</i>
2.3	Gerenciamento de Processos de Negócio	36
2.3.1	<i>Processos de Negócio</i>	<i>36</i>
2.3.2	<i>Conceitos de Gerenciamento de Processos de Negócio</i>	<i>37</i>
2.3.3	<i>Modelagem de Processos.....</i>	<i>37</i>
2.3.4	<i>Notação para Modelagem de Processos.....</i>	<i>39</i>
2.3.5	<i>Business Process Management System (BPMS)</i>	<i>41</i>
2.4	Considerações finais	41
3	Análise da Transparência Ativa e Passiva na Rede Federal de Educação Profissional, Científica e Tecnológica.....	43
3.1	A Rede Federal de Educação Profissional, Científica e Tecnológica.....	43
3.2	Transparência Ativa e Passiva	44
3.3	Metodologia.....	45
3.3.1	<i>Metodologia para análise da transparência ativa.....</i>	<i>45</i>
3.3.2	<i>Metodologia para análise da transparência passiva.....</i>	<i>47</i>
3.4	Resultados.....	49
3.4.1	<i>Resultados da análise da transparência ativa</i>	<i>49</i>
3.4.2	<i>Resultados da análise da transparência passiva</i>	<i>51</i>

3.5	Considerações Finais	54
4	Análise dos processos para publicação de Dados Abertos Conectados.....	55
4.1	Critérios de seleção dos processos para publicação de Dados Abertos Conectados.....	55
4.2	Descrição dos processos para publicação de Dados Abertos Conectados adicionados na nova análise	58
4.3	Metodologia para Análise dos Processos.....	60
4.4	Análise dos resultados	66
4.4.1	<i>Recomendações para “Preparar Partes Interessadas”</i>	<i>66</i>
4.4.2	<i>Recomendações para “Selecionar Conjuntos de Dados”</i>	<i>69</i>
4.4.3	<i>Recomendações para “Modelar os dados”</i>	<i>72</i>
4.4.4	<i>Recomendações para “Especificar uma licença apropriada”</i>	<i>74</i>
4.4.5	<i>Recomendações para “Estabelecer bons identificadores universais (URIs)”</i>	<i>75</i>
4.4.6	<i>Recomendações para “Utilização de vocabulário padrão”</i>	<i>77</i>
4.4.7	<i>Recomendações para “Converter e enriquecer dados”</i>	<i>79</i>
4.4.8	<i>Recomendações para “Prover acesso automatizado aos dados”</i>	<i>80</i>
4.4.9	<i>Recomendações para “Anunciar os novos conjuntos de dados para o público”</i>	<i>81</i>
4.4.10	<i>Recomendações para “Estabelecer um contrato social para os dados publicados”</i>	<i>84</i>
4.5	Considerações Finais	85
5	O Processo ODSP – Open Data Starter Process	87
5.1	Visão geral do modelo ODSP	87
5.2	Principais características do processo ODSP.....	90
5.2.1	<i>Características herdadas do Kit de Dados Abertos do Governo Federal.....</i>	<i>90</i>
5.2.2	<i>Características herdadas dos Processos de Inovação.....</i>	<i>91</i>
5.3	Componentes do processo ODSP	92
5.3.1	<i>Subprocessos e tarefas do processo ODSP</i>	<i>93</i>
5.3.2	<i>Fases do processo ODSP.....</i>	<i>95</i>
5.3.3	<i>Artefatos do processo ODSP</i>	<i>96</i>
5.4	Estrutura geral do processo ODSP.....	98
5.5	Considerações Finais	106
6.	Avaliação	108
6.1	Metodologia para a condução do Estudo de Caso	108
6.2	Estudo de caso	109
6.2.1	<i>Execução do Processo</i>	<i>112</i>
6.3	Considerações Finais	121

7	Considerações Finais	123
7.1	Contribuições.....	123
7.2	Limitações.....	124
7.3	Trabalhos futuros.....	125

REFERÊNCIAS.....	126
-------------------------	------------

APÊNDICE A – Critérios utilizados para análise do nível de sigilo dos dados a serem abertos.....	132
---	------------

APÊNDICE B – Critérios utilizados para a priorização de dados a serem abertos	134
--	------------

ANEXO A – Recomendações para publicação de Dados Abertos e Dados Abertos Conectados.....	139
---	------------

1 Introdução

Em 1994, seis anos após a criação da WWW (*World Wide Web*), o físico inglês Sir Timothy John Berners-Lee chamou atenção para uma revisão de como as informações estavam sendo publicadas na *Web*, ressaltando que os mesmos princípios que permitiram o crescimento da “*Web* de documentos” não foram aplicados aos dados (BIZER; HEATH; BERNERS-LEE, 2009). Tal declaração motivou o surgimento de um movimento transformador denominado “*Web* de Dados”, engajando diversas empresas, governos, institutos de pesquisas, sociedades e outros para a criação de um ecossistema de produção e consumo de dados que facilite a descoberta de novos conhecimentos, aproveitando o máximo dos dados disponíveis. O grande problema é que a maioria dos dados está disponível para consumo humano e não por máquinas (ANTONIOU; HARMELEN, 2008). Mas os seres humanos estão impossibilitados de buscar e acessar todas essas informações de forma eficiente e consistente sem ajuda das máquinas. Os esforços de diversas iniciativas focaram na publicação dos dados de forma estruturada, juntamente com semântica para ambos os atores (Humanos e Máquinas). Tais esforços resultaram em diversas novas abordagens, novos conceitos e novas tecnologias, incluindo os conceitos de Dados Abertos, Dados Abertos Conectados e Dados Abertos Governamentais.

1.1 Motivação

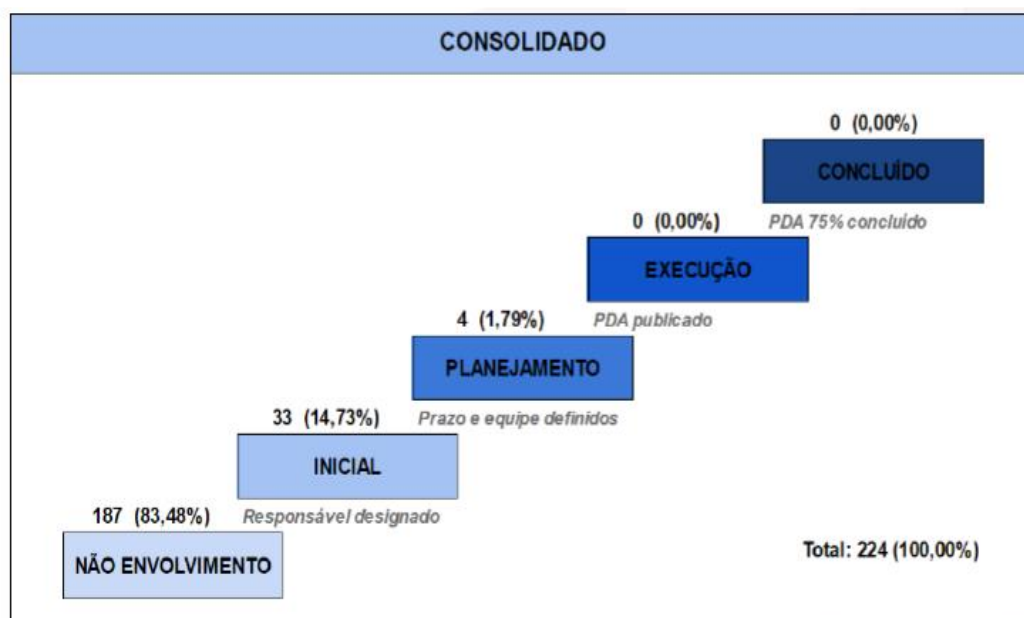
Atualmente, existe um movimento aflorando nos governos para ampliação da sua oferta de dados na Web ora motivado por temáticas como transparência, democracia, crescimento econômico, inovação e valor social aos cidadãos, ora imposto por instrumentos regulatórios. O fornecimento e a utilização de dados no âmbito governamental apresentam um crescimento ao longo dos anos, reflexo da criação de diversas iniciativas em diferentes países, incluindo o Brasil. Existem diversos desdobramentos da Administração Pública Federal para criação de iniciativas com o intuito de publicar as informações de acordo a Lei de Acesso a Informação (LAI). Em alguns casos, as iniciativas atendem aos princípios de Dados Abertos, preocupando-se com a publicação dos dados de forma estruturada, aberta e conectada. Uma parte significativa das iniciativas governamentais está centrada no desenvolvimento de portais como ponto único para todos os dados (KALAMPOKIS; TAMBOURIS; TARABANIS, 2011). Mas isso não exime cada órgão público da responsabilidade no fornecimento desses dados em seus portais. Mesmo com a obrigatoriedade da publicação prevista em legislação, as

entidades governamentais não identificam uma fórmula única para a publicação de Dados Abertos, nem tampouco para conectar esses Dados, revelando uma problematização com diversos desafios legais, organizacionais, tecnológicos e culturais para os órgãos públicos (KALAMPOKIS; TAMBOURIS; TARABANIS, 2011).

1.2 Problemática

Na prática, os órgãos e entidades públicas deveriam divulgar as informações por eles produzidas ou custodiadas em local de fácil acesso, de forma estruturada e aberta (MEDEIROS; MAGALHÃES; PEREIRA, 2013; ALCANTARA et al, 2015). Fomentar a disponibilização e o uso de Dados Abertos é um dos objetivos estratégicos elencados na Estratégia de Governança Digital (BRASIL, 2016), instituída pelo Decreto nº 8.638, de 15 de janeiro de 2016, tendo o papel de balizadora das iniciativas relativas à Governança Digital na administração direta, autárquica e fundacional do Poder Executivo Federal Brasileiro. Tal estratégia foi complementada posteriormente com o decreto lei nº 8.777, que institui a Política de Dados Abertos do Poder Executivo Federal, determinando a publicação de um Plano de Dados Abertos (PDA) pelos órgãos, no prazo de sessenta dias (BRASIL, 2016).

Figura 1.1 - Índice de envolvimento em Dados Abertos pelos órgãos federais.

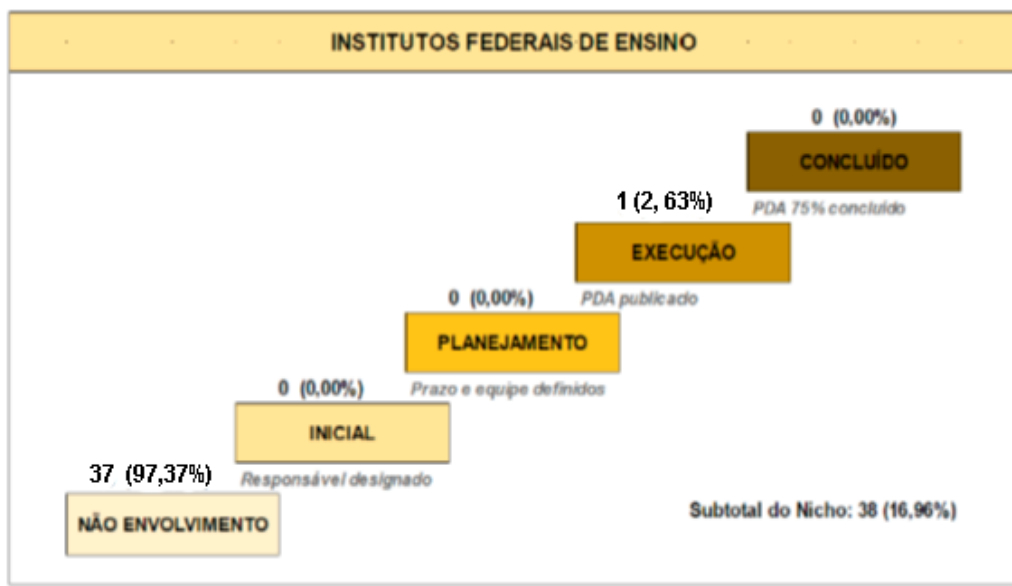


Fonte: Workshop de TIC das IFES. (2016)

A tarefa de abertura dos dados não é algo trivial, realidade perceptível ao observarmos os resultados da Figura 1.1, demonstrando que a taxa dos órgãos que não têm envolvimento com Dados Abertos chega a 83,48%, juntamente com a grande dificuldade de se encontrar dados ou informações nos portais em virtude da não obediência aos princípios de publicação de Dados Abertos. Em geral, os portais utilizam formatos de dados inadequados, além de disponibilizarem dados incompletos, desatualizados, incompreensíveis por máquina, proprietários e com outras inobservâncias (MACHADO; DE OLIVEIRA, 2011; MODESTO; FERNEDA; SANT'ANA, 2014; ALCANTARA et al, 2015; PEDROSO; TANAKA; CAPPELLI, 2013).

A aplicabilidade dos princípios de Dados Abertos não está restrita a um determinado contexto do governo, tendo aplicações em áreas como Educação, Política, Saúde, Segurança, Trânsito e outras. Mas o objetivo da pesquisa foi elevar uma determinada atenção para análise das instituições governamentais do contexto educacional, especificamente a Rede Federal de Educação Profissional. A Figura 1.2 demonstra que a taxa das instituições pertencentes à Rede Federal que não têm envolvimento com Dados Abertos chega a 97,37%.

Figura 1.2 - Índice de envolvimento em Dados Abertos pelos Institutos Federais.



Fonte: Adaptado do Workshop de TIC das IFES. (2016)

Tais dificuldades na criação de um Ecossistema de Dados Abertos nessas instituições refletem a ausência de processos organizacionais e da adoção de padrões existentes, assim como a falta de pessoas capacitadas para a realização das atividades necessárias ao planejamento e implantação do processo, apoiando a criação e manutenção do referido

ecossistema (ALCANTARA et al, 2015; PEDROSO; TANAKA; CAPPELLI, 2013). O Consórcio *World Wide Web* (W3C) tem um papel importante na educação do processo de Publicação de Dados Governamentais Abertos e Conectados, patrocinando treinamento nas tecnologias necessárias para os profissionais de TICs no setor público brasileiro (BREITMAN, 2012). Mas, mesmo assim, os órgãos públicos vêm discutindo problemas básicos como definição de vocabulário comum para a representação dos dados, bem como a adoção de metodologias e ferramentas (BREITMAN, 2012). A percepção que temos das referidas instituições é que as mesmas estão vivendo em um universo desconhecido, buscando referências de métodos e processos de organizações similares, estudando e verificando a possibilidade de adoção dos mesmos, ou seja, buscando melhores práticas.

Neste cenário de oportunidades e desafios, surgiram e prosperaram vários instrumentos para auxiliar no processo de abertura dos dados, como, por exemplo, metodologias, guias, padrões, leis, regulamentos e boas práticas. No entanto, tais instrumentos utilizam diferentes formas de linguagem, não apresentando uma padronização para o mesmo processo. São tantas as opções que chegam a causar dificuldades às instituições imaturas no estabelecimento de um referencial para abertura dos dados (PEDROSO; TANAKA; CAPPELLI, 2013; XAVIER, 2010). Além do mais, nenhuma das opções apresenta uma linguagem apropriada para descrição de processos, limitando o seu detalhamento. Neste caso, existe uma necessidade e um grande desafio que consiste na elaboração de um processo único, padronizado e de referência para todos os setores governamentais (DE FARIA CORDEIRO, 2011).

1.3 Escopo

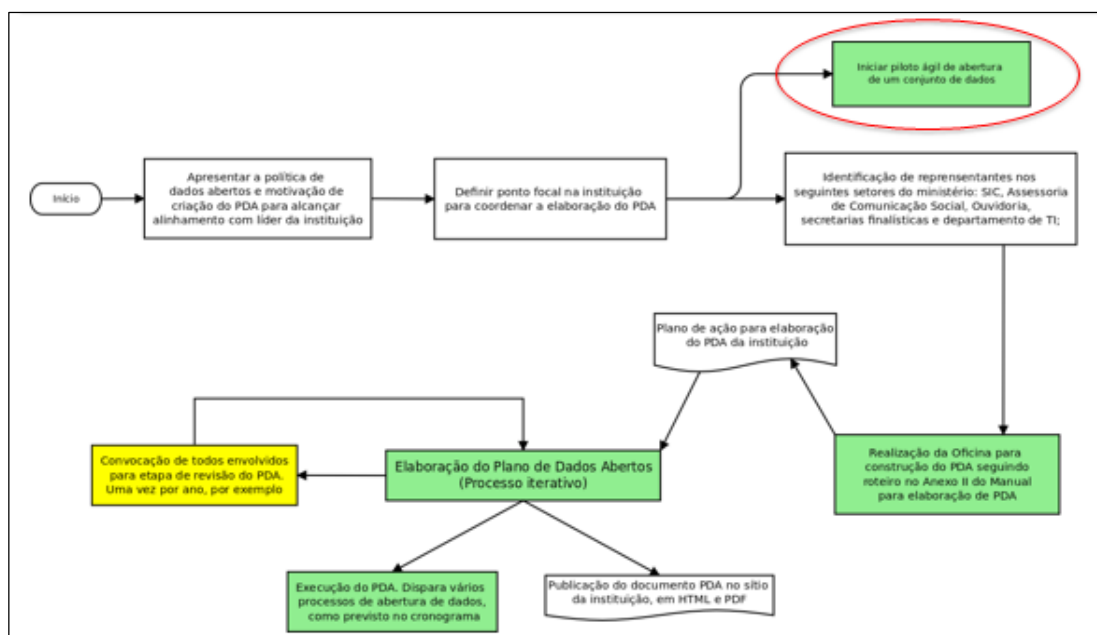
O governo é beneficiado com o processo de abertura dos dados, desenvolvendo pesquisas juntamente com o apoio da população, gerando melhorias para os serviços por meio de novas soluções. Por isso, nada mais justo do que o desenvolvimento de ações relacionadas à Dados Abertos serem uma contrapartida do próprio governo. Assim, o governo consolidou um conjunto de documentos que descreve o processo de abertura de dados (Figura 1.3), métodos e técnicas oferecidas como um kit para implementação de abertura dos dados no âmbito da administração direta, autárquica e fundacional do Poder Executivo Federal (BRASIL, 2014). No entanto, o Kit de Dados Abertos¹ não apresenta um detalhamento necessário para uma instituição imatura, limitado por não utilizar uma linguagem adequada para processos. Além

¹ <http://kit.dados.gov.br/>

do mais, o foco do Kit é a elaboração do Plano de Dados Abertos (PDA), desconsiderando várias questões técnicas e operacionais.

O Kit de Dados Abertos inclui, em um nível macro, todas as atividades que uma instituição deve considerar para programar e executar uma política de dados abertos. Nesse processo se destaca o Plano de Dados Abertos (PDA) como documento central onde são publicados os compromissos e prazos para execução. O problema é que a maioria dos órgãos, principalmente o nicho (Institutos Federais) o qual a pesquisa está focando, faz parte do percentual enquadrado em um patamar baixo (Nível 0) considerando a escala de nível de maturidade (Nível 0 – Nível 3) da Estratégia de Governança Digital (BRASIL, 2016), colocando-as numa condição de desconhecimento e incompreensão para a proposição de um PDA. Mesmo sendo opcional, o Processo “Iniciar piloto ágil de abertura de um conjunto de dados” elencado no Fluxograma do Kit de Dados Abertos (Figura 1.3), antecedente à elaboração do PDA, tem que ser visto como um processo para a elevação da maturidade dessas instituições frente a temática de Dados Abertos. As experimentações e as lições aprendidas, promovidas por esse processo, permitem, em um espaço temporal reduzido, um resultado prático de abertura de dados para a sociedade, além do amadurecimento no processo pela equipe (BRASIL, 2014). As saídas do Processo “Iniciar piloto ágil de abertura de um conjunto de dados” são insumos necessários para a criação de um PDA realista e factível. Como o processo “Iniciar piloto ágil de abertura de um conjunto de dados” está em um nível macro, chamamos atenção para uma proposta de detalhamento utilizando uma linguagem apropriada de processo, adicionando subprocessos e tarefas, no intuito de constituir uma estrutura referencial para os institutos federais.

Figura 1.3 - Fluxograma do Kit de Dados Abertos.



Fonte: <http://kit.dados.gov.br>.

1.4 Proposta

O trabalho desenvolvido e apresentado por essa dissertação tem como principal contribuição um processo para Publicação de Dados Abertos na Rede Federal de Educação Profissional, Científica e Tecnológica, baseado nos conceitos de BPM com a utilização da notação BPMN. A contribuição constitui de uma abordagem para publicação de Dados Abertos, por meio de passos preestabelecidos e que não deve limitar a criatividade profissional, mas deve ser um instrumento que determine um planejamento metódico e harmonize as áreas envolvidas, facilitando e acelerando a implantação do processo de abertura de dados. Este trabalho lida com problemas relacionados à publicação de Dados Abertos vivenciados por profissionais no âmbito governamental. Situa-se na temática de Dados Abertos, baseado na ênfase em Gestão de Tecnologia da Informação do Mestrado Profissional em Ciência da Computação, do Centro de Informática da Universidade Federal de Pernambuco.

Nesta ênfase são abordados temas relacionados à Gestão de Processos, Gestão de Dados, Gestão da Organização e Gestão de Projetos, incluindo metodologias para o desenvolvimento de software. Este trabalho se caracteriza como multidisciplinar, visto que o seu objeto de estudo (Dados Abertos Governamentais) apresenta uma mescla de áreas distintas, reflexo do

envolvimento da área governamental em diferentes contextos, tendo aplicações em áreas como Educação, Política, Saúde, dentre outras. Ademais, a pesquisa está adequada com os objetivos do mestrado profissional, tratando de temas de interesse direto para a organização em que o pesquisador atua. A interação dos acadêmicos com as problemáticas enfrentadas no dia-a-dia nos processos empresariais contribui para a melhoria das instituições e, conseqüentemente, para o desenvolvimento de cada região na qual a mesma esteja inserida.

1.5 Objetivos

Este trabalho tem como objetivo geral propor um processo para publicação de Dados Abertos na Rede Federal de Educação Profissional, Científica e Tecnológica, baseado nos conceitos de BPM e na notação BPMN.

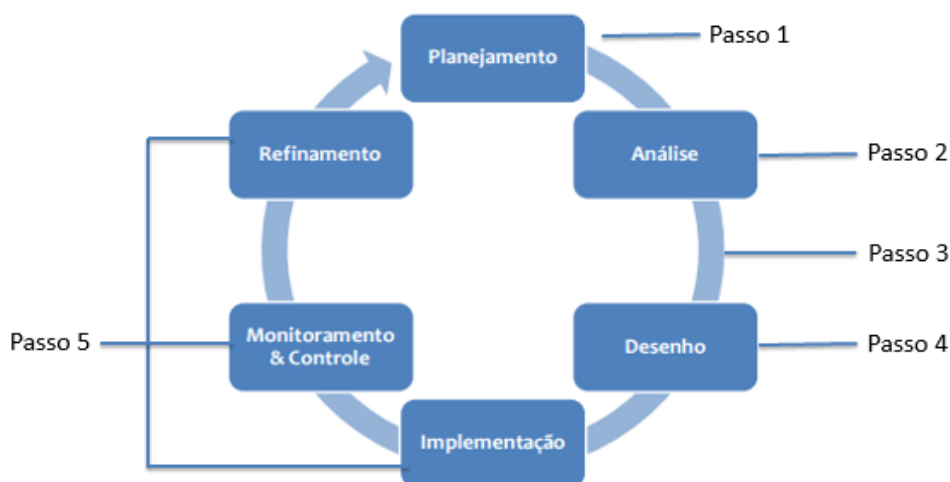
Como objetivos específicos deste trabalho destacam-se:

- Apresentar a maturidade dos Institutos Federais em Dados Abertos, considerando a transparência ativa (Capacidade da instituição em disponibilizar os dados em seus portais independente de requerimentos) e passiva (Capacidade do órgão em disponibilizar os dados motivados por requerimentos);
- Expor recomendações essenciais para publicação de dados a partir de uma revisão da literatura sobre processos de publicação de Dados Abertos, Dados Abertos Governamentais e Dados Abertos Conectados;
- Discorrer os resultados da aplicação do processo proposto em uma instituição real, para fins de avaliação e aperfeiçoamento.

1.6 Metodologia

A pesquisa está pautada na concepção de um novo processo para publicação de Dados Abertos. Por isso, o desenvolvimento do trabalho tomou como base as fases do ciclo de vida típico para processos (Figura 1.4) apresentado pelo CBOK (2013), servindo de orientação para fundamentação, concepção e aplicação prática do processo. Mesmo sendo um ciclo de vida, tal modelo não está limitado ao comportamento cíclico, flexibilizando a aplicação de fases isoladas ou a execução de várias fases sequencialmente, sem a necessidade de completar o ciclo.

Figura 1.4 - Ciclo de vida BPM típico para processos com comportamento previsível.



Fonte: Adaptado do CBOK (2013).

O planejamento (**Passo 1**), antes de qualquer outra ação da pesquisa, é o principal ponto para o bom desenvolvimento da mesma. Neste passo, foram definidos os objetivos; Alinhamento entre a estratégia e a execução do projeto de pesquisa; Procedimentos para condução da pesquisa; Cronograma; Envolvidos; Recursos Necessários e Custos. Em seguida, o **passo 2** (Análise) procurou entender como o nicho (Rede Federal) para qual o processo foi proposto, lidava com as questões relacionadas à gestão e disponibilização de dados, inclusive Dados Abertos. Para isso, analisamos a maturidade dos Institutos Federais frente a temática considerando a transparência ativa (Capacidade da instituição em disponibilizar os dados em seus portais independente de requerimentos) e passiva (Capacidade do órgão em disponibilizar os dados motivado por requerimentos). Já conhecendo o cenário, partimos para um trabalho de análise, discussão e busca de melhores praticas em outros processos. Tais ações estão representadas na metodologia pelo o **passo 3**, no qual realizamos uma revisão da literatura sobre processos de publicação de dados abertos, dados abertos governamentais e dados abertos conectados. Nesta etapa, extraímos recomendações essenciais para publicação de dados a partir dos processos analisados, com intuito de compor o processo proposto. A análise dos processos não foi iniciada da estaca zero, replicamos a metodologia de ÁVILA (2015) para análise de mais 07 processos de publicação de Dados Abertos além dos 15 processos previamente analisados pelo o autor. No **passo 4** (Desenho), o processo proposto para publicação de Dados Abertos foi desenhado na notação BPMN com base nos conceitos de BPM e nas análises anteriores a este passo. O **passo 5** corresponde a avaliação do processo

através de um estudo de caso em 3 etapas. Na etapa de Implementação, aplicamos o processo na organização de acordo com as especificações desenvolvidas no passo 4 (Desenho). Na etapa “Monitoramento e Controle”, extraímos os resultados do processo em comparação ao esperado, além de documentar todas as dificuldades encontradas e as lições aprendidas pela organização do estudo de caso. Na fase de “Refinamento”, corrigimos o processo de acordo com os dados coletados nas evidências do estudo durante a fase de “Monitoramento e Controle”.

1.7 Organização da dissertação

A dissertação aqui apresentada contém sete capítulos organizados e distribuídos da seguinte maneira:

- Capítulo 1: Neste capítulo são apresentadas a motivação e a contextualização do trabalho, juntamente com a problemática, o escopo e os objetivos do trabalho aqui proposto.
- Capítulo 2: É apresentada nesta parte a fundamentação teórica, a qual aborda os principais conhecimentos utilizados nesta dissertação, enfatizando os temas principais desta pesquisa.
- Capítulo 3: Aqui trazemos os resultados da análise da maturidade da Rede Federal de Educação Profissional, Científica e Tecnológica na gestão e publicação de dados, considerando a transparência ativa e passiva.
- Capítulo 4: Nesta etapa é apresentada uma análise de processos existentes na publicação de Dados Abertos ou Dados Conectados, servindo de comparação e complementação para os métodos e processos do modelo de processo proposto.
- Capítulo 5: Esta divisão do trabalho apresenta de forma detalhada o processo de publicação de Dados Abertos proposto para a Rede Federal de Educação Profissional, Científica e Tecnológica, incorporando elementos extraídos dos capítulos anteriores.

- Capítulo 6: Após todo trabalho de concepção do processo apresentado nos capítulos anteriores, aqui apresentamos a avaliação do mesmo em um ambiente real.
- Capítulo 7: Por fim, este capítulo apresenta as conclusões acerca do trabalho apresentado, bem como algumas contribuições geradas pelo o trabalho e as sugestões de trabalhos futuros.

2 Fundamentação Teórica

Este capítulo apresenta os principais conhecimentos utilizados nesta dissertação, enfatizando os temas principais da pesquisa. Inicialmente, na seção 2.1, são apresentados os conceitos de Dados e suas classificações. Em seguida, o tópico 2.2 apresenta os conceitos de Dados Abertos e os princípios que norteiam essa temática. Nesta mesma parte, discorre-se sobre conceitos de Dados Abertos Conectados e Dados Abertos Governamentais, destacando a evolução no cenário mundial e nacional. A área tem se desenvolvido bastante e muitas iniciativas (principalmente governamentais) são elencadas no capítulo. Cada vez mais, as instituições estão se preocupando com a publicação dos dados de forma estruturada, aberta e conectada. No entanto, existe uma preocupação com a qualidade dos dados, que diretamente influenciam na qualidade dos produtos gerados por eles. Por esta razão, um trecho descreve o funcionamento de um esquema de classificação para Dados Abertos, muito utilizado como referência não apenas para classificação dos Dados, mas como roteiro para instituições disponibilizarem os dados com qualidade. A evolução das legislações brasileiras frente à temática finaliza esta etapa destinada aos Dados Abertos. O item 2.3 introduz o conceito básico do que vem ser um processo, em seguida foca nas definições de Gerenciamento de Processos de Negócios (BPM, do inglês Business Process Management) e seus conceitos relacionados, incluindo a Modelagem de Processo, juntamente com os elementos utilizados pela notação BPMN. Finaliza-se com as opções das ferramentas de apoio no processo de modelagem, sendo realizada uma breve e sucinta análise comparativa entre elas.

2.1 Dados

ELMASRI & NAVATHE (2005) definem “**Dado**” como um fato isolado que pode ser gravado e possui um significado implícito. MACHADO & DE OLIVEIRA (2011) trazem a seguinte classificação para os diferentes tipos de Dados:

Dados estruturados: São dados organizados de acordo com critérios definidos, respeitando vários atributos, delimitando escopo, tipos, entre outros.

Dados semiestruturados: São dados que nem sempre é possível prever todos os aspectos. Alguns atributos gerais podem ser conhecidos, enquanto outros poderão ser adicionados posteriormente.

Dados não estruturados: São dados em que nenhum esquema foi especificado, apenas o conteúdo é apresentado.

Quando um dado é agrupado a outros dados com associação de “**Metadados**” (dados sobre dados), formam-se uma coleção com um significado explícito, passando ser uma “**Informação**”. As informações processadas tornam-se “**Conhecimento**”.

2.2 Dados Abertos

Dados Abertos são dados publicados de maneira pública, legível por máquinas, com a finalidade de serem reutilizados e redistribuídos sem nenhuma restrição (BAUER; KALTENBOCK, 2011; KALAMPOKIS; TAMBOURIS; TARABANIS, 2011; MACHADO; DE OLIVEIRA, 2011). Durante o processo de abertura de dados, devem ser obedecidos os oito princípios estabelecidos pela comunidade de Dados Abertos, são eles:

1. **Completos:** Todo o conjunto dos dados deve estar disponível e não somente uma parte deles.
2. **Primários:** Os dados são publicados na forma coletada na fonte e não de forma agregada ou transformada.
3. **Atualizados:** Os dados são disponibilizados o quão rapidamente seja necessário para preservar o seu valor.
4. **Acessíveis:** Os dados são disponibilizados para o público mais amplo possível e para os propósitos mais variados.
5. **Processáveis por máquina:** Os dados são razoavelmente estruturados para possibilitar o seu processamento automatizado.
6. **Não discriminatórios:** Os dados são disponibilizados para qualquer um, sem necessidade de registro.
7. **Não Proprietários:** Os dados são disponibilizados em um formato sobre o qual nenhum ente detenha controle exclusivo.
8. **Livre de Licença:** Os dados não estão sujeitos a regulações de direitos autorais, marcas, patentes ou segredo industrial.

A organização estadunidense *Sunlight Foundation*, focada em transparência governamental e política, adicionou mais dois princípios aos existentes (WONDERLICH, 2010):

9. **Permanentes:** Os dados devem ser preservados ao longo do tempo. Muitas vezes, as informações são atualizadas, alteradas ou removidas sem qualquer indicação de que uma alteração foi feita.

10. **Gratuitos:** Os dados são disponibilizados sem custos.

EAVES (2009) propôs as seguintes leis para Dados Abertos Governamentais, que são aplicáveis a Dados Abertos:

1. Se o dado não pode ser encontrado e indexado na *Web*, ele não existe.

O TRIBUNAL DE CONTAS DA UNIÃO (2015) ressalta que o simples fato de um dado estar disponível não significa necessariamente que seja aberto.

2. Se não estiver aberto e disponível em formato compreensível por máquina, ele não pode ser reaproveitado.

3. Se algum dispositivo legal não permitir sua replicação, ele não é útil.

SHADBOLT & O'HARA (2013) destacam que as prioridades dos governos são para aumentar a quantidade de Dados Abertos de qualquer forma. Para LÓSCIO; OLIVEIRA & BITTENCOURT (2015), não basta apenas publicar os dados na *Web*, é importante que os dados sejam publicados de maneira úteis para seus potenciais consumidores.

2.2.1 Dados Abertos Conectados

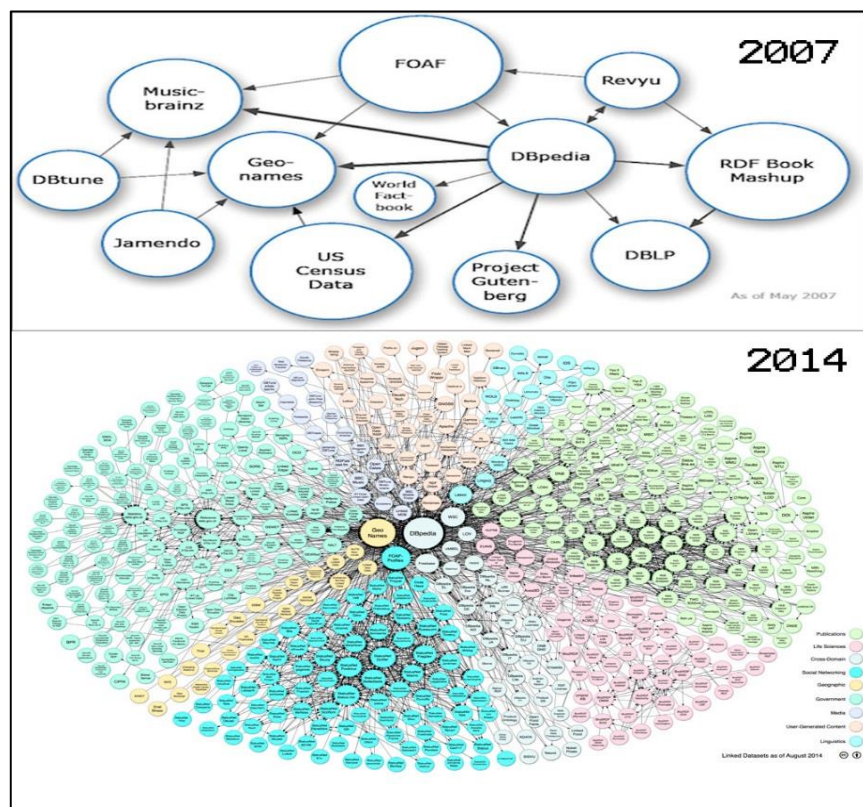
Os Dados Abertos se tornam conectados quando obedecem a um conjunto de boas práticas para serem publicados e conectados a outros conjuntos de dados estruturados na *Web*, criando uma “*Web de Dados*” (BIZER; HEATH; BERNERS-LEE, 2009). Berners-Lee definiu um conjunto de quatro princípios básicos, denominados de Princípios de *Linked Data* (Dados Conectados), os quais deverão ser obedecidos para garantir a criação de uma *Web de Dados*, são eles.

1. Utilizar URIs (*Uniform Resource Identifier*) como nome para “coisas”. Um Identificador Uniforme de Recurso (URI) é uma cadeia de caracteres usada para identificar um recurso.
2. Utilizar URIs HTTP para que as pessoas possam encontrar esses nomes.

3. Quando alguém procura por uma URI, garantir que informações úteis possam ser obtidas por meio dessa URI, as quais devem estar representadas no formato *Resource Description Framework* (RDF).
4. Incluir links para outras URIs de forma que outros recursos possam ser descobertos.

SHADBOLT & O'HARA (2013) ressaltam que a quantidade de Dados Conectados é muito pequena em relação aos Dados Abertos, e chegam a indagar se, mesmo assim, houve avanço em relação à temática. A resposta é perceptível ao visualizar o crescimento do diagrama (Figura 2.1) que representa a nuvem dos conjuntos de dados que foram publicados em formatos de Dados Conectados.

Figura 2.1 - Evolução da Nuvem de Dados Conectados.



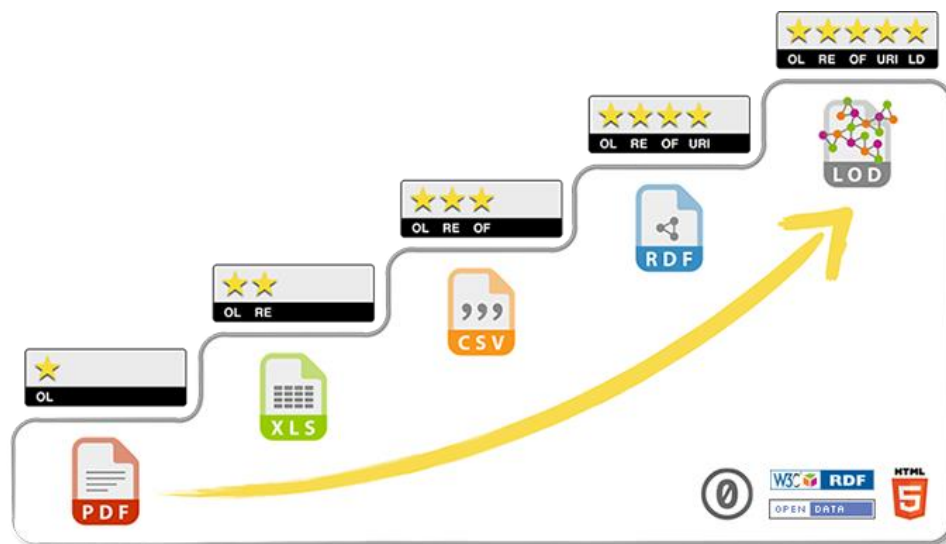
Fonte: Adaptado de <http://lod-cloud.net/>

Um maior quantitativo de Dados Abertos é reflexo da colocação do processo de abertura por diversos autores como o estágio preliminar a conexão dos dados. No entanto, ISOTANI & BITTENCOURT (2015) ressaltam que Dados Conectados não necessariamente precisam ser abertos, mas é uma tendência natural que o próximo passo após a abertura dos dados seja a conexão com outras fontes de dados, possibilitando a combinação entre eles, produzindo novos conhecimentos e aplicações (MACHADO; DE OLIVEIRA, 2011).

2.2.2 Esquema das 5 estrelas para Dados Abertos

Em 2010, Tim Berners-Lee apresentou um esquema de implementação das 5 estrelas para Dados Abertos (Figura 2.2). O sistema utiliza uma escala de 1 a 5 estrelas para medir a qualidade baseada na facilidade de processamento dos dados publicados, ou seja, dados classificados com uma estrela apresentarão uma maior dificuldade para serem processados, enquanto dados com cinco estrelas teriam uma facilidade no seu reúso, além de estarem conectados a outros dados. Basicamente, a escala é definida assim:

Figura 2.2 - Esquema das 5 estrelas para Dados Abertos.



Fonte: <http://5stardata.info/pt-BR/>

- ★ - Disponível na Web, em qualquer formato (por exemplo, uma imagem de um documento digitalizado);
- ★★ - Disponível em um formato estruturado, porém que depende de um software proprietário para a sua utilização (por exemplo, uma planilha do Excel ao invés de uma imagem digitalizada de uma tabela);
- ★★★ - Disponível em um formato estruturado e não proprietário (por exemplo, CSV ao invés de Excel);

★★★★- Usa padrões abertos recomendados pela W3C (RDF e SPARQL) para identificar recursos, de modo que as pessoas possam apontar para eles;

★★★★★- Todos os itens acima se aplicam, e também adicionar links para outras bases de dados relacionados.

ISOTANI & BITTENCOURT (2015) destacam que para o dado ser considerado aberto, é recomendado que o mesmo esteja classificado com no mínimo com três estrelas da escala. Os dados classificados como Dados Conectados deverão receber pelo menos quatro estrelas.

SHADBOLT & O'HARA (2013) ressaltam que o esquema proposto por Tim Berners-Lee é amplamente usado não apenas para classificação, e sim como um roteiro de transição de Dados Abertos para Dados Conectados.

2.2.3 Dados Abertos Governamentais

Os conceitos foram se estendendo em várias ramificações, estando em voga os Dados Abertos Governamentais, os quais preveem abertura de dados e informações da administração pública sem quaisquer restrições, facilitando a exploração pela sociedade, fomentando a interação com o cidadão, entre os órgãos e entre os governos (MACHADO; DE OLIVEIRA, 2011).

O fornecimento e a utilização de dados governamentais foram ampliados através de parcerias intergovernamentais, oficializadas em 2011 através da *Open Government Partnership* (OGP) (SAYOGO; PARDO; COOK, 2014). Vários países enxergaram a necessidade de disponibilizar dados governamentais em formato aberto, visto que a administração pública coleta, produz e dissemina um grande volume de informações valiosas, podendo se tornar matéria prima para outras aplicações, promovendo transparência, democracia, crescimento econômico e valor social aos cidadãos (KALAMPOKIS; TAMBOURIS; TARABANIS, 2011). Entre esses países, o ranking (Figura 2.3) do projeto “*Open Data Barometer*”, que mensura a prevalência e o impacto das iniciativas de dados abertos em todo o mundo, aponta o Reino Unido como o país mais desenvolvido em relação a temática, seguido dos Estados Unidos. O Brasil ocupa a 21ª posição no ranking que contém 86 países.

Figura 2.3 - Ranking do Open Data Barometer.

Country	Barometer Rank	ODB Scaled	Readiness (Scaled)	Implementation (Scaled)	Impact (Scaled)	2013 ODB	ODB Change	2013 Rank
UK	1	100	98	100	100	100	0	1
US	2	92.66	96	88	100	93.38	-0.72	2
Sweden	3	83.7	100	76	88	85.75	-2.05	3
France	4	80.21	91	75	84	63.92	16.29	10
New Zealand	4	80.01	81	88	55	74.34	5.67	4
Netherlands	6	75.79	95	76	57	63.66	12.13	10
Norway	7	74.59	88	73	64	71.86	2.73	5
Canada	7	74.52	90	75	58	65.87	8.65	8
Denmark	9	70.13	94	54	95	71.78	-1.65	5
Australia	10	68.33	92	69	43	67.68	0.65	7
Germany	10	67.63	85	67	53	65.01	2.62	9
Finland	12	66.49	93	54	78	49.44	17.05	14
Estonia	13	60.18	84	51	64	49.45	10.73	14
Spain	13	59.89	78	60	42	48.19	11.7	17
Chile	15	58.7	69	73	8	40.11	18.59	25
Austria	15	58.52	83	42	84	46.03	12.49	18
Czech Republic	17	58.07	64	61	46	43.18	14.89	22
Korea	17	57.65	79	54	48	54.21	3.44	12
Japan	19	53.58	81	53	30	49.17	4.41	14
Israel	20	52.97	70	51	43	45.58	7.39	18
Brazil	21	52.13	66	63	9	36.83	15.3	28

Fonte: <http://opendatabarometer.org/>. (2017)

No Brasil, ALCANTARA (2015) chama atenção para a existência de poucos casos de uso de Dados Abertos, existindo a necessidade de estimular e disseminar a cultura, inclusive no contexto educacional. Segundo ARAÚJO & SOUZA (2011), as esferas e os poderes nacionais oferecem uma boa oferta de dados, no entanto poucos em formatos estruturados e abertos, chamando a atenção para aumentar os esforços e as iniciativas em Dados Abertos no país.

BREITMAN et al (2012) já enxergavam no Brasil, os resultados oriundos dos esforços iniciados em 2009 pelos órgãos do Governo Federal e a Academia. Mesmo com muito trabalho a se fazer, de fato existe uma evolução do país referente à temática, perceptível na Figura 2.3 ao ganhar sete posições no Ranking em um espaço temporal de três anos. O avanço é consequência de eventos como: O País se tornar membro da *Open Government Partnership* (OGP) em 2011; A construção da Infraestrutura Nacional de Dados Abertos (INDA); A organização da CONSOCIAL, uma conferência nacional sobre Transparência e Controle Social (BRITO et al, 2015); A implementação de programas de Governo Eletrônico (e-Gov); A criação do Portal de Dados Abertos² e o Portal da Transparência³; O trabalho educacional

² <http://dados.gov.br/>

³ <http://www.portaldatransparencia.gov.br/>

desenvolvido pelo o escritório da W3C⁴ no Brasil (BREITMAN et al, 2012); A impetração de um arcabouço normativo direcionado à promoção da transparência e da participação social na gestão pública (TRIBUNAL DE CONTAS DA UNIÃO, 2015).

Tais eventos não estão limitados apenas à esfera federal, e cada vez mais vem permeando outras camadas de níveis estaduais e municipais. O levantamento dos portais de Dados Abertos (Tabela 2.1) realizado por OLIVEIRA et al (2016), revela portais em todas as esferas do governo, totalizando 23,520 (Vinte três mil e quinhentos vinte) conjuntos de dados disponibilizados na esfera federal, estadual e municipal. Em mais da metade dos casos são portais muito pequenos em relação aos quantitativos de conjuntos de dados disponibilizados. Mas cabe o destaque para o portal da cidade do Rio de Janeiro, apresentando um maior quantitativo de conjuntos de dados, inclusive em relação ao catalogo nacional. Tais ramificações independentes de serem *top-down* ou *bottom-up* estão expandindo e solidificando o movimento de Dados Abertos Governamentais no país. BRITO at al (2015) destacam que pelo tempo de existência dos repositórios nacionais, esperavam mais dados disponibilizados, principalmente Dados Governamentais. BRITO at al (2015) citam que os Estados Unidos iniciaram seu repositório com 47 (Quarenta e sete) conjuntos de dados, e com um ano de existência, contabilizavam 250,000 (Duzento e cinquenta mil) conjuntos de dados disponibilizados. Mesmo assim, estamos vivendo um movimento crescente. Tal assertiva é fundamentada se compararmos a Tabela 2.1 com os numeros de alguns repositórios elencados por BRITO at al (2015), incluindo Governo Federal (240), Recife (39) e Rio de Janeiro (46).

Tabela 2.1 - Quantitativos de conjuntos de dados dos portais de Dados Abertos nacionais.

Fonte de dados	Quantidade de conjuntos de dados
Alagoas	1,167
Curitiba	33
Distrito Federal	47
Fortaleza	642
Minas Gerais	40
Pernambuco	66
Porto Alegre	149
Recife	384
Rio de Janeiro (Cap.)	16,623
Rio Grande do Sul	162
São Paulo (Cap.)	37
São Paulo (Est.)	119
Governo Federal	4,051

Fonte: Adaptado de OLIVEIRA et al (2016)

⁴ <http://www.w3c.br/Home/WebHome>

2.2.4 Legislação para Dados Abertos no Brasil

O TRIBUNAL DE CONTAS DA UNIÃO (2015) cita que o Brasil vem desenvolvendo no decorrer dos anos, um arcabouço normativo para a temática. Uma evolução iniciada pela Constituição Federal de 1988, ao trazer no inciso XIV do artigo 5º a garantia a todos do acesso à informação, contrapondo-se a cultura do sigilo imposta pela Ditadura Militar (1964-1985) (BRASIL, 1988).

Em 2000, a Lei de Responsabilidade Fiscal (BRASIL, 2000) traz um capítulo tratando especificamente da transparência, do controle e da fiscalização da gestão fiscal. A referida lei foi alterada em 2009 para Lei da Transparência (BRASIL, 2009), determinando a disponibilização, em tempo real, de informações pormenorizadas sobre a execução orçamentária e financeira da União, dos Estados, do Distrito Federal e dos Municípios.

A Lei nº 12.527 conhecida como Lei de Acesso Informação (LAI), regulamenta o direito constitucional de acesso às informações públicas em todas as esferas do governo brasileiro (BRASIL, 2011). A lei foi sancionada em 18 de novembro de 2011, entrando em vigor em 16 de maio de 2012. A lei torna obrigatória a divulgação de todas as informações produzidas ou custodiadas pelo poder público através de portais oficiais da Internet, tendo como principal aspecto o acesso como regra e o sigilo como exceção. Citada por ISOTANI & BITTENCOURT (2015) como arcabouço jurídico para a obrigação governamental de disponibilizar Dados Abertos no país. Existem várias recomendações para a utilização de formatos abertos, estruturados e legíveis por máquinas. Porém, a LAI é omissa quanto à especificação de padrões e formatos, e como tais informações deverão estar disponibilizadas (MODESTO; FERNEDA & SANT'ANA, 2014). Os padrões e formatos estão definidos em outro documento complementar denominado ePING – Padrões de Interoperabilidade de Governo Eletrônico, aplicada apenas ao poder executivo da esfera federal, sem restrição à participação de maneira voluntária de outras esferas de governo.

Em 2016, através do Decreto nº8638 o País institui uma Política de Governança Digital no âmbito dos órgãos e das entidades da administração pública federal direta, autárquica e fundacional (BRASIL, 2016). A política elenca diversos objetivos estratégicos, incluindo fomentar a disponibilização e o uso de dados abertos. No mesmo ano, o decreto lei nº 8.777 institui a Política de Dados Abertos do Poder Executivo Federal, determinando a publicação de um Plano de Dados Abertos (PDA) pelos órgãos, no prazo de sessenta dias

(BRASIL, 2016). O PDA é documento orientador para as ações de implementação e promoção de abertura de dados de cada órgão.

2.3 Gerenciamento de Processos de Negócio

2.3.1 Processos de Negócio

O conceito completo de processo exposto pela Secretaria de Gestão do Ministério do Planejamento (SEGES) e adotado pela *Society for Design and Process Science* (SDPS) diz que um processo é “conjunto integrado e sincrônico de insumos, infraestruturas, regras e transformações, que adiciona valor às pessoas que fazem uso dos produtos e/ou serviços gerados” (MPOG, 2009). PINHEIRO (2010) traz uma visão mais simplificada do conceito, definindo processo (Figura 2.4) como “um conjunto estruturado de atividades elaborado para alcançar um determinado objetivo”. Um processo utiliza uma ou mais entradas, que após processadas entre as atividades, as transformam em saídas definidas. O CBOK (2013) define processo como uma agregação de atividades e comportamentos executados por humanos ou máquinas para alcançar um ou mais resultados. BALDAM (2007) enxerga como o encadeamento de atividades dentro da organização, que transformam entradas em saídas, podendo pertencer a diferentes contextos: Como processos jurídicos, processos químicos, processos de produção siderúrgica, processos de publicação de dados abertos, entre outros. Tais processos se caracterizam como um “processo de negócio” quando entrega valor aos usuários ou clientes.

Figura 2.4 - Elementos básicos de um processo.



Fonte: PINHEIRO (2010).

Os resultados poderão ser uma saída ou produto de uma atividade executada, um seguimento de um processo ou um serviço.

2.3.2 Conceitos de Gerenciamento de Processos de Negócio

BALDAM (2007) destaca que não bastava apenas criar os processos, e sim ter habilidades de manipulação e de mudança dos mesmos. Com esse pensamento, o Gerenciamento de Processos de Negócios (BPM, do inglês *Business Process Management*) surgiu como uma grande onda na virada do século.

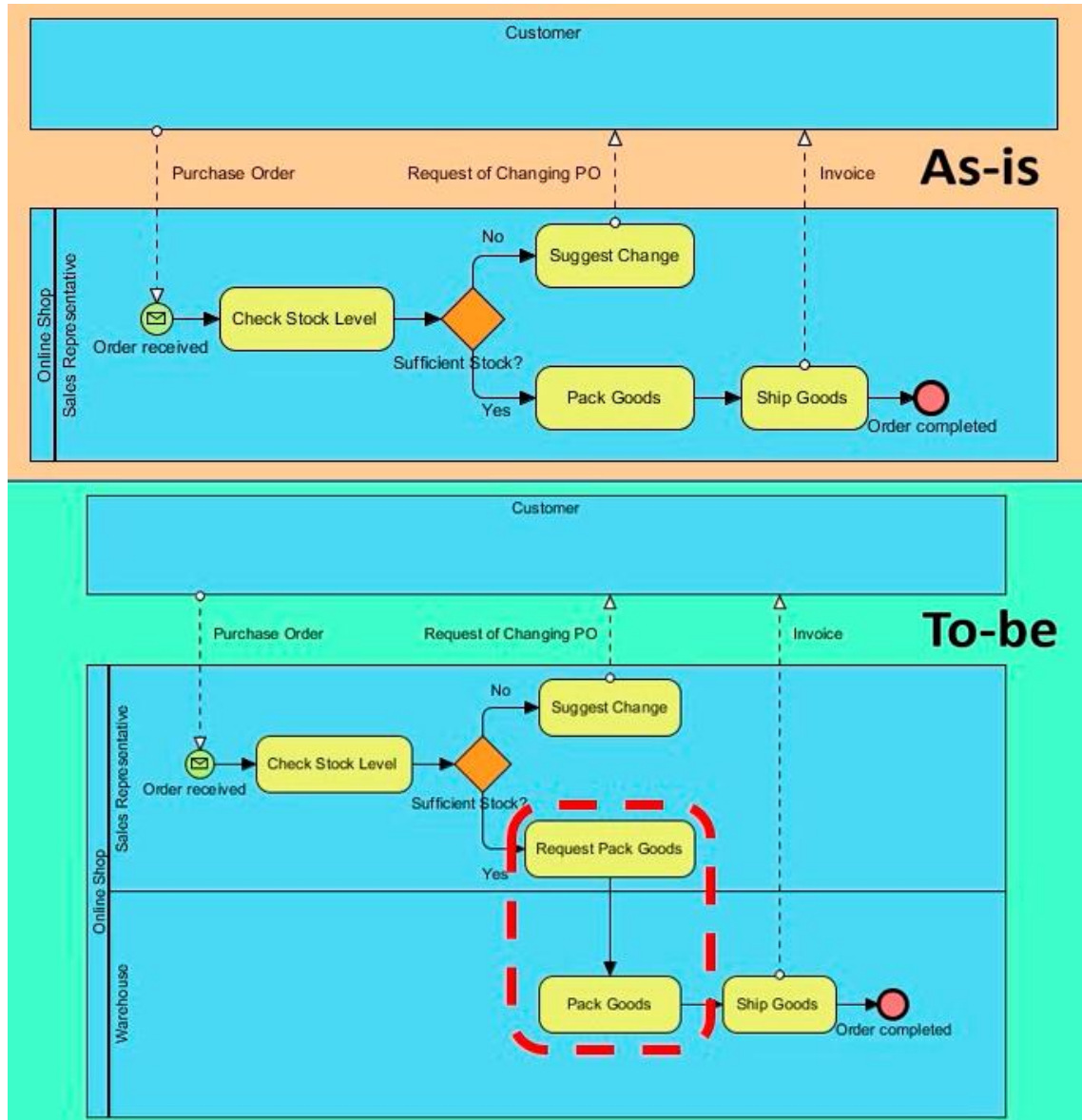
UNDERDAHL (2011) define BPM como uma abordagem desenvolvida para produzir melhores processos. Utilizando essa abordagem, o processo poderá ser modelado, automatizado, integrado, monitorado e otimizado continuamente. O BPM ajuda à organização oferecendo ferramentas que facilite a tomada de decisões, a administração, a operação, o controle e automatização de um determinado processo de forma simples e unificada. O CBOK (2013) diz que BPM é uma cultura que envolve uma nova forma de articular e aplicar de modo integrado abordagens, metodologias, estruturas de trabalho, técnicas e ferramentas para processos que muitas vezes são aplicadas de maneira isolada. O BPM se apresenta como um conceito que une processos e todas as habilidades de gestão dos mesmos, envolvendo a Tecnologia da Informação, ou seja, negócio, processos e tecnologia estão integrados totalmente. Uma estratégia BPM permite integrar os recursos tecnológicos e humanos para oferecer melhores resultados.

2.3.3 Modelagem de Processos

Na estratégia de BPM, os processos são considerados os ativos mais importantes da organização. Por isso, entender e modelar cada processo é um passo crucial para a estratégia de BPM. O entendimento acontece com o levantamento e documentação da situação atual do processo utilizado pela instituição, ou seja, descrevendo o fluxo de trabalho que realmente acontece no âmbito organizacional. Esta atividade é comumente chamada no CBOK (2013) de mapeamento de processos *AS IS*, o qual é representado em fluxo ou diagrama. Nesta mesma oportunidade, levantam-se também os problemas ou fragilidades, bem como as oportunidades de melhoria do processo. Após o mapeamento da situação atual, iniciar-se-á o trabalho de discussão, definição e documentação da situação futura do processo baseado nos problemas e fragilidades da situação atual juntamente com uma proposta de melhoria. O

processo gerado é comumente chamado no CBOK (2013) de *TO BE*, o qual é representada em fluxo ou diagrama – também chamado de modelagem (MARIANO, 2012).

Figura 2.5 - Processos As-is e To Be.



Fonte: <https://www.visual-paradigm.com>

Um processo de e-commerce é representado na figura 2.5. Na parte superior temos o processo denominado *AS IS*, ou seja, é comportamento atual de como a loja executa as atividades para atender uma ordem de compra pelo o site. Ao receber a ordem do cliente, o departamento de venda checa o nível do estoque referente ao produto. Existindo saldo, embala a encomenda e em seguida despacha para o cliente. O processo *TO BE* na parte inferior vem ser uma proposta incorporando melhorias a partir da análise da situação atual. Na exemplificação, a loja recebeu várias reclamações referentes ao serviço de entrega. Após

análise e discussão em cima do processo atual, a loja vislumbrou e modelou no processo *TO BE* a retirada do serviço de estocagem, embalagem e despacho do departamento de vendas. Quando a organização não tem os processos mapeados, a mesma pode iniciar com este trabalho de mapeamento (*AS IS*) e posteriormente incorporar melhorias modelando (*TO BE*) os processos. Existem organizações que não realizam o trabalho de mapeamento, preferindo estudar recomendações e melhores praticas de outras organizações, com intuito de modelar os processos *TO BE* e se adequar aos mesmos.

2.3.4 Notação para Modelagem de Processos

GOUVEIA (2011) elenca diversas técnicas que são utilizadas para a modelagem de processos, incluindo UML⁵ (*Unified Modeling Language*) e EPC⁶ (*Event-driven Process Chain*), no entanto aponta a BPMN⁷ (*Business Process Modeling Notation*) como a mais como padrão de modelagem.

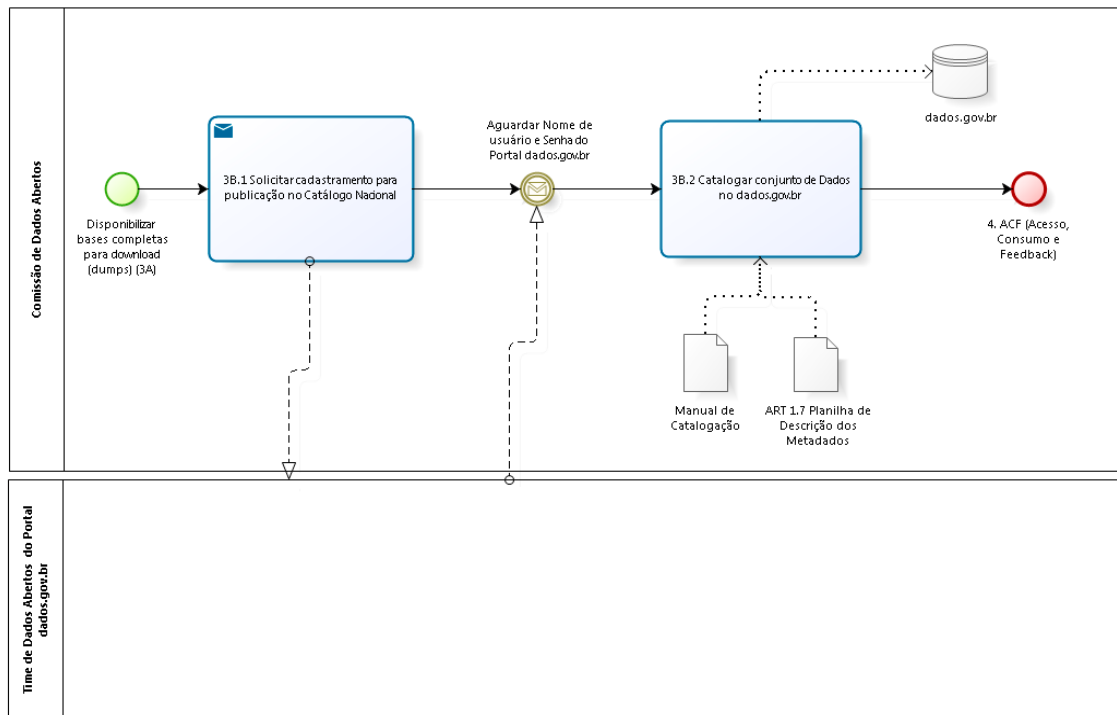
A BPMN foi desenvolvida pela *Business Process Management Initiative* (BPMI), posteriormente adotada como padrão e mantida pelo *Object Management Group* (OMG). A BPMN é uma notação gráfica (Figura 2.6) que descreve as etapas em um processo de negócios. BPMN descreve o fluxo de ponta a ponta de um processo de negócio. A notação foi especificamente projetada para coordenar a sequência de processos e as mensagens que fluem entre os diferentes participantes do processo em um conjunto relacionado de atividades. A Figura 2.6 exemplifica através da notação, o processo de catalogação no portal dados.gov.br, desde criação do cadastro até a sua referida interação com o portal.

⁵ <http://www.uml.org/>

⁶ <http://www.ariscommunity.com/event-driven-process-chain>

⁷ <http://www.bpmn.org/>

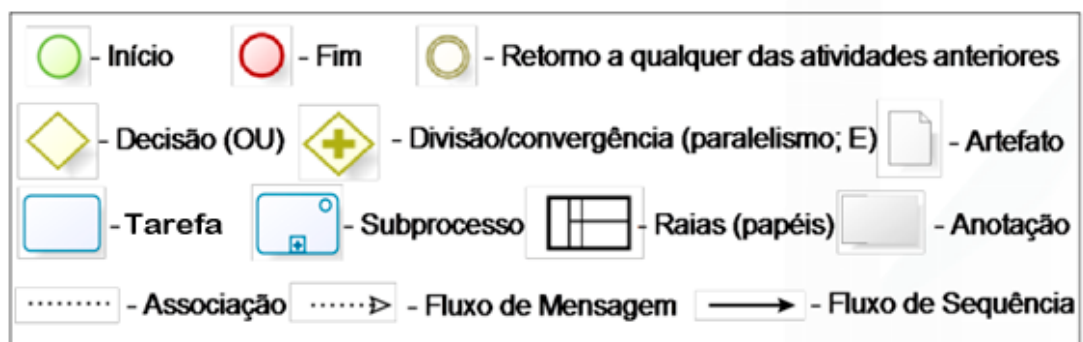
Figura 2.6 - Processo de catalogação no portal dados.gov.br representado com BPMN.



Fonte: Elaborado pelo autor.

Esta notação captura a lógica das atividades, mensagens entre vários participantes e todas as informações necessárias para um processo ser analisado, simulado e implementado. Entre suas características podemos destacar como vantagens: Ser uma linguagem unificada para representar processos; não ser uma metodologia; ser uma linguagem intuitiva e fornecer um entendimento geral dos processos, facilitando a comunicação entre as pessoas. A figura 2.7 apresenta uma breve descrição dos componentes básicos da BPMN utilizados nos diagramas do nosso modelo.

Figura 2.7 - Breve descrição dos elementos básicos da notação BPMN.

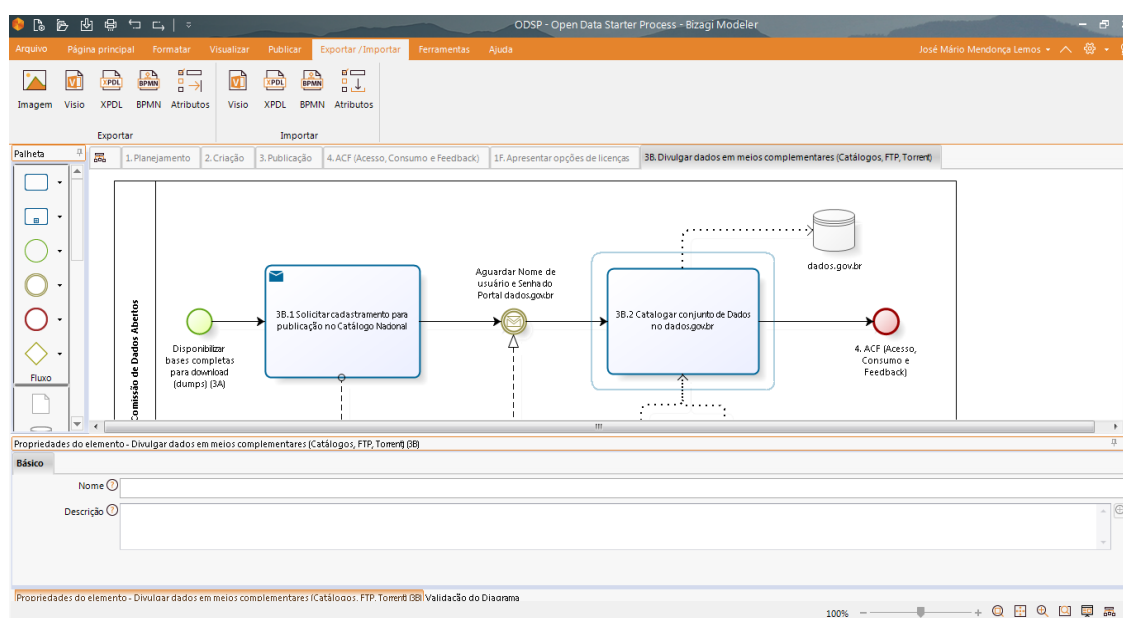


Fonte: Adaptado do Guia de Boas Práticas em Contratação de Soluções de Tecnologia da Informação do SLTI.

2.3.5 Business Process Management System (BPMS)

O *Business Process Management Suite* ou *System* (BPMS) é um conjunto de sistemas ou um sistema isolado que auxilia na automação de gestão dos processos incluindo funcionalidades para modelagem, execução, controle e monitoração. Existem diversas ferramentas BPMS livres e proprietárias como Bizagi⁸, Bonita⁹, Orquestra¹⁰, Lecom Atos¹¹ entre outras. Baseado na análise comparativa entre as ferramentas realizada pela SIX MINDS (2015), o Bizagi (Figura 2.8) desponta como a melhor ferramenta com 97 pontos, seguido do Orquestra com 95 pontos, Bonita com 94 pontos e Lecon Atos com 89 pontos.

Figura 2.8 - BPMS Bizagi versão Modeler.



Fonte: Elaborado pelo autor.

2.4 Considerações finais

As temáticas apresentadas neste capítulo se resumiram nos conceitos fundamentais do desenvolvimento até o resultado final da pesquisa. O decorrimento do capítulo obedeceu uma sequência lógica e gradativa na explicitação de cada conceito. A seção inicial traz a diferenciação de “Dados”, “Informação” e “Conhecimento”, nos quais são conceitos que lidamos o tempo todo, mas de certa forma são empregados erroneamente em alguns

⁸ <http://www.bizagi.com/pt>

⁹ <http://www.bonitasoft.com/>

¹⁰ <http://orquestrabpm.com.br/>

¹¹ <http://www.lecom.com.br/>

contextos. Após os conceitos bases, foram percorridos os conceitos relacionados ao objeto de estudo da pesquisa, incluindo Dados Abertos, Dados Abertos Conectados e Dados Abertos Governamentais. Com isso, tivemos o entendimento de que os dados se tornam abertos quando obedecem a um conjunto de princípios e leis, não bastando apenas disponibilizar os mesmos de maneira pública. Os princípios incluem que os dados sejam completos, primários, atualizados, acessíveis, processáveis por máquina, livre de licença, permanentes e gratuitos. Além disso, não devem ser discriminatórios e proprietários. Quando a origem dos Dados Abertos é oriunda das instituições governamentais, os mesmos são rotulados como “Dados Abertos Governamentais”. Após de abertos, a tendência natural é conectar os dados a outros conjuntos de dados, possibilitando novos conhecimentos e aplicações. O capítulo não estava restrito apenas a explicitação de conceitos, trazendo consigo uma visão do panorama atual frente a temática, considerando também a evolução histórica. Com isso, percebemos a evolução tanto para Dados Abertos quanto para Dados Conectados no cenário nacional e internacional, com vantagem nos números das iniciativas em Dados Abertos. As iniciativas que proporcionaram a evolução do Brasil frente a temática são evidenciadas no capítulo, com destaque para o mapeamento dos momentos históricos do arcabouço jurídico. Como o produto gerado a partir desse trabalho de pesquisa é um processo, também foi apresentado os principais conceitos do Gerenciamento de Processos de Negócios (BPM, do inglês *Business Process Management*). Existem diversas notações e ferramentas (BPMS) para modelar processos, a notação mais adotada é a BPMN e a ferramenta melhor pontuada é o Bizagi.

3 Análise da Transparência Ativa e Passiva na Rede Federal de Educação Profissional, Científica e Tecnológica

Nesta subdivisão do trabalho, mostra-se os resultados da análise da maturidade da Rede Federal de Educação Profissional, Científica e Tecnológica na gestão e publicação de dados. O tópico 3.1 discorre sobre o histórico e a contextualização da Rede Federal. Em seguida, na seção 3.2, são apresentados os conceitos de transparência ativa e passiva. Os processos metodológicos utilizados para as análises (transparência ativa e passiva) na Rede Federal são descritos no segmento seguinte, o 3.3. Por fim, nos itens 3.4 e 3.5, são expostos e discutidos os resultados das referidas análises.

3.1 A Rede Federal de Educação Profissional, Científica e Tecnológica

Desde o surgimento da Lei 11.892 de 29 de dezembro de 2008 que institui a Rede Federal de Educação Profissional, Científica e Tecnológica, temos um novo domínio (Figura 3.1) afluindo na educação brasileira. Em relação aos eixos educação profissional, ciência e tecnologia, a expansão histórica dos números de unidades atinge todas as unidades federativas caracterizando uma rede federal que oferece ensino médio convencional, ensino médio profissionalizante, ensino superior de graduação e pós-graduação. São 38 Institutos Federais subdivididos em 562 unidades que integram a rede, e 29 unidades conveniadas subdivididas em 2 Cefets, 25 escolas vinculadas a Universidades, o Colégio Pedro II e uma Universidade Tecnológica.

Figura 3.1 - Instituições da Rede Federal.



Fonte: <http://redefederal.mec.gov.br/>.

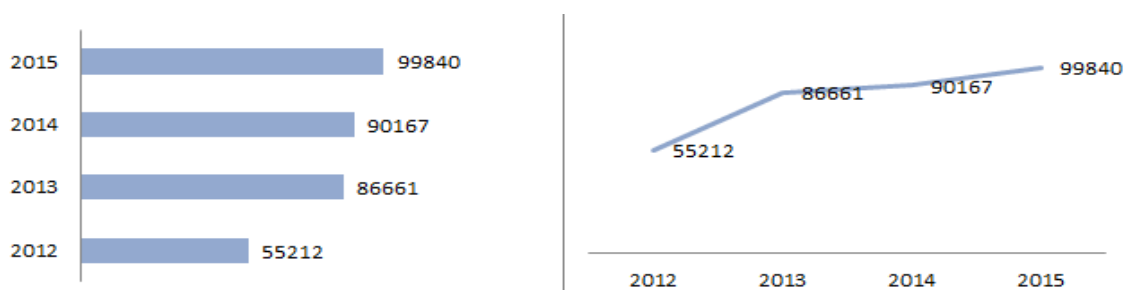
3.2 Transparência Ativa e Passiva

A ação das instituições públicas de divulgarem de maneira espontânea as informações produzidas ou custodiadas por elas, se caracteriza como um ato de transparência ativa (PEDROSO; TANAKA; CAPPELLI, 2013). De acordo com a LAI e seu respectivo regulamento (BRASIL, 2012), é “dever dos órgãos e entidades promover, independente de requerimento, a divulgação em seus sítios na Internet de informações de interesse coletivo ou geral por eles produzidas ou custodiadas”.

Idealmente, todos os dados produzidos e custodiados por organizações públicas deveriam estar disponíveis de maneira aberta e conectada, mas raramente é possível ter acesso a esse conjunto de dados ideal, devido a diversos fatores como: A cultura de sigilo das instituições; a falta de habilidades dos servidores em trabalhar com os dados; a ausência de padronização, estruturação e conexão dos dados. Como não existe obrigatoriedade na abertura de todo o conjunto de dados, as instituições geralmente disponibilizam em seus portais o rol mínimo, conforme dispõe o § 3º do Art. 7º, do Decreto 7.724/2012.

A transparência passiva se caracteriza pela capacidade dos órgãos públicos em fornecer informações que são solicitadas pelo cidadão. Devido à grande dificuldade de se encontrar dados ou informações nos portais de dados, os quantitativos de pedidos de usuários reivindicando informações através do e-SIC¹² (Sistema Eletrônico do Serviço de Informação ao Cidadão) apresentam um crescimento ao longo dos anos. A Figura 3.2 mostra este crescimento para todos os órgãos públicos, enquanto a Figura 3.3 traz os resultados para os órgãos do nosso estudo.

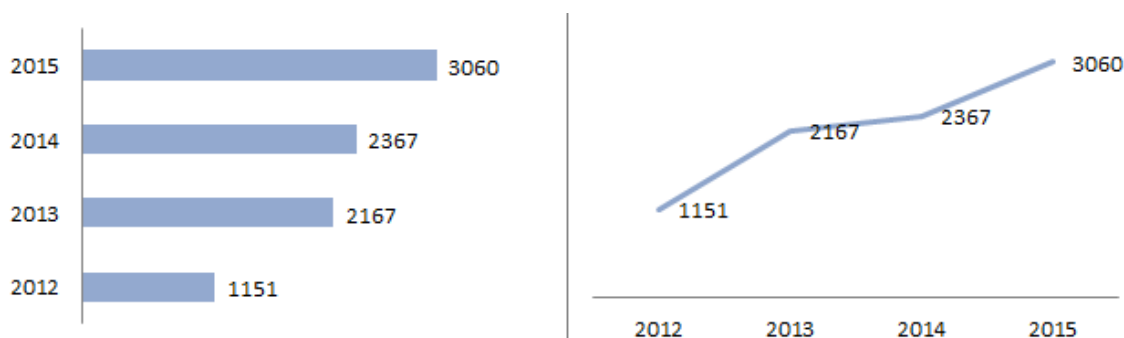
Figura 3.2 - Quantitativo de pedidos de acesso à informação de todos os órgãos públicos no e-SIC.



Fonte: Elaborado pelo autor.

¹² <https://esic.cgu.gov.br/sistema/site/index.html>

Figura 3.3 - Quantitativo de Pedidos de acesso à informação de todos Institutos Federais no e-SIC.



Fonte: Elaborado pelo autor.

Os resultados permitem algumas reflexões, como: Caso as instituições estivessem aplicando a LAI em conjunto com os princípios de Dados Abertos, os números de requerimentos seriam os mesmos? Haveria um decréscimo anual nos requerimentos? O que queremos ressaltar é que o órgão tem obrigação de fazer uma transparência ativa, publicando e divulgando, de maneira efetiva e proativa, sem requerimentos (MEDEIROS; MAGALHÃES; PEREIRA, 2013).

3.3 Metodologia

Esta seção tem como objetivo apresentar a metodologia empregada nas análises de transparência ativa e passiva da Rede Federal de Educação Profissional, Científica e Tecnológica.

3.3.1 Metodologia para análise da transparência ativa

A amostragem para a referida análise foi de 38 (trinta e oito) portais pertencentes aos Institutos Federais de Educação, Ciência e Tecnologia, contemplando todas as unidades federativas. Foram levantadas todas as informações disponibilizadas apenas na seção de Acesso à Informação e demais páginas dos Portais da Reitoria. O escopo foi centrado no conteúdo mínimo e obrigatório da LAI, incluindo informações sobre competências, estrutura organizacional, localização, horário de atendimento, repasses financeiros, registro de despesas, licitações, programas, ações e projetos. Utilizamos como base o Indicador D3 (Proporção de Órgãos Públicos Federais e Estaduais, por formato de arquivo) da pesquisa sobre o uso das Tecnologias da Informação e Comunicação no setor público brasileiro (CGI, 2013). Em relação ao indicador D3, a forma de coleta na pesquisa TIC Governo Eletrônico 2013 aconteceu por meio de questionário direcionado apenas ao responsável pelos conteúdos

digitais dos órgãos. Nesse caso, os entrevistados foram questionados quanto à disponibilidade de arquivos para *download* em seus formatos. Em nossa pesquisa, a forma de coleta aconteceu por meio de acesso aos portais, observando em qual formato (PDF, CSV, XLS OU XLSX, ODT, DOC OU DOCX, RDF, XML, PPT ou PPTX, OUTRO) a informação estava publicada. O autor acessava o site (Figura 3.4) de acordo com o escopo e visualizava se a determinada informação estava disponível, registrando por meio de planilha eletrônica em qual formato a informação observada estava publicada.

Figura 3.4 - Forma de coleta na análise da transparência ativa



Fonte: Elaborado pelo autor.

Para a análise da transparência ativa, o autor definiu 05 (cinco) critérios, os quais foram estabelecidos de acordo com o Sistema de 5 Estrelas dos Dados Abertos proposto por Tim Berners-Lee e apresentado no capítulo anterior. Os critérios definidos estão descritos a seguir:

1º Critério: Ao clicar na sessão correspondente a informação pesquisada, as informações cujos links se encontram inexistentes ou desabilitados, serão classificadas na análise dos dados com nenhuma estrela.

2º Critério: Será considerado 1 estrela, todos os dados disponibilizados de forma on-line com uma licença aberta e apresentando os seguintes formatos: PDF, PPT/PPTX, DOC/DOCX, ODT e ZIP.

3º Critério: Convencionamos considerar com 1 estrela, os links que encaminham o usuário para outra página externa. Desde que a página externa contenha as informações relacionadas ao aspecto analisado.

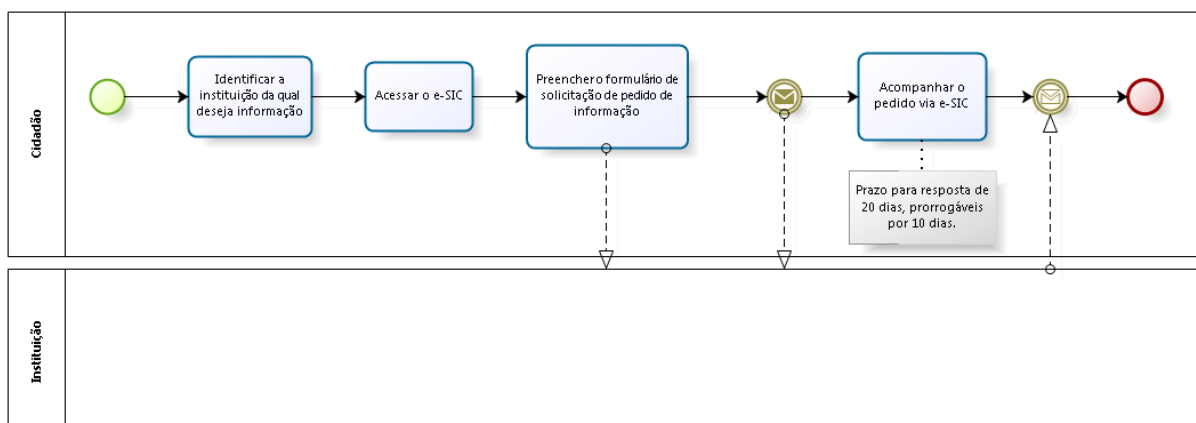
4º Critério: Todos os formatos que não estejam elencados na classificação do Sistema de 5 Estrelas dos Dados Abertos, serão classificados como “outro”. Um outro formato será classificado com 1 estrela, tendo o pressuposto de que independente de formato, é melhor ter a informação disponibilizada do que não ter.

5º Critério: Todos os outros formatos encontrados obedecerão normalmente a sequência da classificação do Sistema de 5 Estrelas dos Dados Abertos: 2 estrelas para formatos XLS ou XLSX; 3 estrelas para formatos CSV, JSON, ODS e XML.

3.3.2 Metodologia para análise da transparência passiva

A situação atual da transparência ativa na qual o órgão se enquadra não pode se tornar uma assertiva limítrofe da capacidade do órgão no fornecimento de informações. Sabemos que diferentes instituições estarão em níveis de maturidade distintos, mas a cultura de sigilo praticada pelos governos fornece falsos positivos da capacidade real de cada órgão (PEDROSO; TANAKA; CAPPELLI, 2013). Para analisar a capacidade real no fornecimento de dados, utilizamos o mecanismo de solicitação de informações proporcionado pela LAI ao cidadão. Especificamente, utilizamos o Sistema Eletrônico do Serviço de Informações ao Cidadão (e-SIC), encaminhando pedidos de informações de acesso à informação de acordo com o fluxo do processo elencado na Figura 3.5.

Figura 3.5 - Processo de pedido de informação via e-SIC.



Fonte: Elaborado pelo autor.

Formalizamos via sistema a solicitação dos pedidos de informação para 38 (trinta e oito) Institutos Federais de Educação, Ciência e Tecnologia, contemplando a amostragem citada na análise da transparência ativa. Os pedidos foram estruturados de acordo com as perguntas elencadas abaixo. As perguntas foram elaboradas com base no conteúdo mínimo obrigatório (competências, estrutura organizacional, localização, horário de atendimento, repasses financeiros, registro de despesas, licitações, programas, ações e projetos) previsto na LAI. Ressaltamos que, preferencialmente, as respostas fossem enviadas em um formato estruturado e aberto como CSV, JSON, ODS ou XML.

1. Quais competências chaves são desenvolvidas pelo instituto?
2. Como é formada a hierarquia do instituto?
3. Quais os telefones, endereços e e-mails das principais unidades de negócios do instituto?
4. Qual o horário de funcionamento e atendimento ao público das principais unidades de negócios do instituto?
5. Quais Programas, Projetos e Ações foram executados pelo instituto no exercício corrente e em exercícios anteriores?
6. Quem são os responsáveis por cada Programa, Projeto e Ação?
7. Quando o instituto recebeu receitas e quais são os valores recebidos?
8. Qual o valor liquidado pelo instituto no mês corrente? Qual o valor liquidado no ano corrente? Qual o valor liquidado nos exercícios encerrados?
9. Qual o valor pago pelo Instituto no mês corrente? Qual o valor pago no ano corrente? Qual o valor pago nos exercícios encerrados?

10. Quais são as despesas do Instituto com diárias e passagens pagas a servidores públicos?
11. Quais as informações (modalidade, objeto, situação etc.) das licitações realizadas em todos os exercícios pelo instituto?
12. Quais as informações (modalidade, objeto, situação etc.) das licitações em andamento do instituto?
13. Quantos e quais são os contratos firmados em todos os exercícios pelo instituto?

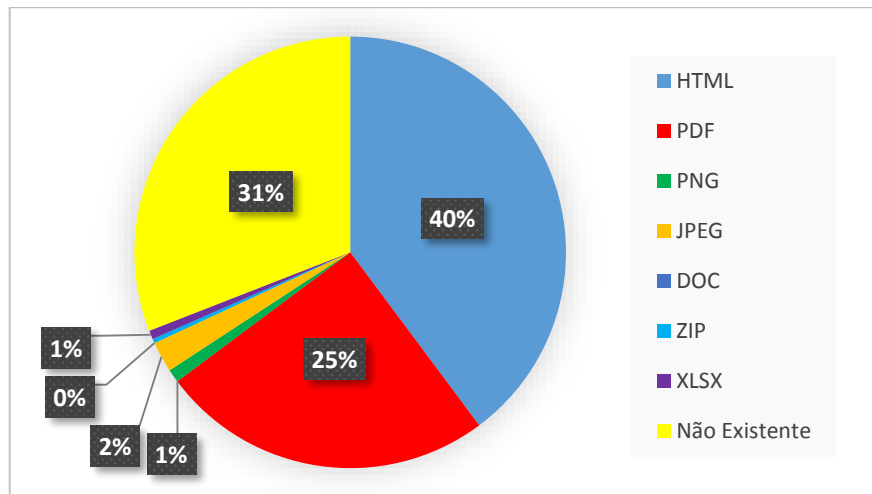
Após o envio, o acompanhamento das solicitações aconteceu pelo próprio sistema E-sic e por meio de planilha eletrônica considerando: Prazo, Conteúdo e o Formato de Arquivo da Resposta.

3.4 Resultados

Esta seção discutirá os resultados das análises de transparência da Rede Federal. Inicialmente serão expostos os resultados para transparência ativa, sempre cruzando tais indicadores com a literatura, no intuito de validar, discutir ou gerar novas observações a partir dessa intersecção. Para transparência passiva, expomos os resultados trazendo para o texto, fatos acometidos pelas instituições durante o processo. Por fim, é realizada uma análise geral, destacando os principais achados e a importância dessa análise para o trabalho.

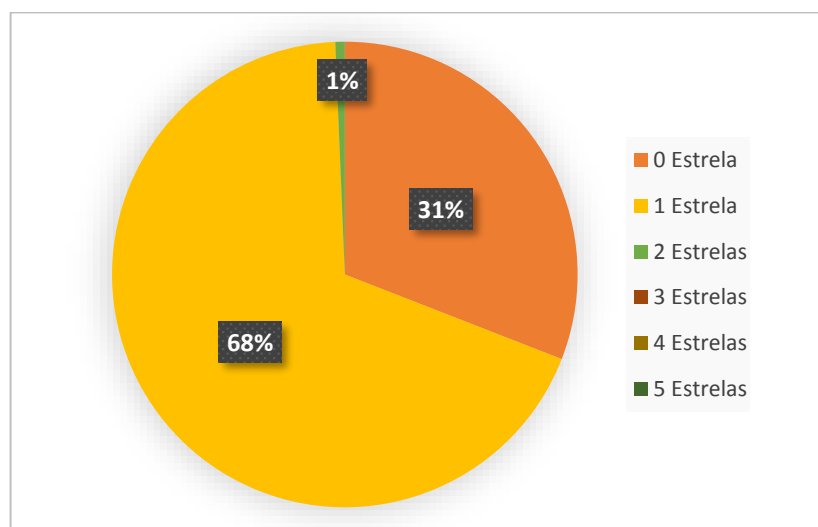
3.4.1 Resultados da análise da transparência ativa

Os resultados (Figura 3.6 e Figura 3.7) enquadram os Institutos Federais como imaturas, considerando a escala do esquema de 5 estrelas para Dados Abertos. A figura 3.6 reflete a grande dificuldade de encontrar parte das informações nos portais, chegando a uma taxa de 31% de indisponibilidade. As informações não foram encontradas por estarem realmente indisponíveis, ou pela dificuldade imposta em virtude da não obediência aos princípios de publicação de dados abertos, e também porque cada portal apresentava estrutura distinta.

Figura 3.6 - Transparência Ativa da Rede Federal por formato.

Fonte: Elaborado pelo autor.

Os números revelados na figura 3.7 mostraram que 68% dos conteúdos publicados na Rede Federal estão classificados com 1 estrela, apresentando formatos semiestruturados e não estruturados. Os conteúdos classificados por uma única estrela apresentaram uma predominância dos formatos HTML (40%) e PDF (25%). Uma similaridade com esse resultado é encontrada na maioria dos órgãos do governo brasileiro, os quais apresentam um índice equivalente a 80% de conteúdos publicados no formato HTML em seus portais (ARAÚJO; DE SOUZA, 2011; MACHADO; DE OLIVEIRA, 2011). Delimitando o escopo para as páginas de Acesso à Informação dos órgãos, a maioria apresenta as informações em formato PDF (PEDROSO; TANAKA; CAPPELLI, 2013).

Figura 3.7 - Transparência Ativa da Rede Federal de acordo com esquema de 5 estrelas.

Fonte: Elaborado pelo autor.

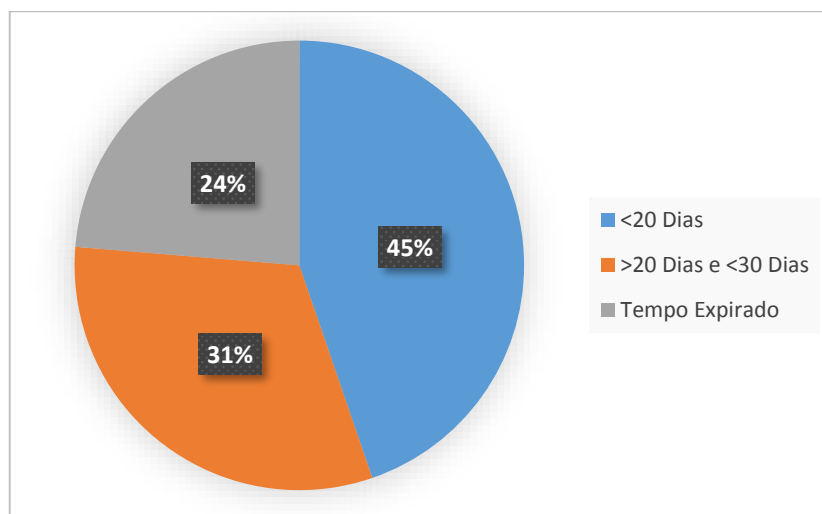
O maior nível na escala alcançado pelas instituições foi de apenas 2 estrelas, representando 1% do conteúdo apresentado em um formato estruturado e proprietário. Um reposicionamento dessas instituições para alcançar o nível máximo (5 Estrelas) exigirá um custo alto em comparação ao custo de se alcançar o nível mínimo recomendado (3 Estrelas) (ISOTANI; BITTENCOURT, 2015). Mas um primeiro passo é pensar em qual formato o dado será publicado, visto que o formato dos dados tem influência na escala do esquema de implementação das 5 estrelas para Dados Abertos (SAYOGO; PARDO; COOK, 2014). Uma solução simples seria a exportação dos dados em um formato estruturado e não proprietário como o CSV, facilitando a redistribuição e o reúso. Idealmente, espera-se o fornecimento de semântica aos dados com base em uma ontologia, juntamente com a publicação obedecendo aos princípios dos Dados Conectados (HOXHA; BRAHAJ, 2011).

3.4.2 Resultados da análise da transparência passiva

Mesmo se tratando de uma análise da transparência passiva, muitas instituições entenderam o questionamento como uma forma de avaliação da atuação do Serviço de Informação ao Cidadão (SIC) no que diz respeito à transparência ativa, justificando que é um assunto que está sendo estudado e aperfeiçoado nos institutos. Contudo, tal entendimento por parte das instituições, foi reflexo das perguntas submetidas apresentarem uma relação com a disponibilização em seus portais do rol mínimo conforme dispõe o § 3º do Art. 7º, do Decreto 7.724/2012. Mas, independente de analisar a atividade ou reatividade, o objetivo central é entender a gestão dos dados nas referidas instituições utilizando ambas as análises.

Os resultados apresentados nas figuras 3.8 e 3.9 demonstram a falta de gestão de dados dos Institutos Federais, em virtude do espaço temporal para responder questões sobre dados que já deveriam estar disponibilizados em seus respectivos portais. Além disso, cada instituto apresenta estruturas de dados distintas e metodologias distintas para cada resposta. Mesmo sendo participante de uma rede (Rede Federal) com características similares, tais instituições não apresentam um padrão para o mesmo processo.

Figura 3.8 - Reatividade das instituições em relação ao espaço temporal para respostas.



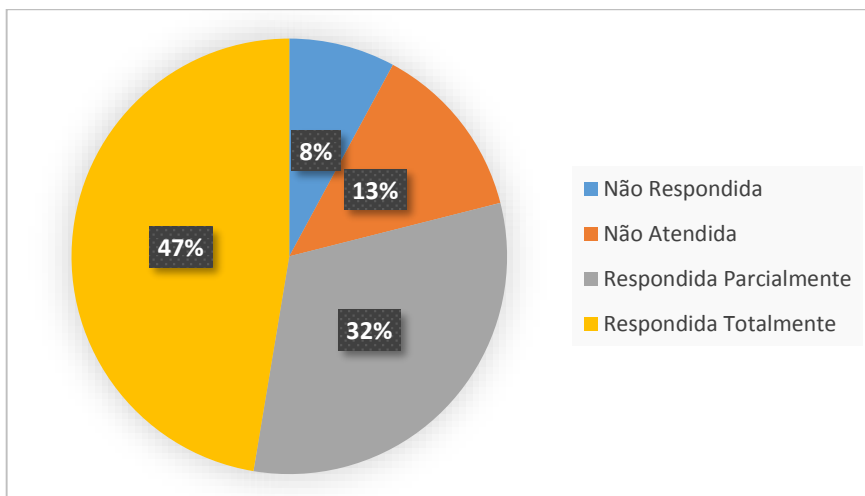
Fonte: Elaborado pelo autor.

Como apresentado na figura 3.8, 31% das instituições levaram mais de 20 dias para apresentar uma resposta, enquanto 24% tiveram o tempo expirado. Além disso, como mostrado na figura 3.9, algumas instituições ainda apresentaram respostas de maneira parcial (32%), e no pior dos casos, nem responderam (8%). As justificativas para as abstenções são inúmeras, incluindo que o Serviço de Informação ao Cidadão encontra-se ainda em implantação; ou estão passando por uma reformulação dos portais para publicação das informações solicitadas de forma mais clara. Outra justificativa apresentada diz respeito ao volume de dados solicitado, ou seja, o pedido apresentava uma dimensão que inviabilizava o trabalho da unidade responsável pela produção de resposta. Citamos o exemplo do Instituto Federal Catarinense:

“O pedido é desproporcional, pois demanda um trabalho de busca e compilação de dados de todas as unidades que compõem o Instituto Federal Catarinense, num período de 7 anos, o que demandaria um trabalho além das atividades rotineiras das unidades, de aproximadamente 30 a 60 dias. ”

Além disso, uma instituição ressaltou que o e-SIC destina-se, por sua natureza legal, à obtenção de informação pontual e específica e não se destinava à efetivação de pesquisa histórica ou acadêmica. No entanto, uma instituição que se justificou em relação ao não atendimento, disponibilizou acesso ao sistema corporativo para consulta direta das informações, bastando solicitar agendamento no setor desejado pelo próprio e-SIC.

Figura 3.9 - Transparência Passiva em relação ao conteúdo.

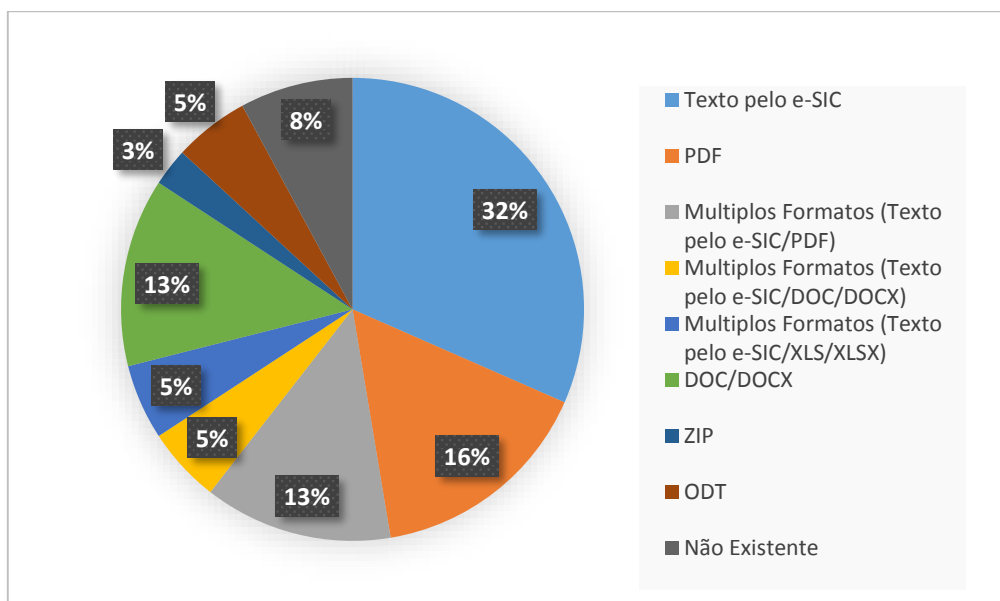


Fonte: Elaborado pelo autor.

Como apresentado na figura 3.9, 13% das solicitações não foram atendidas em relação aos conteúdos, mesmo as instituições atendendo dentro do prazo. Em vários casos, as instituições informavam que a solicitação se tratava de informações já disponíveis em seus portais. Mas as referidas informações não eram encontradas, dificuldade criada pelo fato da instituição não indicar a localização exata das informações em sua página.

A figura 3.10 está relacionada aos formatos de resposta, mesmo sendo solicitado pelo pesquisador, nenhuma das instituições atendeu a solicitação de responder em um formato estruturado e aberto como CSV, JSON, ODS ou XML. Os formatos semiestruturados e não estruturados englobam 95% das respostas, predominando a utilização de texto via o sistema e-SIC e o envio de arquivo em PDF. Algumas instituições justificaram o envio no formato PDF para garantir a fidedignidade das informações.

Figura 3.10 - Transparência Passiva em relação ao formato.



Fonte: Elaborado pelo autor.

3.5 Considerações Finais

Os resultados apresentados neste capítulo ressaltam o não envolvimento dos Institutos Federais com as práticas para a publicação de Dados Abertos. As análises de transparência ativa e passiva do capítulo atual ratificaram tais resultados. Além de, proporcionar uma determinada visão do processo de gestão e publicação de dados, inclusive abertos. Os resultados demonstraram que as instituições apresentam muitos problemas e insuficiências em relação à temática. Esse fato demonstra que ainda falta um longo caminho a percorrer pelos Institutos Federais, justificando a necessidade de se ter uma proposta para guiar a abertura de dados. Os principais achados das análises de transparência ativa e passiva foram:

- A confirmação do não envolvimento dos Institutos Federais com Dados Abertos;
- O entendimento de que o atendimento a instrumentos regulatórios não garante necessariamente a qualidade das informações;
- O conhecimento dos formatos utilizados e a predominância dos formatos semiestruturados e não estruturados em ambos os tipos de transparência;
- O despreparo e a má vontade da maioria das equipes no fornecimento de informações, além da desorganização das instituições na gestão dos dados.

4 Análise dos processos para publicação de Dados Abertos Conectados

Este capítulo apresenta uma análise de trabalhos relacionados à nossa proposta, a qual foi utilizada como base para a criação do processo proposto. LUNA (2011) reitera que é natural adotar processos guiados pelo o que há de melhor no mercado com objetivo de amenizar erros e atribulações já vivenciadas e superadas por outras instituições. CARNEIRO (2010) explana que é preciso identificar as recomendações de organizações similares, estudar essas recomendações e verificar a possibilidade de adoção das mesmas. Com este intuito, ÁVILA (2015) realizou uma análise de 15 processos para publicação de Dados Abertos Conectados, correlacionando as 70 recomendações encontradas com as melhores práticas para publicação de Dados Conectados propostas pela W3C. Replicamos a metodologia desta análise adicionando um quantitativo de 07 processos aos 15 já analisados por Ávila (2015). A análise com um conjunto maior de processos de publicação já era vislumbrada pelo próprio autor como oportunidade de pesquisa.

4.1 Critérios de seleção dos processos para publicação de Dados Abertos Conectados

ÁVILA (2015) utilizou diversos critérios para selecionar os processos, como a definição de que os mesmos deveriam ter como objetivo a publicação de dados abertos governamentais. Além disso, ÁVILA (2015) considerou a questão geográfica como critério, estabelecendo que os processos deveriam pertencer a um contexto socioeconômico e geopolítico similar ao do processo proposto pelo pesquisador, que, nesse caso, foram processos oriundos do Brasil e de Países da América do Sul. Para não trazer uma visão limitada para o processo em relação ao que está acontecendo no cenário mundial, ÁVILA (2015) adicionou artigos científicos e documentos técnicos de países europeus.

Para nossa análise, incorporamos os processos analisados por ÁVILA (2015) e acrescentamos novos processos não contemplados por ele, sem nos limitar a nenhum recorte geográfico. Os novos processos foram encontrados através de levantamento bibliográfico nas bases científicas (ACM Digital Library¹³, Google Scholar¹⁴, IEEE Xplore Digital Library¹⁵ e Portal de Periódicos CAPES¹⁶). Além disso, não foram consideradas apenas busca por fontes

¹³ <http://dl.acm.org/>

¹⁴ <https://scholar.google.com.br/>

¹⁵ <http://ieeexplore.ieee.org/Xplore/home.jsp>

¹⁶ www.periodicos.capes.gov.br

científicas (artigos, teses e dissertações), mas também por literatura cinza nos repositórios governamentais. Os seguintes critérios foram considerados: o processo deveria apresentar características de um guia, e que, mesmo utilizando qualquer tipo de representação (Literal ou Gráfica), deveria visar uma abordagem de passo a passo ou a criação de fluxos de trabalho, servindo de roteiro para quem almeja disponibilizar seus conjuntos de dados de forma aberta e conectada. Visto que o processo de saída da pesquisa é focado em um contexto educacional, poderíamos ter um filtro relevante na seleção dos processos, com intuito de privilegiar processos de publicação de dados educacionais. No entanto este filtro tornou a busca deserta, ou seja, não foi encontrado nenhum processo específico para este contexto. Desta maneira, foram estabelecidos os processos de publicação de dados abertos expostos no Quadro 4.1 para servirem de objeto de estudo.

Quadro 4.1- Processos de publicação de Dados Abertos analisados.

Código	Origem	Documentos (processos) Avaliados	Tipo
P1	Brasil	Manual para Elaboração de Plano de Dados Abertos; Plano de Dados Abertos - Ministério do Planejamento, Orçamento e Gestão do Brasil; Kit de Dados Abertos; Cartilha Técnica para Publicação de Dados Abertos no Brasil v1.0 (BRASIL, 2014)	Dados Abertos Governamentais
P2	Chile	Norma Técnica para Publicación de Datos Abiertos en Chile (CHILE, 2013).	Dados Abertos Governamentais
P3	Colômbia	Guía para la Apertura de Datos em Colombia (COLOMBIA, 2012)	Dados Abertos Governamentais
P4	Equador	Guía de Política Pública de Datos Abiertos – Ecuador (ECUADOR, 2014).	Dados Abertos Governamentais
P5	Uruguai	Guía rápida de publicación em datos.gub.uy (URUGUAY, 2012)	Dados Abertos Governamentais
P6	Itália	Geolinked Open Data for the Municipality of Catania (CONSOLI et al., 2014)	Dados Abertos Conectados Governamentais

Código	Origem	Documentos (processos) Avaliados	Tipo
P7	Internacional	LOP – Capturing and Linking Open Provenance on LOD Cycle (MENDONÇA et al., 2013)	Dados Abertos Conectados
P8	Internacional	TWC LOGD: A portal for linked open government data ecosystems (DING et al., 2011)	Dados Abertos Conectados Governamentais.
P9	Internacional	Linked Open Data: The Essentials - A Quick Start Guide for Decision Makers (BAUER; KALTENBÖCK, 2011)	Dados Abertos Conectados
P10	Internacional	The Joy of Data - A Cookbook for Publishing Linked Government Data on the web (HYLAND; WOOD, 2011)	Dados Abertos Conectados
P11	Grécia	Applying Linked Data Technologies to Greek Open Government Data: A Case Study (GALIOTOU; FRAGKOU, 2013)	Dados Abertos Conectados Governamentais
P12	Internacional	Managing the Life-Cycle of Linked Data with the LOD2 Stack (AUER et al., 2012)	Dados Abertos Conectados
P13	Espanha	Methodological Guidelines for Publishing Government Linked Data (VILLAZÓN-TERRAZAS et al., 2011)	Dados Abertos Conectados Governamentais
P14	União Europeia	Methodology for publishing datasets as open data - COMSODE; Documents of Practice for Methodology for publishing datasets as open data - COMSODE (COMSODE, 2014a; COMSODE, 2014a)	Dados Abertos Governamentais
P15	Internacional	Open Data Handbook (OKF, 2015)	Dados Abertos
P16	Brasil	Guia de Dados Abertos; Guia de Incentivo ao Re uso de Dados Abertos; Guia de Web Semântica; Guia do Modelo de Maturidade de Dados Abertos (SÃO PAULO; REINO UNIDO, 2015)	Dados Abertos Conectados Governamentais

Código	Origem	Documentos (processos) Avaliados	Tipo
P17	Internacional	Guidelines on Open Government Data for Citizen Engagement (NAÇÕES UNIDAS, 2013)	Dados Abertos Governamentais
P18	Internacional	Open Government Data Toolkit (BANCO MUNDIAL, 2016)	Dados Abertos Governamentais
P19	Brasil	Manual Dos Dados Abertos: Governo; Manual dos Dados Abertos: Desenvolvedores. (W3C BRASIL, 2011)	Dados Abertos Governamentais
P20	Internacional	Guidelines for Open Data Policies (SUNLIGHT FOUNDATIONS, 2014)	Dados Abertos
P21	Internacional	How to Publish Linked Data on the Web (BIZER; CYGANIAK; HEATH); Chapter Recipes for Publishing Linked Data in Linked Data: Evolving the Web into a Global Data Space (HEATH; BIZER, 2011)	Dados Abertos Conectados
P22	Brasil	Piece of Cake (ÁVILA, 2015)	Dados Abertos Conectados Governamentais
Legenda:  Processos analisados por Ávila (2015)  Processos adicionados na nova análise			

Fonte: Elaborado pelo autor.

4.2 Descrição dos processos para publicação de Dados Abertos Conectados adicionados na nova análise

O P16 é constituído de vários documentos oriundos do projeto de cooperação técnica entre o Governo do Estado de São Paulo e o Reino Unido. O “Guia de Dados Abertos” detalhando os aspectos técnicos e de gestão tecnológica do processo com foco na abertura das bases; O “Guia de Web Semântica” enfocando as conexões entre os dados; O “Guia do Modelo de Maturidade de Dados Abertos” enfatizando a gestão estratégica e a organização das atividades no órgão publicador e o “Guia de Incentivo ao Reuso de Dados Abertos”

tratando das dimensões de governança e interação com a comunidade usuária interna e externa ao governo. Mesmo em um contexto no Poder Executivo Estadual, a relevância da produção técnica está pautada na capacidade de transição na aplicação de padrões internacionais em um contexto regional. Um processo rico em conhecimento, incorporando diversos elementos do conhecimento tácito, como experiências e vivências da equipe do projeto. Os documentos apresentam ineditismo ao documentar conhecimento subjetivo, não mensurável, e de difícil captura e transmissão, se tornando muito valiosos.

O P17 é o processo que está pautado em recomendações internacionais das Nações Unidas, abordando aspectos técnicos e legais, mas com um foco maior nos aspectos políticos, por entender a relevância do tema para as esferas econômica, social e ambiental de uma nação, inclusive países em desenvolvimento. Sendo um órgão que funciona como uma interface entre diversas nações, o processo se torna relevante por trabalhar com a mudança cultural em um nível global, e ao mesmo tempo consegue ser minimalista ao ponto de não atender apenas decisores políticos, contemplando também técnicos, cidadãos e outros. O ponto forte do processo está no desafio da adequação das recomendações a distintos contextos geográficos e culturais.

O “Open Government Data Toolkit” (P18) é um processo desenvolvido pelo Banco Mundial, uma importante organização no cenário internacional, que tem como objetivo fomentar o crescimento econômico dos países em desenvolvimento. Os benefícios na área econômica oriundos da abertura e a conexão dos dados fizeram o Banco Mundial desenvolver um conjunto de ferramentas para ajudar as nações a iniciar e desenvolver um programa de dados abertos. Tal programa poderá despertar o estímulo à inovação, a competição, o acesso ao conhecimento coletivo. O diferencial do processo está na concepção baseada na abordagem “*work in progress*”, ou seja, um processo inacabado que estará sempre em revisão e atualização, retroalimentado pelo *feedback* oriundo das experiências do mundo real.

O Manual de Dados Abertos (P19) é uma série com duas publicações de propostas distintas, publicada pelo W3C, resultado do acordo de cooperação técnica-científica entre o Laboratório Brasileiro de Cultura Digital e o Núcleo de Informação e Coordenação do Ponto BR (NIC.br). O primeiro volume traz uma leitura com o teor básico sobre o tema Dados Abertos, tendo como público-alvo leitores com perfis de agentes públicos, pessoas interessadas, especialistas de dados e profissionais de áreas afins. O segundo volume oferece uma visão mais técnica, tendo como público-alvo toda a comunidade de desenvolvedores

Web. Um material com uma concepção coletiva, com reunião de artigos específicos, troca de ideias, sugestões e revisões. Os manuais apresentam distintas características, reflexo do resultado de um grupo de trabalho heterogêneo, que reuniu representantes governamentais, representantes da sociedade, membros da academia, ONGs e desenvolvedores.

O processo P20 é um trabalho da *Sunlight Foundation*, uma organização estadunidense, apartidária e sem fins lucrativos. Iniciou suas atividades focando no congresso americano, e extrapolou fronteiras ao expandir seu foco para níveis locais, estaduais, federais e internacionais. Um processo que traz elementos construídos por uma instituição do terceiro setor para construção de uma política de dados abertos no âmbito governamental. O processo oferece exemplos destacando maneiras de implementar cada uma das recomendações. Os exemplos não são exaustivos, refletindo o que a empresa acredita serem ideias que deverão ser experimentadas pelas instituições.

Os documentos do processo P21 são esboçados em forma de várias receitas práticas para a publicação de informações como dados conectados na Web, ou seja, um tutorial. Mesmo com a característica de passo-a-passo, não percebemos uma abordagem conceitual básica, e sim uma abordagem puramente técnica e de forma detalhada. A relação entre os dois trabalhos está pautada em uma evolução histórica onde “How to Publish Linked Data on the Web” foi publicado em 2007 e “Linked Data: Evolving the Web into a Global Data Space” foi publicado em 2011. Mesmo o documento atual estando disponível com mais detalhes e atualizado, a versão antiga nos fornece elementos não contemplados na versão atualizada. Provavelmente, um reflexo da presença de mais um autor no trabalho anterior.

O processo P22, proposto por ÁVILA (2015), incorporou diversas características do esquema 5-Estrelas dos Dados Abertos, dos modelos de processo de software iterativos e do PMBoK. Além de incorporar elementos extraídos da revisão de literatura abrangendo 15 processos de Publicação de Dados Abertos Conectados. O “Piece of Cake” tem uma abordagem em camadas que são implementadas de acordo com o nível de maturidade da instituição, ou seja, cada camada fornece um conjunto de atividades com uma curva de aprendizagem diferenciada.

4.3 Metodologia para Análise dos Processos

Para cada processo de publicação de dados, ÁVILA (2015) identificou a existência de recomendações que implementavam no mínimo uma das melhores práticas sugerida pelo W3C, elencadas no documento “Best Practices for Publishing Linked Data (BPLD)”¹⁷. Ao longo das 10 BPLDs, foram identificadas 70 recomendações nos processos, conforme o quadro 4.2. Os identificadores das recomendações entre parênteses foram herdados do trabalho de ÁVILA (2015). O anexo A demonstra de qual trabalho cada recomendação foi retirada.

Quadro 4.2 - Recomendações para publicação de Dados Abertos Governamentais.

BPLD	Recomendações
1. Preparar partes interessadas (<i>stakeholders</i>)	Identificar os benefícios para a abertura de dados (1A); Identificar as Partes Interessadas (1B); Definir perfis profissionais a serem envolvidos (1C); Definir grupos de usuários dos dados (1D); Elaborar um plano de ações para publicação dos dados (1E); Capacitar os envolvidos (1F)
2. Selecionar conjuntos de dados	Analisar a estrutura organizacional (2A); Estabelecer diretrizes que orientem a priorização de dados a serem abertos (2B); Realizar consultas aos usuários sobre a demanda de dados (2C); Identificar os dados que serão abertos (2D); Definir nível de maturidade da abertura (1-5 estrelas) (2E); Analisar o nível de sigilo dos dados e informações (2F); Analisar relatórios anuais e documentação existente (2G); Analisar o esforço para abertura de dados (2H); Fazer e validar mapa de responsabilidades entre conjuntos de dados e unidades de negócio responsáveis (2I); Identificar e analisar sistemas de informação que poderão ser objeto da abertura de dados (2J); Identificar dados que podem ser conectados (2K);
3. Modelar os dados	Gerar cópias de segurança das bases de dados que serão abertas (3A); Higienizar os dados (3B); Estabelecer rotinas de conversão de dados para formatos legíveis por máquina (3C); Anonimizar dados sensíveis (3D); Modelar rotinas automatizadas (ETL) (3E); Analisar se os dados serão conectados ou não (3F); Estabelecer ou aprimorar documentação de dados (esquemas, vocabulários e ontologias) (3G);

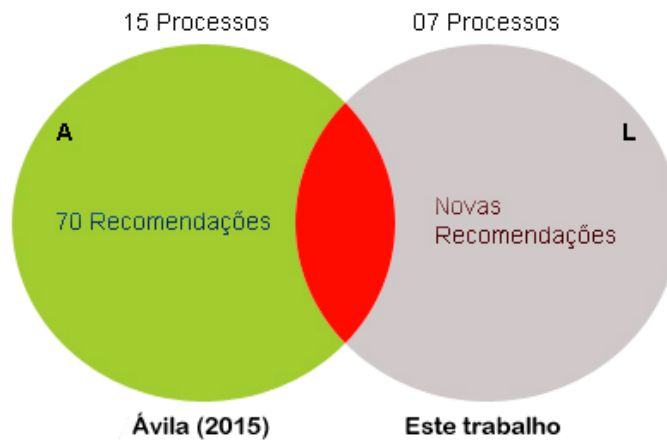
¹⁷ <https://www.w3.org/TR/ld-bp/>

BPLD	Recomendações
4. Especificar uma licença apropriada	Adotar Licenças Não restritivas (4A); Estabelecer de questões-chave para definição de licenças (4B); Apresentar opções de licenças a serem adotadas (4C)
5. Estabelecer bons identificadores Universais (URIs) para dados conectados	Utilizar URIs para conectar os dados (5A); Estabelecer URIs persistentes, que não se alterem em nenhum momento (5B); Proporcionar pelo menos um recurso de dados em formato que seja legível por máquina para cada URI (5C); Usar URIs como nomes para as coisas (5D); Estabelecer Design simplificado de URIs (5E); Utilizar identificadores relacionados a informações do mundo real (5F); Usar URIs HTTP para que recursos de dados possam ser encontrados via <i>Web</i> por pessoas e máquinas (5G); Estabelecer URIs neutras (5H); Utilizar datas em URIs com moderação (5I); Utilizar <i>hashs</i> (#) em URIs cautelosamente (5J); URIs das entidades (conjuntos de dados ou recursos) sejam diferentes das URIs das páginas que apresentam estes recursos para a leitura feita por humanos (5K);
6. Utilizar vocabulários padrão	Estabelecer metadados obrigatórios (6A); Criar um esquema de dados para cada conjunto de dados (6B); Incentivar o re-uso de vocabulários (6C); Publicar esquemas de dados em arquivos diferentes (6D); Determinar linguagens para expressar esquemas de dados (6E); Estabelecer critérios de escolha de vocabulários (6F); Certificar que os dados estão conectados a outros conjuntos de dados (6G); Desenvolver ou utilizar ontologias para estruturar a semântica dos dados (6H);
7. Converter e enriquecer dados	Converter dados para múltiplas finalidades e usos (7A); Adotar rotinas ETL para enriquecimento de dados (7B); Estabelecer bons <i>links</i> com outros conjuntos de dados (7C); Permitir o envolvimento de várias pessoas na identificação de como os dados a serem convertidos se relacionam com outros dados (7D); Utilizar rotinas automatizadas de conversão de dados, como a triplificação, quando possível (7E); Converter dados em várias serializações RDF (7F);
8. Prover acesso automatizado aos dados	Disponibilizar bases completas para <i>download</i> (<i>dumps</i>) (8A); Estabelecer um Mapa de Decisões Tecnológicas (8B); Desenvolver uma API (8C); Desenvolver um endpoint SPARQL (8D);

BPLD	Recomendações
9. Anunciar os conjuntos de dados para o público	Publicar metadados junto aos dados (9A); Estabelecer dados tecnicamente e legalmente abertos (9B); Disponibilizar os dados com o menor custo possível ao usuário, preferencialmente de modo gratuito na internet (9C); Divulgar dados em meios complementares (Catálogos, FTP, Torrent) (9D); Divulgar dados em seções destacadas de sítios de governo (9E); Estabelecer recursos de consulta parcial da base de dados como uma API ou <i>Webservice</i> (9F); Estabelecer visualizações e demais recursos de exploração dos dados (9G); Melhorar os dados para que sejam mais facilmente encontrados por máquinas (9H); Disponibilizar dados conectados em servidores de triplas (9I);
10. Estabelecer um contrato social para os dados publicados	Estabelecer mecanismos de monitoramento e avaliação da oferta de dados disponibilizados ao público (10A); Estabelecer espaços para recebimento do <i>feedback</i> do usuário, preferencialmente publicando dados de uma pessoa e/ou telefone de contato para esclarecimento de dúvidas sobre o uso e disponibilidade dos dados (10B); Disponibilizar leis e atos normativos que explicitem aos usuários quanto às obrigações dos governos em publicarem dados com qualidade e disponibilidade (10C); Estabelecer com clareza que o processo de publicação contempla etapas de manutenção e atualização dos dados (10D); Utilizar tecnologias que mantenham os dados conectados disponíveis, atualizados e abertos (10E);

Fonte: Elaborado pelo autor.

Para os novos processos, identificamos as recomendações considerando na análise a interseção (Figura 4.1) com os processos analisados por ÁVILA (2015), ou seja, foram consideradas apenas recomendações que já apareciam no trabalho de ÁVILA (2015). ÁVILA (2015) não estabeleceu um critério objetivo sobre o que seria ou não seria uma recomendação. O estabelecimento de critério objetivo foi dificultado porque cada iniciativa faz de um jeito diferente, não apresentando uma linguagem padronizada. Neste caso, ele fez uma interpretação subjetiva, especialmente considerando quando a ocorrência do que era considerada uma recomendação acontecia em mais de um processo. O trabalho de ÁVILA (2015) já foi validado pela comunidade científica. Por esta razão, adotamos o mesmo procedimento na identificação das recomendações.

Figura 4.1 - Relação dos processos analisados

Fonte: Elaborado pelo autor

Então para cada recomendação encontrada nos novos processos, a mesma deveria satisfazer a equação abaixo:

$$A \cap L = \{x \mid x \in A \text{ e } x \in L\}$$

Sendo:

$A \cap L$ = intersecção

x = recomendações encontradas nos processos.

A = conjunto dos processos analisados por Ávila (2015).

L = conjunto dos processos analisados por este trabalho.

Naturalmente, foram encontradas novas recomendações que não faziam parte dessa intersecção. No entanto, não incluímos as novas recomendações na análise por dois motivos: primeiramente, o objetivo da análise consistiu em classificar as recomendações existentes em essenciais (Obrigatórias) ou opcionais (Desejáveis), e não a inclusão de uma nova recomendação. Além do quantitativo de 70 recomendações do universo de ÁVILA (2015) ser considerada suficiente para a nossa análise.

Considerando cada recomendação identificada como uma atividade de um processo de publicação, as mesmas foram classificadas por ÁVILA (2015) em:

• **Obrigatórias:** Atividades que devem ser executadas por qualquer instituição publicadora de Dados Abertos (conectados) governamentais, independente do nível de maturidade. Aplicáveis principalmente para instituições que possuem baixa maturidade em publicação de dados e desejam um conjunto sucinto de atividades para disponibilizar Dados Abertos aos seus usuários.

• **Desejáveis:** Atividades que devem ser executadas por instituições com nível mais alto de maturidade. Este grupo de atividades busca melhorar, significativamente, a oferta de dados mediante a incorporação de requisitos de maior qualidade, disponibilidade e conformidade dos dados que serão disponibilizados.

Com a adição de novos processos existe a probabilidade de alteração na classificação, fazendo com que os elementos classificados sofram alterações, ou seja, uma recomendação que antes era considerada “Obrigatória” pode se tornar “Desejável” e vice-versa. As fórmulas do cálculo de classificação foram criadas pelo autor deste trabalho com base na análise dos gráficos do trabalho de ÁVILA (2015). Para classificação, calculamos o desvio em relação à média aritmética entre as recomendações identificadas e a presença nos processos que a contemplaram, como descrito a seguir.

$$\bar{x} = \frac{\sum x_i}{n}$$

Sendo:

x_i = valores referentes ao quantitativo de vezes que cada recomendação foi contemplada nos processos. Exemplo: A recomendação “identificar as partes interessadas (1B)” foi observada no processo P1, P3, P13 e P14, então $x_i = 4$. A recomendação “Definir grupos de usuário de dados (1D)” foi observada nos processos P1 e P14, então $x_i = 2$. Neste caso, $\sum x_i = 6$.

n = o número do quantitativo de valores x_i . Para nosso exemplo, $n = 2$.

$\bar{x}$ = a média aritmética. Para nosso exemplo, $\frac{6}{2} = 3$.

Então, achamos o valor de x_i para cada recomendação e calcula-se a diferença desse valor com a média aritmética dos conjuntos de recomendações de uma das melhores práticas a qual ela pertence. Esta diferença é calculada pelo desvio em relação à média:

$$d_i = x_i - \bar{x}$$

Com o desvio calculado, o fator de classificação é utilizado com a seguinte lógica. Se o desvio for maior ou igual à média ($d_i \geq \bar{x} \rightarrow o$), então a recomendação é classificada como obrigatória. Se o desvio for menor em relação à média ($d_i < \bar{x} \rightarrow d$), então a recomendação é classificada como desejável. Em nosso exemplo, O d_i da recomendação 1B é igual 1($d_{i=4-3}$), ou seja, uma recomendação obrigatória. O d_i da recomendação 1D é igual -1($d_{i=2-3}$), ou seja, uma recomendação desejável.

4.4 Análise dos resultados

Esta seção apresenta os resultados provenientes da análise de cada recomendação em relação a sua classificação em “Obrigatória” ou “Desejável”. Serão apresentados os resultados do trabalho de ÁVILA (2015) e os resultados da análise do autor referente aos novos processos. Os resultados demonstram como cada recomendação foi afetada em relação a classificação após adição dos novos processos.

4.4.1 Recomendações para “Preparar Partes Interessadas”

Na análise desta BPLD, 10 dos 22 processos apresentaram recomendações de preparação das partes interessadas. A adição de novos processos não alterou os resultados de Ávila (2015), permanecendo as atividades de “Identificar as partes interessadas (1B)”, “Definir perfis profissionais a serem envolvidos (1C)” e “Elaborar um plano de ações para publicação dos dados (1E)” como fortemente recomendadas. A recomendação 1B é encontrada nos processos P1, P3, P13, P14, P16, P17, P18 e P22. A 1C é encontrada nos processos P1, P3, P4, P14, P16 e P22. A 1E é encontrada nos processos P1, P3, P13, P14, P16, P17, P20 e P22. A tabela 3 demonstra as recomendações com os valores destacados em verde como “Obrigatórias” e os valores destacados em vermelho como “Desejáveis”.

Tabela 4.1 - Quantitativo de recomendações para "Preparar Partes Interessadas" nos processos de publicação de dados abertos analisados.

Recomendações	Aparições nos processos analisados por Ávila (2015)	Aparições na análise dos novos processos	Total de Aparições	Desvio em relação à média	Classificação
Identificar os benefícios para aberturas de dados (1A)	1	3	4	-1,66667	Desejável
Identificar partes interessadas (1B)	4	4	8	2,333333	Obrigatória

Recomendações	Aparições nos processos analisados por Ávila (2015)	Aparições na análise dos novos processos	Total de Aparições	Desvio em relação à média	Classificação
Definir profissionais a serem envolvidos (1C)	4	2	6	0,333333	Obrigatória
Definir grupos de usuários dos dados (1D)	2	1	3	-2,66667	Desejável
Elaborar um plano de ações para publicação dos dados (1E)	4	4	8	2,333333	Obrigatória
Capacitar os envolvidos (1F)	1	4	5	-0,66667	Desejável
Média	2,666666667	3	5, 666666667		

Fonte: Elaborado pelo autor.

A seguir, apresentamos uma breve análise dos processos que foram adicionados à nossa análise. Observamos que as atividades de escolha e identificação dos atores nos processos P16 e P17 são influenciadas pelo nível de maturidade da organização. O processo P17 elenca potenciais perfis que estarão engajados ao processo, ressaltando que o diálogo e o entendimento pelas partes interessadas são imprescindíveis para o sucesso e a sustentabilidade das iniciativas de dados abertos governamentais. O processo P18 traz uma consideração importante, ao ressaltar que a identificação das partes interessadas não se resume apenas no lado dos publicadores de dados, e sim na identificação no lado dos consumidores de dados. Para ambos os públicos, o envolvimento precoce muitas vezes implica na crescente conscientização sobre o uso, benefícios e oportunidades apresentadas pela iniciativa.

Existe uma importância em definir o escopo dos dados que serão abertos, bem como os recursos e os prazos em um projeto de abertura de bases. Os processos P16 e P17 trazem a necessidade de delimitar o projeto, devido aos desafios técnicos e humanos para as instituições no processo. A fase de planejamento deverá definir os objetivos estratégicos (Engajamento do cidadão, transparência, prestação de contas, inovação, entre outros) que a organização pretende com a abertura e conexão dos dados, além de identificar e planejar os recursos humanos, tecnológicos e financeiros necessários. A delimitação do escopo no processo P20 não está pautada em um nível tático, pela natureza da proposta não ser um plano e sim uma política. Mas apresenta da forma, a necessidade de traçar um caminho com os objetivos a serem alcançados com a abertura e conexão dos dados.

Mesmo com um capítulo dedicado aos benefícios para abertura dos dados, o processo P16 não elenca recomendações para identificação desses benefícios dentro da organização.

Ao invés disso, o capítulo expõe os diversos benefícios ao se abrir os dados juntamente com exemplos de casos bem-sucedidos de outras organizações. Mas a linguagem utilizada é meramente informativa ao leitor, não sendo vislumbrada como uma prática a ser seguida. Ocorre um comportamento inverso nos processos P17 e P18, transmitindo implicitamente a recomendação, por causa dos diferentes interesses das partes interessadas. A saída dessa recomendação será argumento persuasivo para convencimento das partes interessadas no apoio da iniciativa de abertura das bases. Precisa-se de cuidados na identificação e apresentação desses benefícios, porque quando são apresentadas falsas expectativas para as partes interessadas, pode gerar frustrações desligamento, enfraquecendo a iniciativa.

Existe uma diversidade de categorias de usuários com necessidades específicas. A recomendação “Definir grupos de usuários dos dados (1D)” é de conhecer o grupo alvo que utilizarão os dados. Tal ação ajudará o publicador na tomada de decisão na escolha de quais canais de comunicação, quais serão suas estratégias de promoção, como irá proceder seus relacionamentos e suas atividades de suporte. Somente um dos novos processos apresentou claramente esta atividade como recomendação. Existe um destaque em alguns trechos nos outros processos, comentando a adoção desta prática. Mas que se confunde com a identificação das partes interessadas, visto que um grupo de usuários é uma das partes interessadas no processo.

Cabe ressaltar nas variações ocorridas com a adição dos novos processos, que a atividade de “Capacitar os envolvidos (1F)”, apresentou um aumento significativo em sua relevância para sair da classificação de “Desejável” para “Obrigatória”, mas não conseguiu devido o desvio ter sido negativo em relação à média. Mesmo não tendo um desvio positivo em relação à média, a recomendação 1F ganha importância com os novos processos, aparecendo nos processos P1, P16, P17, P18 e P22. Para criar um ambiente sustentável para as iniciativas de Dados Abertos Governamentais, o processo P17 ressalta que é preciso explorar opções facilitadoras na capacitação e formação dos envolvidos. Para o processo P18 é necessário desenvolver a capacidade, aumentando o entendimento dos conceitos juntamente com a capacidade técnica de manipulação, para extração de conhecimento e um melhor uso dos dados pelos os usuários.

4.4.2 Recomendações para “Selecionar Conjuntos de Dados”

Na análise desta BPLD, 18 dos 22 processos apresentaram recomendações de seleção dos conjuntos de dados. A adição de novos processos fez com que a recomendação “Identificar e analisar sistemas de informação que poderão ser objeto da abertura de dados (2C)” fosse rebaixada de “Obrigatória” para “Desejável”, e a recomendação “Analisar o nível de sigilo dos dados e informações (2E)” foi promovida de “Desejável” para “Obrigatória”. A tabela 4.2 demonstra os resultados para as recomendações dessa BPLD.

Tabela 4.2 - Quantitativo de recomendações para "Selecionar Conjuntos de Dados" nos processos de publicação de dados abertos analisados.

Recomendações	Aparições nos processos analisados por Ávila (2015)	Aparições na análise dos novos processos	Total	Desvio em relação à média	Classificação
Analisar relatórios anuais e documentação existente (2A)	1	1	2	-2,58333	Desejável
Analisar a estrutura organizacional (2B)	2	1	3	-1,58333	Desejável
Identificar e analisar sistemas de informação que poderão ser objeto da abertura de dados (2C)	3	1	4	-0,58333	Desejável
Estabelecer diretrizes que orientem a priorização de dados a serem abertos (2D)	4	5	9	4,416667	Obrigatória
Analisar o nível de sigilo dos dados a serem abertos (2E)	2	4	6	1,416667	Obrigatória
Realizar consultas aos usuários sobre a demanda de dados (2F)	3	3	6	1,416667	Obrigatória
Identificar os dados que serão	7	5	12	7,416667	Obrigatória

Recomendações	Aparições nos processos analisados por Ávila (2015)	Aparições na análise dos novos processos	Total	Desvio em relação à média	Classificação
abertos (2G)					
Identificar dados que podem ser conectados (2H)	1	1	2	-2,58333	Desejável
Definir nível de enriquecimento dos dados a serem publicados (1-5 estrelas) (2I)	1	1	2	-2,58333	Desejável
Analisar o esforço para abertura de dados (2J)	1	1	2	-2,58333	Desejável
Fazer e validar mapa de responsabilidades entre conjuntos de dados e unidades de negócio responsáveis (2K)	1	2	3	-1,58333	Desejável
Não apresentou recomendação, apenas destacando a adoção desta prática (2L)	4	0	4	-0,58333	
Média	2,5	2,083333333	4,5833		

Fonte: Elaborado pelo autor.

A delimitação do escopo dos dados a serem abertos é sugerida nos processos P2, P4, P14, P15, P16, P17, P18, P20 E P22 com a recomendação “Estabelecer diretrizes que orientem a priorização de dados a serem abertos (2D)”. O processo P17 ressalta que nem todos conjuntos de dados são relevantes para os objetivos do órgão, sejam eles o engajamento do cidadão, a transparência governamental, o atendimento a legislação, entre outros. O processo P20 coloca que diversas variáveis como objetivos, atores e eventos influenciam na determinação dos dados que serão divulgados, no entanto o processo se torna particionado por razões como a insuficiência de recursos humanos, tecnológicos e financeiros. Mas o fator custo deve estar balanceado com outros métodos de priorização, como relevância, demanda do público, entre outros.

A recomendação “Analisar o nível de sigilo dos dados a serem abertos (2G)” é contemplada nos processos P1, P3, P16, P17, P20 e P22. O processo P20 ressalta que quando

temos um plano ou política de dados abertos bem definida, estes se tornam complementares às legislações existentes, contemplando as exclusões previstas em leis. Observamos isso no processo P16, o qual referencia a Lei de Acesso à Informação, elencando as categorias de classificação do que pode ser considerado sigiloso e recomendando a análise dos dados para evitar a publicação de informações cujo sigilo é considerado imprescindível. São informações que podem pôr em risco a segurança da sociedade e do estado, incluindo pesquisas científicas e tecnológicas com este fim, segredos de Estado, além de informações pessoais capazes de identificar indivíduos. O processo P20 ressalta que o nível de sigilo pode estar associado a camadas de acesso. Dessa forma, uma informação que pode ser muito sensível para liberação ao público, muitas vezes pode ser disponibilizada para um determinado grupo, como pesquisadores acadêmicos, unidades internas entre outros.

A necessidade de “Realizar consultas aos usuários sobre a demanda dos dados (2F)” é destacada pelos os processos P4, P14, P15, P16, P17 e P22. Tendo em vista que o processo P16 afirma que haverá pontos de contato do publicador com seus usuários, de certa forma existirá uma transversalidade desta recomendação com as recomendações da BPLD “Estabelecer um contrato social para os dados publicados”, devido às práticas, estratégias e tecnologias aplicadas para gerenciar e analisar as interações com os usuários. Essas ferramentas devem ser utilizadas pelas organizações para extraírem das interações, o reconhecimento da existência de segmentos de usuários com necessidades específicas e demandas específicas. Para o processo P17, é necessário ter uma boa compreensão de quais dados podem ser valiosos sob o prisma da comunidade. A intenção principal é obter uma sensação de quais dados serão demandados.

Dentro desta BPLD, a recomendação com um maior quantitativo de aparições entre os processos foi a “Identificação dos dados que serão abertos (2G)” em P2, P3, P4, P5, P13, P14, P15, P16, P18, P19, P20 e P22. O processo P16 ressalta que não podemos tomar como base apenas a relevância dos referidos dados, e que a recomendação 2G, selecionando as bases, ou suas partes, deverá está alinhada de acordo com a maturidade de cada organização. O processo P19 destaca a necessidade da escolha do conjunto de dados que se deseja abrir, mas descarta a priorização, visto que para um movimento embrionário, a consideração da viabilidade da publicação dos dados e o interesse das pessoas são mais importantes nas organizações com baixo nível de maturidade. Para o fortalecimento do processo, a recomendação para o P20 se resume em inventariar o que pode ser aberto. O P16 elenca

perguntas como: *Quais bases poderão ser publicadas imediatamente? Com quantas estrelas? Quais bases precisarão passar por algum tratamento antes de serem publicadas? Quais bases serão publicadas no longo prazo? Qual será o mínimo de estrelas que o órgão deseja conseguir em todas as bases que forem publicadas?* Essas perguntas são insumos condicionantes para diversas recomendações, porque existirá a definição de onde estamos e para onde queremos caminhar em nível de maturidade.

4.4.3 Recomendações para “Modelar os dados”

Recomendações que remetem à modelagem dos dados são encontradas em 18 processos (P1, P2, P3, P4, P6, P7, P8, P9, P10, P11, P13, P14, P16, P17, P18, P20, P21 e P22). Conforme a Tabela 4.3, os novos processos acrescentaram a atividade de “Anonimizar dados sensíveis (3E)” ao grupo de recomendações obrigatórias. O grupo era composto apenas pelas recomendações “Higienizar os dados (3D)”, com intuito de excluir dos seus dados às informações irrelevantes ou possíveis erros; “Estabelecer rotinas de conversão de dados para formatos legíveis por máquinas (3F)”, ofertar dados em vários formatos legíveis por máquinas como XML, CSV, TXT, JSON ou KML; e “Modelar rotinas automatizadas (ETL) (3G)”, automatizando as rotinas de publicação e manutenção dos dados. A recomendação “Modelar rotinas automatizadas (ETL) (3G)” foi rebaixada de “Obrigatória” para “Desejável”.

Tabela 4.3 - Quantitativo de recomendações para "Modelar os dados" nos processos de publicação de dados abertos analisados.

Recomendações	Aparições nos processos analisados por Ávila (2015)	Aparições na análise dos novos processos	Total	Desvio em relação à média	Classificação
Gerar cópia de segurança das bases de dados que serão abertas (3A)	3	1	4	-1,42286	Desejável
Analisar quais dados serão conectados (3B)	3	1	4	-1,42286	Desejável

Recomendações	Aparições nos processos analisados por Ávila (2015)	Aparições na análise dos novos processos	Total	Desvio em relação à média	Classificação
Estabelecer ou aprimorar documentação de dados (esquemas, vocabulários e ontologias) (3C)	1	2	3	-2,14286	Desejável
Higienizar os dados (3D)	6	1	7	1,857143	Obrigatória
Anonimizar dados sensíveis (3E)	2	5	7	1,857143	Obrigatória
Estabelecer rotinas de conversão de dados para formatos legíveis por máquinas (3F)	4	2	6	0,857143	Obrigatória
Modelar rotinas automatizadas (ETL) (3G)	4	1	5	-0,14286	Desejável
Media	3,285714286	1,857142857	5,142857143		

Fonte: Elaborado pelo o autor.

A recomendação “Analisar o nível de sigilo dos dados a serem abertos (2E)” pode ter gerado um reflexo em “Anonimizar dados sensíveis (3E)” por apresentarem um grau de relacionamento entre as duas atividades, independente de uma estar em nível de planejamento e a outra em um nível operacional respectivamente. As duas recomendações são encontradas simultaneamente nos novos processos P16, P17, P20 e P22. Não existe um detalhamento da recomendação no processo P16, a qual é identificada em frases curtas dentro de um capítulo relacionado com a elaboração do plano de ações para publicação dos dados. O processo P17 recomenda o estabelecimento de uma política para classificação de quais dados estão passíveis para anonimização. Sendo considerada uma questão de risco a ser mitigado, caso dados de natureza restrita venha a público. A anonimização para o processo P18, não deverá consumir um nível de esforço alto para equipe, visto que muitos tipos de dados do governo não envolvem informações confidenciais e, portanto, têm pouca ou nenhuma necessidade de técnicas de anonimato.

4.4.4 Recomendações para “Especificar uma licença apropriada”

Com relação a melhor prática “Especificar uma licença apropriada”, os processos P2, P5, P10, P13, P14, P16, P17, P19 e P20 apresentaram recomendações. Enquanto os processos P4, P9, P15 e P18 não apresentaram recomendação, mas destacaram a adoção desta prática. Conforme a Tabela 4.4, a adição dos novos processos, elevou na classificação a recomendação “Adotar licenças não restritivas (4B)” de desejável para obrigatória. Os processos P17 e P20 recomendam que as organizações adotem licenças não restritivas, permitindo a reutilização e redistribuição dos dados por qualquer pessoa para qualquer finalidade. O processo P16 destaca que na formulação de termos de utilização ou de uma licença aberta para uso de dados, devemos verificar compatibilidade da licença com outras licenças de dados abertos. No entanto, ao adotar uma licença que seja menos restritiva, maior será a interoperabilidade jurídica com outros conjuntos de dados que adotam outras licenças.

Tabela 4.4 - Quantitativo de recomendações para "Especificar uma licença apropriada" nos processos de publicação de dados abertos analisados.

Recomendações	Aparições nos processos analisados por Ávila (2015)	Aparições na análise dos novos processos	Total	Desvio em relação à média	Classificação
Estabelecer questões-chave para definição de licenças (4A)	1	1	2	-2,25	Desejável
Adotar licenças não restritivas (4B)	1	5	6	1,75	Obrigatória
Apresentar opções de licenças a serem adotadas (4C)	4	1	5	0,75	Obrigatória
Não apresentou recomendação, apenas destacando a adoção desta prática (4D)	3	1	4	-0,25	
Media	2,25	2	4,25		

Fonte: Elaborado pelo autor.

A permanência como obrigatória da recomendação “Apresentar opções de licenças a serem adotadas (4C)”, ocorreu pelos altos quantitativos de aparições nos processos P2, P5, P10 e P13 analisados por ÁVILA (2015). Mesmo tendo um decréscimo no desvio em relação à média com a adição dos novos processos, a mesma continuou em sua posição. Torna-se uma opção fortemente recomendada, o órgão publicador ser apresentado a diversas opções de

licença de uso, não somente para adotar a que for mais conveniente ao conjunto de dados a ser publicado, mas responder indagações nas quais o processo P16 elenca, como: *O ordenamento jurídico brasileiro é suficiente para tratar os dados de forma aberta, sem necessidade de licença? Outras licenças que já existem se enquadram no contexto brasileiro? Será necessário criar uma licença específica?* O processo P18 destaca que a utilização de licenças comumente utilizadas pela comunidade oferece várias vantagens em relação a uma nova licença criada, incluindo um maior entendimento entre os usuários, uma maior interoperabilidade e uma maior facilidade de cumprimento da mesma.

4.4.5 Recomendações para “Estabelecer bons identificadores universais (URIs)”

ÁVILA (2015) encontrou as recomendações bem distribuídas para o estabelecimento de bons identificadores universais, dificultando a identificação da predominância das atividades que seriam fortemente recomendadas. Os novos processos criaram uma maior disparidade entre essas recomendações, facilitando a definição entre a obrigatoriedade ou não das recomendações. As recomendações contabilizadas na Tabela 4.5 foram encontradas nos processos P10, P11, P12, P13, P14, P16, P19, P21 e P22. Cabe ressaltar que tais recomendações foram encontradas em aproximadamente 77% dos processos com foco em Dados Abertos Conectados, concluindo-se que as referidas atividades estão mais associadas com a conexão dos dados do que com a abertura dos dados.

Tabela 4.5 - Quantitativo de recomendações para "Estabelecer bons identificadores universais (URIs)" nos processos de publicação de dados abertos analisados.

Recomendações	Aparições nos processos analisados por Ávila(2015)	Aparições na análise dos novos processos	Total	Desvio em relação à média	Classificação
Estabelecer design simplificado de URIs (5A)	2	1	3	-0,75	Desejável
Usar URIs como nomes para as coisas (5B)	3	4	7	3,25	Obrigatória
Utilizar identificadores relacionados a informações do mundo real (5C)	3	1	4	0,25	Obrigatória
Usar URIs HTTP para que pessoas e máquinas possam encontrá-las via Web utilizando estes endereços (5D)	3	4	7	3,25	Obrigatória

Recomendações	Aparições nos processos analisados por Ávila(2015)	Aparições na análise dos novos processos	Total	Desvio em relação à média	Classificação
URIs das entidades (conjuntos de dados ou recursos) sejam diferentes das URIs das páginas que apresentam estes recursos para a leitura feita por humanos (5E)	2	1	3	-0,75	Desejável
Proporcionar pelo menos um recurso de dados em formato que seja legível por máquina para cada URI. (5F)	1	1	2	-1,75	Desejável
Estabelecer URIs persistentes, que não se alterem em nenhum momento (5G)	3	4	7	3,25	Obrigatória
Estabelecer URIs neutras (5I)	1	1	2	-1,75	Desejável
Utilizar datas em URIs com moderação (5I)	1	1	2	-1,75	Desejável
Utilizar <i>hashs</i> (#) em URIs cautelosamente (5J)	1	1	2	-1,75	Desejável
Utilizar URIs para conectar os dados (5K)	1	4	5	1,25	Obrigatória
Não apresentou recomendação, apenas destacando a adoção desta prática (5L)	1	0	1	-2,75	
Média	1,833333333	1,916666667	3,75		

Fonte: Elaborado pelo autor.

As recomendações “Usar URIs como nomes para as coisas (5B)”;

“Usar URIs HTTP para que pessoas e máquinas possam encontrá-las via Web utilizando estes endereços (5D)”;

“Estabelecer URIs persistentes, que não se alterem em nenhum momento (5D)” ou, pelo menos, dure o máximo de tempo possível; e “Utilizar URIs para conectar os dados (5K) apenas foram consideradas em alguns processos como atividades a serem realizadas, pelo fato das publicações citarem que para se ter Dados Abertos Conectados, a organização deverá obedecer os 4 princípios de *Linked Data* (BIZER; HEATH & BERNERS-LEE, 2009). Como o esgotamento desse tema não fazia parte do escopo dos processos, os mesmos referenciavam diversos esquemas e artigos para guiar o publicador na definição de URIs.

4.4.6 Recomendações para “Utilização de vocabulário padrão”

Para esta BPLD, ÁVILA (2015) encontrou as recomendações bem distribuídas dificultando a identificação das recomendações mais ou menos frequentes. No entanto ele ressalta que os órgãos publicadores poderiam referenciar o processo P14 quando houver a necessidade de estabelecer vocabulários, visto que o mesmo apresentou a maioria das recomendações. Com a adição dos novos processos, tivemos uma maior disparidade entre as recomendações, identificando as atividades entre obrigatórias e desejáveis. Das oito recomendações consideradas na análise, “Estabelecer metadados obrigatórios (6A)”, “Incentivar o reuso de vocabulário (6E)” e “Desenvolver ou utilizar ontologias para estruturar a semântica dos dados (6G)” aparecem em sete dos treze processos que apresentaram algum tipo de recomendação para esta BPLD. A Tabela 4.6 demonstra o quantitativo de recomendações, destacando as “Obrigatórias” e as “Desejáveis” para esta BPLD.

Tabela 4.6 - Quantitativo de recomendações para "Utilização de vocabulário padrão" nos processos de publicação de dados abertos analisados.

Recomendações	Aparições nos processos analisados por Ávila (2015)	Aparições na análise dos novos processos	Total	Desvio em relação à média	Classificação
Estabelecer metadados obrigatórios (6A)	3	4	7	2,875	Obrigatória
Determinar linguagens para expressar esquemas de dados (6B)	1	1	2	-2,125	Desejável
Criar um esquema de dados para cada conjunto de dados (6C)	1	1	2	-2,125	Desejável
Publicar esquemas de dados em arquivos diferentes (6D)	1	1	2	-2,125	Desejável
Incentivar o reuso de vocabulários (6E)	3	3	6	1,875	Obrigatória

Recomendações	Aparições nos processos analisados por Ávila (2015)	Aparições na análise dos novos processos	Total	Desvio em relação à média	Classificação
Estabelecer critérios de escolha de vocabulários (6F)	2	1	3	-1,125	Desejável
Desenvolver ou utilizar ontologias para estruturar a semântica dos dados (6G)	3	4	7	2,875	Obrigatória
Certificar que os dados estão conectados a outros conjuntos de dados (6H)	2	2	4	-0,125	Desejável
Média	2	2,125	4,125		

Fonte: Elaborado pelo o autor.

O processo P20 destaca que o “Estabelecimento de metadados obrigatórios (6A)”, juntamente com outras documentações, pode ser útil para diversas partes interessadas, provendo ajuda não apenas no entendimento sobre a criação, a qualidade e a utilização dos dados, mas também na descoberta e reutilização do conjunto de dados. O P16 ressalta a importância do órgão publicador disponibilizar os metadados em formatos que possibilitem a interpretação tanto para seres humanos quanto para máquinas. Sob o olhar do processo P21, os dados em si precisam ser auto descritivos sempre que possível.

A recomendação “Incentivar o reuso de vocabulários (6E)” é citada no processo P16 como um fator facilitador para fazer com que o significado pretendido pelo publicador dos dados seja o mesmo significado entendido pelo consumidor dos dados, ou seja, universalizar os conceitos para todos os envolvidos. Não existe um impedimento caso o publicador não queira reutilizar vocabulários comumente usados por outras iniciativas, mas seu objetivo do entendimento comum entre ambas às partes estará prejudicado. Para cada publicação específica, é preciso que se estabeleça um conjunto de vocabulários de referência, caso não exista nenhum que satisfaça as necessidades, uma nova ontologia ou vocabulário poderão ser criados (6G), com atenção aos cuidados do processo P16 de reutilizar o maior número possível de elementos de ontologias já existentes.

4.4.7 Recomendações para “Converter e enriquecer dados”

Conforme Tabela 4.7, a adição de novos processos para esta BPLD não causou alterações nos resultados obtidos por ÁVILA (2015). A maioria das ocorrências de recomendações nos novos processos apenas ocorreu no processo P22, cujo autor é o mesmo da análise anterior, nas quais reafirmaram suas convicções. As atividades “Utilizar rotinas automatizadas de conversão de dados, como a triplicação, quando possível” e “Conectar conjuntos de dados com outros dados relacionados” foram as que tiveram maior número de recomendações.

Tabela 4.7 - Quantitativo de recomendações para "Converter e enriquecer dados" nos processos de publicação de dados abertos analisados.

Recomendações	Aparições nos processos analisados por Ávila (2015)	Aparições na análise dos novos processos	Total	Desvio em relação à média	Classificação
Utilizar rotinas automatizadas de conversão de dados, como a triplicação, quando possível (7A)	6	1	7	3,166667	Obrigatória
Adotar rotinas ETL para enriquecimento de dados (7B)	2	1	3	-0,83333	Desejável
Conectar conjuntos de dados com outros dados relacionados (7C)	4	1	5	1,166667	Obrigatória
Permitir o envolvimento de várias pessoas na identificação de como os dados a serem convertidos se relacionam com outros dados (7D)	1	1	2	-1,83333	Desejável
Converter dados em várias serializações RDF (7E)	2	1	3	-0,83333	Desejável
Converter dados para múltiplas finalidades e usos (7F)	1	2	3	-0,83333	Desejável
Media	2,666666667	1	3,833333333		

Fonte: Elaborado pelo autor.

É importante observar que, para pequenos volumes de dados, a necessidade de automação no processo de conversão não seja algo latente. No entanto, à medida que a organização vai amadurecendo e deseja disponibilizar grandes volumes de dados, os seres humanos não serão capazes de realizar a conversão manualmente a partir desses dados. O processo P22 destaca três abordagens que poderão ser utilizadas para a recomendação 7A, a **conversão automática**, conhecida como triplificação, à **conversão parcial** em parte suportada por scripts automatizados e finalização mediante trabalho manual e a **modelagem**, que se baseia na elaboração de um modelo de conhecimento desenvolvido por especialistas humanos com posterior conversão automatizada dos dados a partir deste modelo.

Conforme os princípios de *Linked Data* e adoção de alguns processos com a recomendação “Utilizar URIs para conectar os dados (5K)”, um conjunto de dados deve fazer referência a outros conjuntos de dados, de forma a permitir a criação de novos conhecimento. No entanto, o processo P22 chama atenção para a realização de uma avaliação mais cuidadosa com intuito de escolher os conjuntos de dados mais relevantes a serem vinculados com o seu próprio conjunto.

4.4.8 Recomendações para “Prover acesso automatizado aos dados”

Para o provimento de acesso automatizado aos dados, a recomendação “Disponibilizar bases completas para *download (dumps)* (8D)” foi apresentada na Tabela 4.8 como sendo a mais relevante. Cabe ressaltar que, com adição dos novos processos, houve alteração da classificação dos resultados fornecidos por Ávila (2015), excluindo a recomendação “Desenvolver um endpoint SPARQL” do grupo de obrigatórias e tendo um aumento significativo no quantitativo de aparições da recomendação 8D. A ocorrência desse fato é devido à recomendação se apresentar como uma atividade básica em relação às outras para quem almeja utilizar esta BPLD. O processo P17 ressalta que uma API, *endpoint* SPARQL ou serviço mais avançado também pode ser muito útil, mas não substitue as formas de acesso mais básicas, como distribuição em massa, download de arquivos e outros.

Tabela 4.8 - Quantitativo de recomendações para "Prover acesso automatizado aos dados" nos processos de publicação de dados abertos analisados

Recomendações	Aparições nos processos analisados por Ávila (2015)	Aparições na análise dos novos processos	Total	Desvio em relação à média	Classificação
Estabelecer um mapa de decisões tecnológicas (8A)	1	3	4	-2,25	Desejável
Desenvolver uma API (8B)	2	3	5	-1,25	Desejável
Desenvolver um <i>endpoint</i> SPARQL (8C)	4	2	6	-0,25	Desejável
Disponibilizar bases completas para download (<i>dumps</i>) (8D)	4	6	10	3,75	Obrigatório
Media	2,75	3,5	6,25		

Fonte: Elaborado pelo autor.

O desenvolvimento de uma API (8A) para o acesso à informação é recomendado pelo o processo P1, P15, P19, P20 e P22. Embora os dados possam ser publicados de maneira estruturada, aberta e conectada em formas de acesso mais básicas como distribuição em massa, arquivos e outros. O uso de uma API pode otimizar o processo de recuperação e reuso dos dados pelos intermediários (Comunidade de Desenvolvedores, empreendedores e outros). No entanto, o processo P16 faz uma provocação com a pergunta “*O desenvolvimento de uma API é necessário?*”. Dependendo do caso, essa pode ser uma questão para ser tratada cuidadosamente na esfera governamental, por trazer um custo mais elevado, uma complexidade técnica mais elevada e benefícios duvidosos. Uma alternativa elencada pela maioria dos processos, considerada menos onerosa a este modelo, é utilizar técnicas mais básicas como “Disponibilizar bases completas para *download (dumps)* (8D)”. Mesmo o processo P19 defendendo o desenvolvimento de uma API, o mesmo reconhece que tal prática se torna secundária quando se trata da disponibilização completa dos dados brutos para *download*.

4.4.9 Recomendações para “Anunciar os novos conjuntos de dados para o público”

Na análise desta BPLD, 17 dos 22 processos apresentaram recomendações para anunciar os novos conjuntos de dados para o público. Conforme a Tabela 4.9, a adição de novos processos confirmou a unanimidade visualizada por ÁVILA (2015) para uma etapa de divulgação dos dados, apenas para cinco processos não foi possível extrair recomendações. Considerando os processos analisados, as recomendações que figuram como obrigatórias

foram: “Permitir que os dados sejam mais facilmente encontrados por máquinas (9A)”, “Estabelecer dados tecnicamente e legalmente abertos (9B)”, “Publicar metadados junto aos dados (9E)”, e “Divulgar dados em meios complementares (Catálogos, FTP, *Torrent*) (9H)”. Destaca-se a elevação de 9H para obrigatória e a queda de 9D para desejável.

Tabela 4.9 - Quantitativo de recomendações para "Anunciar os novos conjuntos de dados para o público" nos processos de publicação de dados abertos analisados.

Recomendações	Aparições nos processos analisados por Ávila(2015)	Aparições na análise dos novos processos	Total	Desvio em relação à média	Classificação
Permitir que os dados sejam mais facilmente encontrados por máquinas (9A)	3	4	7	1,888889	Obrigatória
Estabelecer dados tecnicamente e legalmente abertos (9B)	1	5	6	0,888889	Obrigatória
Disponibilizar dados com o menor custo possível ao usuário, preferencialmente de modo gratuito na internet (9C)	1	3	4	-1,111111	Desejável
Disponibilizar dados conectados em servidores de triplas (9D)	4	1	5	-0,111111	Desejável
Publicar metadados junto aos dados (9E)	6	1	7	1,888889	Obrigatória
Estabelecer recursos de consulta parcial da base de dados como uma API, <i>webservice</i> ou <i>endpoints</i> SPARQL (9F)	1	3	4	-1,111111	Desejável
Divulgar dados em seções destacadas de sítios de governo (9G)	3	1	4	-1,111111	Desejável

Recomendações	Aparições nos processos analisados por Ávila(2015)	Aparições na análise dos novos processos	Total	Desvio em relação à média	Classificação
Divulgar dados em meios complementares (Catálogos, FTP, <i>Torrent</i>) (9H)	3	4	7	1,888889	Obrigatória
Estabelecer visualizações e demais recursos de exploração dos dados (9I)	1	1	2	-3,11111	Desejável
Media	2,555555556	2,555555556	5,111111111		

Fonte: Elaborado pelo autor.

A recomendação de permitir que os dados sejam mais facilmente encontrados por máquinas (9A) é justificada no processo P17 pelo grande volume de informações, no qual os seres humanos não são capazes de extrair conhecimento a partir desses dados. Neste caso, a capacidade de processar e analisar os dados com máquinas é necessária. A recomendação “Estabelecer dados tecnicamente e legalmente abertos (9B)” é definida com os termos “Abertura Técnica” e “Abertura Legal dos Dados” no processo P17 e P18, ganhando uma seção específica. Do ponto de vista legal, o processo P22 faz uma ponte com o estabelecimento de uma licença que permita o acesso, uso e reúso dos dados de forma livre e irrestrita. Do ponto de vista técnico, a composição dos dados por formatos limitados tecnicamente, não são recomendados. Os dados devem ser liberados em formatos facilmente reutilizados. São dois fatores que influenciam esta abordagem para “Abertura Técnica”, grandes volumes de dados e a incapacidade dos seres humanos de tratar esses dados. Então, é imprescindível tanto para o processo P17, como para o processo P20, a recomendação de permitir que os dados sejam mais facilmente encontrados por máquinas (9A).

Para a recomendação “Divulgar dados em meios complementares (Catálogos, FTP, *Torrent*) (9H)”, os processos P16 e P19 destacam que ao redor do mundo as iniciativas estão tendo êxito ao reunir as bases de dados num catálogo central, promovendo praticidade para o cidadão, além de instrumentos para medição e monitoramento dessas bases pelo órgão publicador. Além do catálogo, o processo P17 cita como opções a disponibilização em sites próprios, sites de terceiros, via servidores FTP ou como uma API. No entanto, um órgão com

um nível de maturidade baixo, não precisa se assustar com esse quantitativo de opções e nem tão pouco implementar sistemas complexos. O P16 ressalta que uma página simples da *Web* com uma lista bem estruturada de documentos pode servir como um bom repositório de dados ou um catálogo de dados.

4.4.10 Recomendações para “Estabelecer um contrato social para os dados publicados”

Conforme Tabela 4.10, das cinco recomendações extraídas dos processos para esta BPLD, apenas “Estabelecer espaços para recebimento do *feedback* do usuário (10E)” despontou como a mais relevante. Cabe ressaltar que o avanço em relação às outras recomendações ocorreu pós-adição dos novos processos, até então a mesma encontrava-se bem distribuída com as demais, sendo considerada empatada tecnicamente.

Tabela 4.10 - Quantitativo de recomendações para "Estabelecer um contrato social para os dados publicados" nos processos de publicação de dados abertos analisados

Recomendações	Aparições nos processos analisados por Ávila (2015)	Aparições na análise dos novos processos	Total	Desvio em relação à média	Classificação
Disponibilizar leis e atos normativos que explicitem aos usuários quanto às obrigações dos governos em publicarem dados com qualidade e disponibilidade (10A)	1	2	3	-1	Desejável
Utilizar tecnologias que mantenham os dados conectados disponíveis, atualizados e abertos (10B)	1	2	3	-1	Desejável
Estabelecer com clareza que o processo de publicação contempla etapas de manutenção e atualização dos dados. (10C)	2	1	3	-1	Desejável
Estabelecer mecanismos de monitoramento e avaliação da oferta de dados disponibilizados ao público (10D)	1	2	3	-1	Desejável
Estabelecer espaços para recebimento do <i>feedback</i> do usuário (10E)	2	6	8	4	Obrigatória

Recomendações	Aparições nos processos analisados por Ávila (2015)	Aparições na análise dos novos processos	Total	Desvio em relação à média	Classificação
Media	1,4	2,6	4		

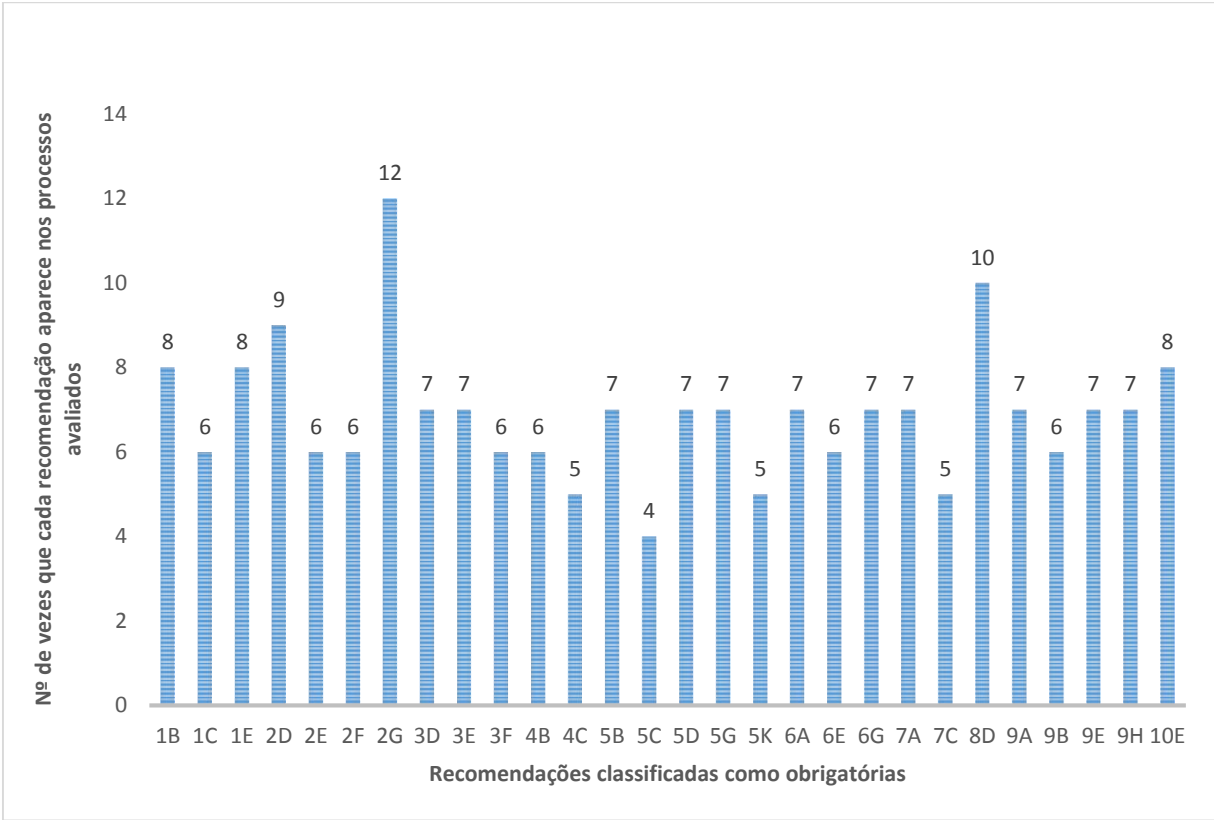
Fonte: Elaborado pelo autor.

Estratégias são propostas nos processos P17 e P20 para criação de um ecossistema sustentável de dados abertos conectados, elencando como um dos elementos fundamentais, a criação ou o desenvolvimento de um ambiente favorável de fomento ao diálogo e ao envolvimento dos cidadãos em torno dos Dados Abertos Conectados, ou seja, estabelecer espaços para recebimento do feedback do usuário (10E). Esta retroalimentação não pode ser vista como uma atividade que ocorre posteriormente a publicação, a fim de avaliar o conjunto de dados, e sim como um retorno da comunidade que poderá influenciar desde o planejamento na delimitação do escopo. Conforme o processo P19, a organização deverá encarar como uma atividade transversal por todo o processo, incorporando o sentimento do público desde o início no desenvolvimento de um plano de dados abertos; na priorização de conjuntos de dados, entre outras atividades.

4.5 Considerações Finais

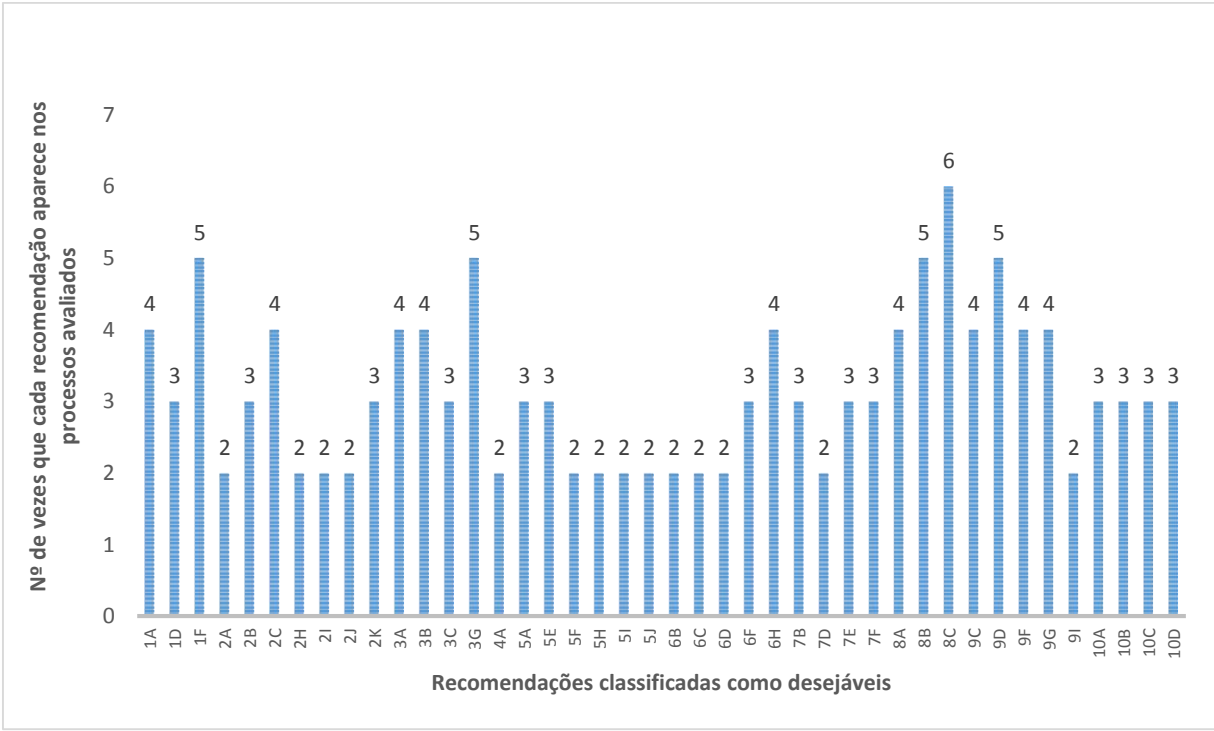
Os resultados da análise dos 22 processos para publicação de Dados Abertos Conectados elencaram 28 (vinte e oito) recomendações como “Obrigatórias” (Figura 4.2) e 42 (Quarenta e dois) como “Desejáveis” (Figura 4.3). O modelo “Piece of Cake” elenca 41 recomendações como “Obrigatórias” (ÁVILA, 2015). No entanto, ÁVILA (2015) não considerou apenas a metodologia utilizada neste capítulo para classificação. Para efeito de composição do modelo “Piece of Cake”, ÁVILA (2015) também utilizou estudos empíricos para complementar a classificação. Mas um novo subconjunto de recomendações obrigatórias pode ser resultado de uma discussão mais ampla. Tal fato ocorreu em nossos resultados com a adição dos novos processos, nos quais ocasionaram na redução dos quantitativos de recomendações obrigatórias. Como o processo proposto por essa pesquisa tem o objetivo de ser um piloto, apresentando um foco estritamente nas atividades essenciais, obtemos os resultados esperados (Redução do quantitativo das recomendações obrigatórias) da análise acometida neste capítulo.

Figura 4.2 - Sumarização dos resultados das recomendações classificadas como obrigatórias.



Fonte: Elaborado pelo autor.

Figura 4.3 - Sumarização dos resultados das recomendações classificadas como desejáveis.



Fonte:Elaborado pelo autor.

5 O Processo ODSP – Open Data Starter Process

Os resultados do estudo apresentado no terceiro capítulo, referente ao cenário dos Institutos Federais em relação a Dados Abertos, constataram nenhum envolvimento com as práticas para a publicação de Dados Abertos, além de não apresentarem nenhum planejamento para tal. Essas conclusões só foram possíveis a partir do levantamento em relação à publicação de Plano de Dados Abertos (PDA) nas instituições, além de pesquisa analítica nos portais e análise no atendimento a solicitações de informações de maneira passiva. No entanto, a chegada de legislações (Decreto Lei nº 8.638 e nº 8.777) com o objetivo de fomentar a disponibilização e o uso de Dados Abertos, pressionam tais instituições frente a temática. Tais instituições buscam se adequarem aos objetivos e prazos impostos pela nova legislação de Dados Abertos, e para isso tentam adotar processos de organizações similares, como também da literatura existente. A busca de um referencial é uma ação natural de organizações inexperientes em qualquer temática. Contudo, a ação de construção de um referencial é dificultada, quando as alternativas existentes para publicação de Dados Abertos, utilizam diferentes formas de linguagem, não apresentando uma padronização para o mesmo processo. Além disso, nenhuma das opções existentes apresenta uma notação ou linguagem específica de processo. Assim, como parte desta pesquisa, este capítulo apresenta a principal contribuição desta dissertação, ou seja, um processo, denominado ODSP (OPEN DATA STARTER PROCESS), proposto com intuito de guiar as referidas instituições na etapa embrionária da implantação de Dados Abertos. Esta divisão do trabalho apresenta de forma detalhada o processo, mas também mostra a concepção de criação e padronização do mesmo com base nos conceitos de BPM e na notação BPMN.

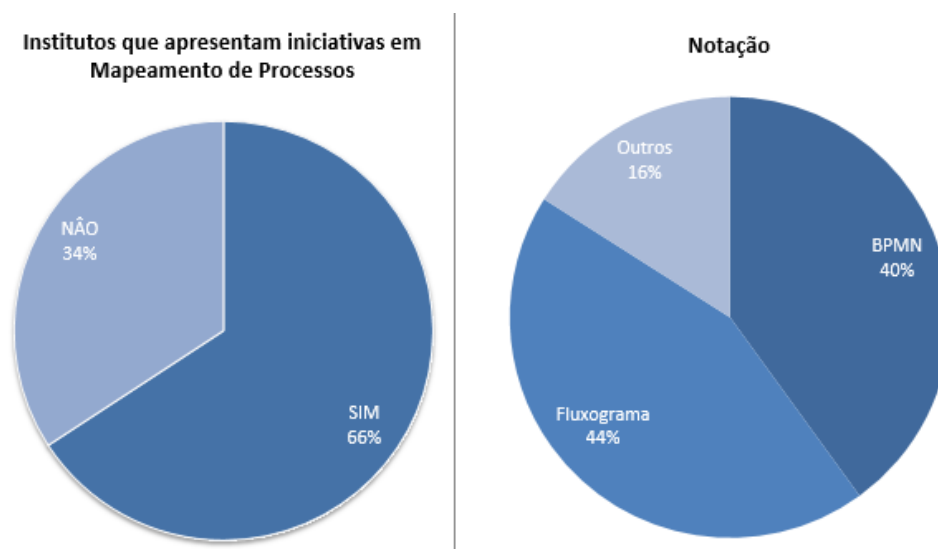
5.1 Visão geral do modelo ODSP

Com intuito de auxiliar no processo de abertura e conexão dos dados, diversas iniciativas têm sido desenvolvidas, incluindo a elaboração de metodologias, guias, padrões, leis, regulamentos e boas práticas. Por um lado, isso contribui para o amadurecimento da área de Dados Abertos. Por outro lado, esta variedade se torna prejudicial para uma instituição imatura, que pode ficar perdida meio a um turbilhão de possibilidades. Percebemos que mesmo utilizando notações e linguagens distintas, a maioria das alternativas visa à criação de fluxos de trabalho, servindo de referência na publicação de dados abertos ou dados

conectados. O Kit de Dados Abertos (BRASIL, 2014) ressalta que o processo tanto de planejamento como de execução na publicação de dados poderá ser organizado por meio de um fluxo de atividades (Processo). No entanto, nenhuma das opções existentes apresenta uma notação ou linguagem específica de processo.

Diante dessa necessidade, o processo ODSP propõe realizar uma espécie de curadoria de tudo que já foi proposto para facilitar a abertura de dados, porém com foco nos aspectos essenciais. Além disso, a fim de apresentar uma descrição em uma notação padrão, adotamos a abordagem de BPM (*Business Process Management*) para composição do processo. É importante destacar alguns fatores que foram determinantes para escolha de BPM, incluindo a observância nas políticas, premissas e especificações técnicas definidas pelo o Padrão de Interoperabilidade do Governo Eletrônico ePING, conforme portaria normativa SLTI nº 05, de 14 de julho de 2005, que define como padrão para processos, a Notação de Modelagem de Processo de Negócio (BPMN) visando à sustentação do intercâmbio de informações entre áreas transversais do governo. Outro determinante partiu dos arranjos das instituições para qual o processo é destinado. Os institutos federais têm um arranjo tradicional, o qual pode ser percebido em sua hierarquia com áreas verticalizadas. Cada área tem seus próprios objetivos e metas. Porém, o serviço ou produto final entregue ao usuário final é consequência da interação das diversas áreas que compõem a organização. Então, independente do que seja, o objetivo geral da organização deverá ser o foco, transversalizando os setores. Com isso, percebemos que os institutos mesmo não sendo uma organização centrada em processos, utilizam os mesmos para alcançar o objetivo geral, mantendo a verticalização para organização do pessoal. Algo perceptível ao observarmos os resultados do levantamento apresentado na Figura 5.1, demonstrando que 66% das instituições para qual o processo é destinado apresentam iniciativa em mapeamento de processos, e 40% utilizam a notação específica de BPM, facilitando a adoção do mesmo.

Figura 5.1 - Levantamento das iniciativas em Mapeamento de Processos nos Institutos Federais.



Fonte: Elaborado pelo autor.

O processo ODSP¹⁸ está estruturado em duas versões. Uma versão em PDF e outra versão interativa na Web, utilizando a concepção baseada na abordagem “*work in progress*”, ou seja, um processo inacabado que estará sempre em revisão e atualização, retroalimentado pelos os feedbacks oriundos das experiências do mundo real.

Para a representação do processo ODSP adotamos o padrão de Notação de Modelagem de Processo de Negócio (BPMN), desenvolvida pelo *Business Process Management Initiative* (BPMI), mantida pelo *Object Management Group* (OMG) e adotada pelo o ePING. No entanto, ePING utiliza a versão 1.2 da notação e estuda sua atualização para a 2.0, a versão utilizada pelo o processo. A utilização de BPMN facilita a adoção do processo por ser uma notação gráfica, segundo por conseguir capturar a logica das atividades, além das mensagens entre os vários participantes e todas as informações necessárias para o processo ser analisado, simulado e implementado.

Para a construção do processo, optamos pelo BPMS Bizagi¹⁹ pela sua superioridade em relação às outras em alguns critérios importantes para nossa proposta de processo, como: desempenho, permitir simulação do funcionamento do processo na versão gratuita, permitir publicação do processo via Web na versão gratuita e realizar exportação em múltiplos formatos na versão gratuita. A versão do Bizagi utilizada foi a Modeler.

¹⁸ <https://sites.google.com/cin.ufpe.br/odsp>

¹⁹ MODELER, BizAgi. Disponível em:< <http://www.bizagi.com/pt/produtos/bpm-suite/modeler>>.

Em relação à estrutura e documentação do processo, utilizamos o padrão adotado pela Secretaria de Logística e Tecnologia da Informação (SLTI) do Ministério do Planejamento, e apresentado no “Guia de Boas Práticas em Contratação de Soluções de Tecnologia da Informação” (MPOG, 2014). A escolha do padrão aconteceu devido os Institutos Federais serem órgãos integrantes do Sistema de Administração dos Recursos de Informação e Informática - SISP do Poder Executivo Federal, por isso estão condicionados aos padrões impostos pela SLTI.

5.2 Principais características do processo ODSP

Baseado no nível de Dados Abertos dos Institutos Federais, iniciou-se um trabalho de discussão, definição e documentação para o ODSP. Foram considerados os problemas e as fragilidades da situação atual, juntamente com as recomendações de outros processos existentes para publicação de Dados Abertos. O processo herda algumas características de outras propostas para facilitar a abertura de dados. No entanto algumas características se manifestam com mais força em relação a outras. Com isso, surge a necessidade de entender os fatores motivadores da herança. A herança das características do Kit de Dados Abertos do Governo Federal é motivada pelo o fato do processo ODSP ser um subproduto do processo “Iniciar piloto ágil de abertura de um conjunto de dados” presente no Kit, apresentando-se como um detalhamento de tal processo, que está em um nível macro e abstrato. Outras características predominantes são herdadas do processo de inovação, motivada pelo fato do nicho não conhecer ou vivenciar a temática de Dados Abertos, tendo o papel de catalisadoras no momento de adotar ou aderir à mudança proposta pelo o modelo. A seguir essas características serão explicadas.

5.2.1 Características herdadas do Kit de Dados Abertos do Governo Federal

O processo ODSP herda características experimentais e ágeis do processo “Iniciar piloto ágil de abertura de um conjunto de dados” presente no Kit de Dados Abertos, suprimindo diversas atividades que no momento embrionário da iniciativa de Dados Abertos são dispensáveis. A ideia principal é que a equipe possa experimentar, de forma muito rápida, todo o processo de publicação de Dados Abertos. Sendo essencial a reflexão após a execução do processo, aprendendo com erros e acertos que servirão como insumos para a instituição em relação ao tema e para construção de um Plano de Dados Abertos com estimativa de prazo,

escopo e custo mais realista. A apresentação do processo ODSP com características de um fluxo de trabalho também é reflexo do processo sistêmico representado graficamente por meio de fluxogramas no Kit de Dados Abertos. A escolha da notação BPMN também teve influências da representação adotada no kit.

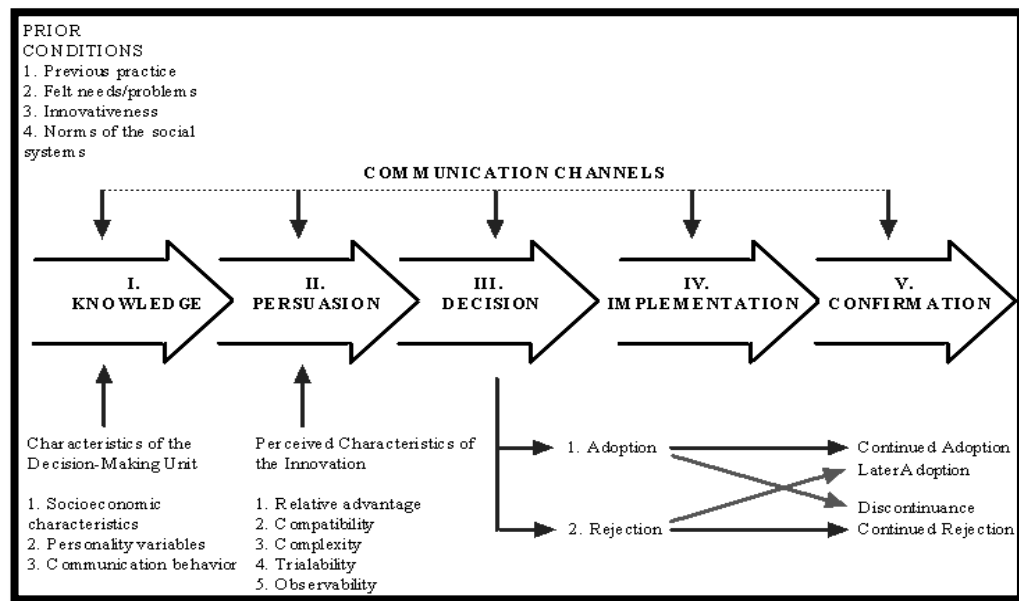
Além disso, os objetivos do processo ODSP são reflexos dos objetivos do processo “Iniciar piloto ágil de abertura de um conjunto de dados” do Kit de Dados Abertos do Governo Federal, que são:

- Ter, em curto prazo, um resultado prático de abertura de dados à disposição da sociedade.
- Preparar as equipes técnicas da instituição, familiarizando-as com os conceitos, terminologias, métodos e ferramentas necessários ao processo de abertura de dados.
- Iniciar o preparo da infraestrutura tecnológica de suporte aos dados abertos na instituição.
- Colher, a partir das lições aprendidas no processo, as necessidades para os futuros projetos de abertura de dados, tais como capacitação de pessoal, aquisição de infraestrutura, entre outros.

5.2.2 Características herdadas dos Processos de Inovação

A produção e publicação de dados abertos implicam em custos, mudanças culturais, institucionais, organizacionais e técnicas que se defrontarão com barreiras de adoção tanto para o órgão divulgador como para a comunidade de usuários (SÃO PAULO & REINO UNIDO, 2015). A OKF (2015) declara que uma iniciativa de dados abertos se trata de uma inovação. Por isso, ganhar celeridade no movimento é benéfico, visto que você ganha energia cinética e aprende com a experiência. A inovação se dá tanto pelo fracasso quanto pelo sucesso. Segundo ROGERS (2003) um processo de inovação é aquele no qual os tomadores de decisão conduzem uma formação de uma atitude em direção à inovação. No entanto, fatores podem influenciar a decisão de aprovar ou rejeitar a implementação da nova ideia. Então, nossa proposta mesclou alguns elementos do modelo de processo inovação-decisão proposto por ROGERS (2003) e representado pela figura 5.2.

Figura 5.2 - Modelo de processo inovação-decisão



Fonte: ROGERS, E. M. Diffusion of innovations.

Os indivíduos e tomadores de decisão, que nos Institutos Federais estão figurados como reitores, pró-reitores e diretores, precisam passar do primeiro estágio de conhecimento da inovação, para formar uma atitude em direção à inovação, que nesse caso seria a publicação de Dados Abertos. Então, no estágio da decisão, os tomadores de decisão chegariam à conclusão de adotar ou rejeitar a referida inovação.

5.3 Componentes do processo ODSP

Os componentes do processo ODSP podem ser: Fases, Subprocessos, Tarefas e Artefatos. Uma fase representa um estágio particular do processo de publicação de dados abertos. Os subprocessos são divididos em conjuntos de tarefas, de tal maneira que tarefas representam passos no processo de publicação. Para cada uma das tarefas são descritas orientações mais detalhadas sobre como as tarefas devem ser realizadas. Os artefatos são o último bloco de construção do modelo. Um artefato pode ser tanto de entrada como de saída, ao mesmo tempo. Normalmente, um artefato de saída de uma tarefa pode contribuir para uma tarefa subsequente.

A seguir, apresentamos detalhes de como foram identificados os subprocessos, as tarefas e as fases do processo ODSP.

5.3.1 Subprocessos e tarefas do processo ODSP

O Capítulo 4 demonstra os resultados da análise de 22 processos para publicação de Dados Abertos Conectados, servindo de referencial para construção do processo ODSP. Para cada processo, foram identificadas recomendações. De acordo com reincidência dessas recomendações em vários processos, as mesmas foram classificadas em obrigatórias e desejáveis. Como o processo tem o objetivo de ser um piloto, focamos nas tarefas essenciais que devem ser executadas por qualquer instituição publicadora de Dados Abertos, independente do nível de maturidade. Foram encontradas vinte e oito recomendações obrigatórias, das quais quinze foram reaproveitadas e alocadas em nosso processo como subprocessos, enquanto cinco foram classificadas como tarefas. Essas cinco tarefas se juntaram as outras sessenta e quatro tarefas elencadas no Quadro 5.2. As outras tarefas são oriundas dos passos que os processos traziam para atender a cada recomendação.

As tarefas são representadas no processo ODSP por retângulos arredondados e representam as tarefas que são desenvolvidas por um ator. Quando uma tarefa é finalizada, a próxima tarefa inicia seguindo o fluxo do processo. Os subprocessos são tarefas compostas representadas no processo ODSP por retângulos arredondados e o símbolo de soma no centro da figura. Devido à sua complexidade, os subprocessos representam um nível acima na hierarquia, apresentando um conjunto de tarefas para alcançar um determinado objetivo. O Quadro 5.1 elenca a classificação de cada recomendação no processo.

Quadro 5.1 - Classificação das recomendações no processo.

Recomendação	Classificação
1. Identificar partes interessadas	Subprocesso
2. Definir profissionais a serem envolvidos	Subprocesso
3. Elaborar um plano de ações para publicação dos dados	Subprocesso
4. Estabelecer diretrizes que orientem a priorização de dados a serem abertos	Subprocesso
5. Analisar o nível de sigilo dos dados a serem abertos	Subprocesso
6. Realizar consultas aos usuários sobre a demanda de dados	Subprocesso
7. Identificar os dados que serão abertos	Subprocesso

Recomendação	Classificação
8. Higienizar os dados	Subprocesso
9. Anonimizar dados sensíveis	Subprocesso
10. Estabelecer rotinas de conversão de dados para formatos legíveis por máquinas	Tarefa
11. Adotar licenças não restritivas	Tarefa
12. Apresentar opções de licenças a serem adotadas	Subprocesso
13. Usar URIs como nomes para as coisas	Não utilizada no processo
14. Utilizar identificadores relacionados a informações do mundo real	Não utilizada no processo
15. Usar URIs HTTP para que pessoas e máquinas possam encontrá-las via web utilizando estes endereços	Não utilizada no processo
16. Estabelecer URIs persistentes, que não se alterem em nenhum momento	Não utilizada no processo
17. Utilizar URIs para conectar os dados	Não utilizada no processo
18. Estabelecer metadados obrigatórios	Subprocesso
19. Incentivar o reuso de vocabulários	Tarefa
20. Desenvolver ou utilizar ontologias para estruturar a semântica dos dados	Subprocesso
21. Utilizar rotinas automatizadas de conversão de dados, como a triplicação, quando possível	Tarefa
22. Conectar conjuntos de dados com outros dados relacionados	Não utilizada no processo
23. Disponibilizar bases completas para download (dumps)	Subprocesso
24. Permitir que os dados sejam mais facilmente encontrados por máquinas	Recomendação dentro de uma tarefa
25. Estabelecer dados tecnicamente e legalmente abertos	Recomendação dentro de uma tarefa

Recomendação	Classificação
26. Publicar metadados junto aos dados	Tarefa
27. Divulgar dados em meios complementares (Catálogos, FTP, Torrent)	Subprocesso
28. Estabelecer espaços para recebimento do feedback do usuário	Subprocesso

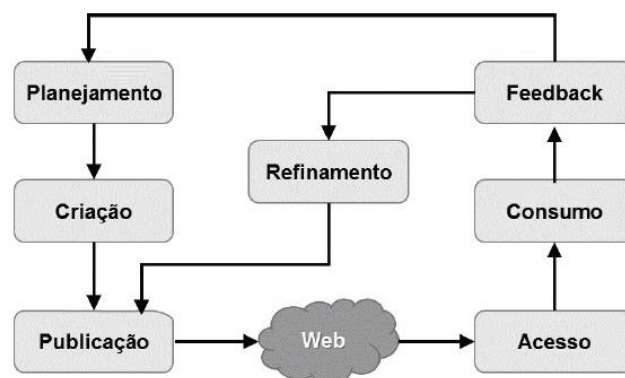
Fonte: Elaborado pelo autor.

É importante notar que seis recomendações não foram reaproveitadas porque, considerando a maturidade do nicho para qual o modelo está sendo proposto, a conexão dos dados poderia criar um nível adicional de complexidade para esta fase inicial, comprometendo os objetivos do processo.

5.3.2 Fases do processo ODSP

Com intuito de organizar as tarefas e os subprocessos que compõem o processo ODSP, apresentados na Tabela 5.1, utilizamos o ciclo de vida proposto por LÓSCIO; OLIVEIRA & BITTENCOURT (2015). Para LÓSCIO; OLIVEIRA & BITTENCOURT (2015), um ciclo de vida é um conjunto de fases que compõem o processo de publicação e consumo de dados. Os autores trazem uma proposta de ciclo de vida dos dados na *Web* (Figura 5.3), o qual, além de descrever as principais fases envolvidas no processo de publicação e consumo de dados, também oferece uma terminologia comum que auxilia a comunicação entre provedores e consumidores de dados. Dessa forma, adotamos as fases do Ciclo de Vida dos Dados na *Web* como as fases do processo ODSP.

Figura 5.3 - Ciclo de Vida dos Dados na Web



Fonte: LÓSCIO; OLIVEIRA & BITTENCOURT (2015)

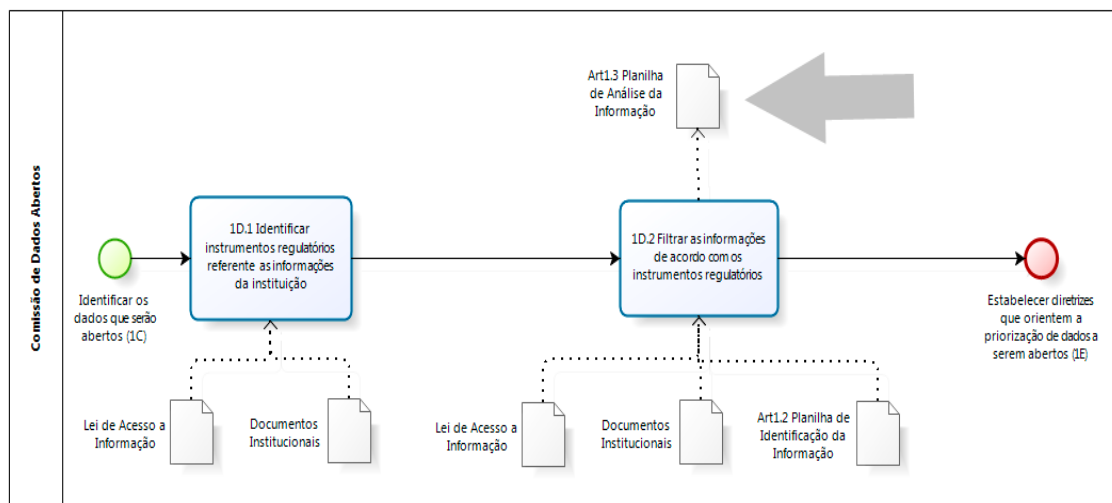
Sendo assim, cada um dos subprocessos identificados no quadro 5.1 foram alocados nas fases Planejamento, Criação e Publicação. Na fase de Planejamento consideramos os subprocessos que surgem desde a intenção de publicar até a publicação de um plano de ação para execução do processo de abertura na instituição. A fase de Criação compreende os subprocessos relacionados com a criação dos dados, desde sua extração até a sua transformação para o formato adequado para publicação na *Web*. A fase de Publicação abriga os subprocessos responsáveis pela disponibilização dos dados de forma pública na *Web*. A fase de Acesso é referente ao momento em que os usuários ganham acesso aos dados. A fase de Consumo é o momento em que os dados são usados pelo público para criação de visualizações, aplicações, gráficos entre outros. O *feedback* compreende o recebimento de comentários dos usuários referentes aos dados publicados. Realizamos a junção das fases acesso, consumo e *feedback* na fase publicação, visto que os quantitativos de recomendações encontradas não foram suficientes para povoar cada uma delas. A fase de Refinamento compreende as atividades relacionadas a adições ou atualizações nos dados posteriores à publicação. Esta fase não foi considerada em nosso processo, por não termos recomendações obrigatórias com estes objetivos.

Sabemos do caráter finito de um projeto piloto, o qual gerará insumos para processos posteriores. No entanto, não impede de se enxergar o processo como um ciclo, visto que após a primeira interação, a instituição poderá adotar como seu processo permanente adicionando alguns elementos para o enriquecimento dos dados.

5.3.3 Artefatos do processo ODSP

O artefato prover informações de como documentos, dados e outros objetos são usados e atualizados durante o processo, sendo representado em nosso processo por uma folha de papel (Figura 5.4). Utilizamos vários artefatos incluindo manuais, modelos de documentos e outros. No entanto, o mais relevante é a planilha herdada do “Guia para la apertura de datos em Colombia”

Figura 5.4 - Alocação dos artefatos no modelo.



Fonte: Elaborado pelo autor.

O guia do Governo Colombiano associou o fluxo de trabalho com uma ferramenta auxiliar a partir da qual a instituição documenta as informações pertinentes a cada passo do processo. A cada etapa do guia, a instituição faz uso de um documento (Figura 5.5) no formato XLS contendo várias abas, as quais auxiliam o processo de Identificação da Informação, Análise da Informação, Priorização dos Dados e Descrição dos Metadados. O próprio guia vislumbra que tal instrumento pode ser adaptado pelas instituições com o propósito de criar elementos alinhados a cada legislação, permitindo um suporte jurídico. Então, incorporamos o referido instrumento ao nosso processo utilizando quatro passos: i) Realizando a tradução do idioma da língua espanhola para língua portuguesa; ii) Realizando as modificações gráficas para identidade visual dos Institutos Federais; iii) Alinhando os critérios utilizados pelo o instrumento com a Lei de Acesso à Informação do Brasil; e por último, iv) incorporando o instrumento no processo como artefato de entrada ou saída das tarefas do processo.

Figura 5.5 - Planilha de Abertura dos Dados.



Ministerio de Tecnologías de la Información y las Comunicaciones
República de Colombia
www.mintic.gov.co

Plantilla para la Identificación de la Información

Información de la entidad

Entidad:

Sector:

Nombre del funcionario(s):

Área(s) responsable(s):

Información de contacto del funcionario(s)

Teléfono:

E-mail:

Nombre de la información	Descripción	Área responsable de la información	Tipología de Información	Ámbito geográfico	Idioma	Fuente Primaria
			---	---	---	---
			---	---	---	---
			---	---	---	---
			---	---	---	---

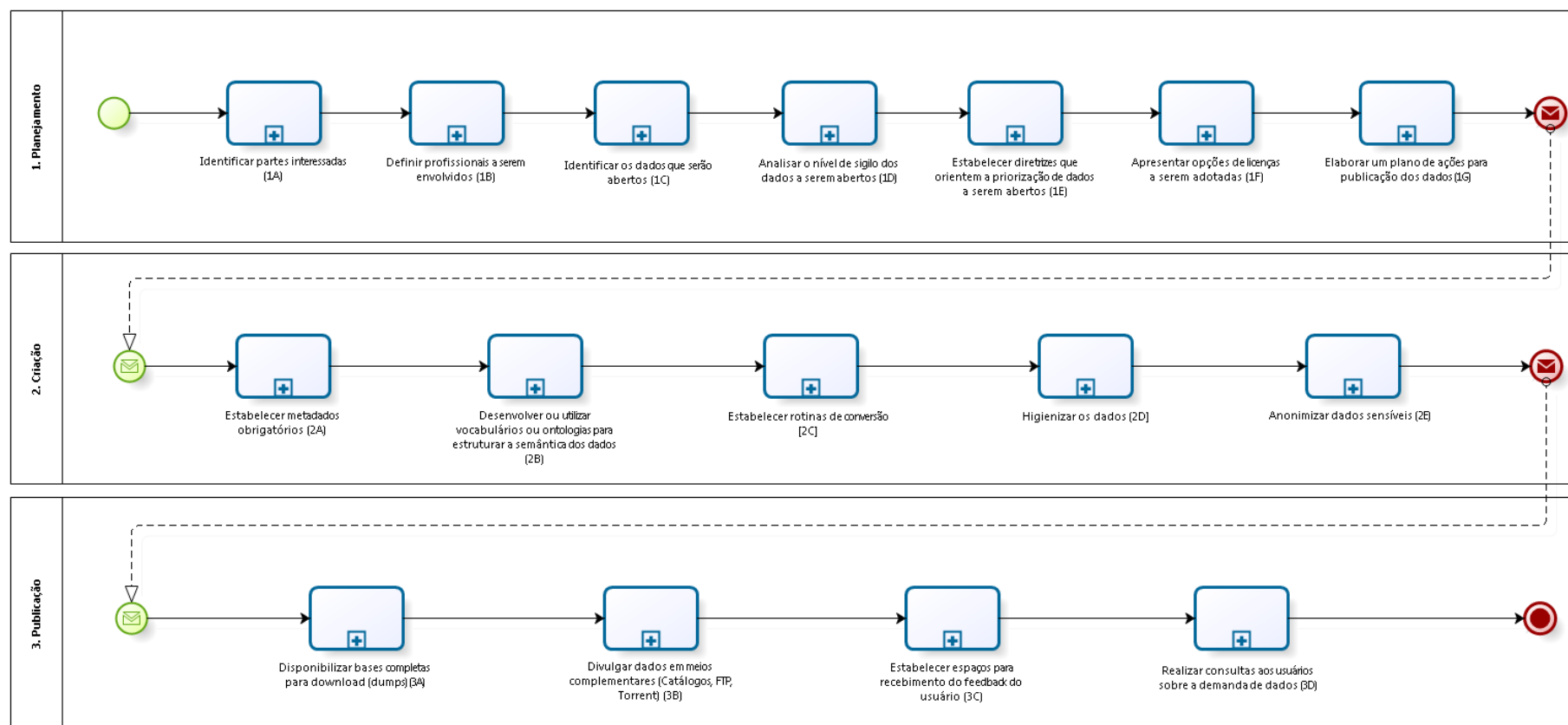
Hoja de Control
Generalidades
Identificación de información
Análisis de la información
Priorización de los datos
Inventario de datos

Fonte: Guía para la apertura de datos em Colombia.

5.4 Estrutura geral do processo ODSP

Considerando as fases, subprocessos, tarefas e artefatos apresentados nas seções anteriores, a Figura 5.6 ilustra a visão global do processo ODSP. Para cada uma das fases identificadas são descritos os subprocessos que compõem aquela fase. Em seguida, o Quadro 5.2 detalha as tarefas e os artefatos de cada subprocesso.

Figura 5.6 - Visão Global do Processo ODSP



Fonte: Elaborado pelo autor.

Quadro 5.2 - Sequenciamento de fases, subprocessos, tarefas e artefatos do processo.

Fases	Subprocessos	Tarefas	Artefatos
1. Planejamento	1A. Identificar partes interessadas	1A.1 Identificar <i>early-adopters</i>	Art 1. Organograma da instituição
		1A.2 Identificar partes envolvidas	
		1A.3 Solicitar ao Dirigente Geral um momento para apresentação da iniciativa	Art 2. Memorando de Solicitação para Apresentação da Iniciativa
		1A.4 Convidar <i>early-adopters</i> e partes envolvidas para a reunião	
		1A.5 Realizar a reunião de apresentação da iniciativa	Art 3. Legislação Art 4. Documentos institucionais Art 5. Exemplos de iniciativas Art 6. Materiais auxiliares Art 7. Ata de reunião
		1A.6 Aprovar implementação da iniciativa de Dados Abertos	
		1A.7 Cuidar dos receios e mal-entendidos	Art 7. Ata de reunião
	1B. Definir profissionais a serem envolvidos	1B.1 Indicar integrante técnico	
		1B.2 Indicar integrante do SIC	
		1B.3 Indicar integrante de comunicação	

Fases	Subprocessos	Tarefas	Artefatos
		1B.4 Indicar integrante da ouvidoria	
		1B.5 Indicar integrante pesquisador institucional	
		1B.6 Indicar integrante administrativo	
		1B.7 Indicar integrante acadêmico	
		1B.8 Indicar outros interessados na composição da Comissão de Dados Abertos Conectados	
		1B.9 Definir ponto focal	
		1B.10 Instituir Comissão de Dados Abertos	Art 8. Portaria de nomeação da Comissão de Dados Abertos
	1C. Identificar os dados que serão abertos	1C.1 Realizar levantamento de dados, junto aos setores	Art 9. Planilha de identificação da informação
		1C.2 Obter lista com os dados pedidos pelo SIC	
		1C.3 Obter Lista de Sistemas	
		1C.4 Inventariar Dados	
	1D. Analisar o nível de sigilo dos dados a serem abertos	1D.1 Identificar instrumentos regulatórios referente as informações da instituição	Art 4. Documentos institucionais Art 10. Lei de Acesso a Informação
		1D.2 Filtrar as informações de acordo com os instrumentos	Art 4. Documentos institucionais

Fases	Subprocessos	Tarefas	Artefatos
		regulatórios	Art 9. Planilha de identificação da informação. Art 10. Lei de acesso a informação Art 11. Planilha de análise da informação
	1E. Estabelecer diretrizes que orientem a priorização de dados a serem abertos	1E.1 Identificar impactos na publicação dos dados	Art 12. Planilha de priorização dos dados
		1E.2 Identificar dificuldades na publicação dos dados	
		1E.3 Realizar cálculo de priorização dos dados	
	1F. Apresentar opções de licenças a serem adotadas	1F.1 Identificar as licenças existentes	Art 13. Licenças
		1F.2 Adotar licenças não restritivas	
		1F.3 Criar licença própria	
		1F.4 Associar licença aos dados	Art 14. Planilha de identificação de licença
	1G. Elaborar um plano de ações para publicação dos dados	1G.1 Identificar e planejar recursos tecnológicos	Art 15. PDTI Art 16. Soluções livres disponíveis no Mercado Art 17. Soluções proprietárias disponíveis no

Fases	Subprocessos	Tarefas	Artefatos
			Mercado Art 18. Projetos Similares Art 19. Padrões Tecnológicos
		1G.2 Identificar e planejar recursos financeiros	
		1G.3 Realizar uma oficina para construção do Plano de Ações para publicação dos Dados Abertos	Art 20. Manual para elaboração de PDA
		1G.4 Identificar e planejar os recursos humanos	
		1G.5 Definir cronograma e marcos	
		1G.6 Elaborar o Plano de Ações para publicação dos Dados Abertos	Art 20. Manual para elaboração de PDA Art 21. Plano de ações para publicação dos Dados Abertos
		1G.7 Realizar reunião de alinhamento com partes interessada	
2. Criação	2A. Estabelecer metadados obrigatórios	2A.1 Identificar os conjuntos de dados a documentar	Art 12. Planilha de priorização dos dados
		2A.2 Identificar padrões/formatos para publicação das informações	
		2A.3 Estabelecer os metadados comuns	

Fases	Subprocessos	Tarefas	Artefatos
		2A.4 Estabelecer os metadados específicos	
		2A.5 Documentar os metadados	Art 21. Planilha de descrição dos metadados
	2B. Desenvolver ou utilizar vocabulários ou ontologias para estruturar a semântica dos dados	2B.1 Solicitar esquema e documentação da base de dados	
		2B.2 Procurar vocabulários existentes para reutilização	Art 22. Vocabulários e ontologias
		2B.3 Reusar vocabulários	
		2B.4 Criar vocabulários	
	2C. Estabelecer rotinas de conversão	2C.1 Definir formato e padrão de publicação	
		2C.2 Realizar extração dos dados	
		2C.3 Estabelecer rotinas de conversão de dados para formatos legíveis por máquinas.	
		2C.4 Utilizar rotinas automatizadas de conversão de dados.	
	2D. Higienizar os dados	2D.1 Familiarizar com o conjunto de dados	
		2D.2 Identificar e encontrar possíveis erros	
		2D.3 Consertar erros identificados	
	2E. Anonimizar dados	2E.1 Obter Consentimento	

Fases	Subprocessos	Tarefas	Artefatos
	sensíveis	2E.2 Subdividir os dados	
		2E.3 Remover os dados sensíveis	
		2E.4 Sintetizar os dados	
3. Publicação	3A. Disponibilizar bases completas para download (<i>dumps</i>)	3A.1 Criar repositório de dados	
		3A.2 Publicar conjunto de dados	
		3A.3 Publicar licença junto aos dados	Art 14. Planilha de identificação de licença
		3A.4 Publicar metadados junto aos dados	Art 21. Planilha de descrição dos metadados
	3B. Divulgar dados em meios complementares (Catálogos, FTP, <i>Torrent</i>)	3B.1 Solicitar cadastramento para publicação no Catálogo Nacional	
		3B.2 Catalogar conjunto de Dados no dados.gov.br	Art 21. Planilha de descrição dos metadados Art 23. Manual de catalogação
	3C. Estabelecer espaços para recebimento do feedback do usuário	3C.1 Definir estratégia para recebimento do feedback do usuário	
		3C.2 Definir canal de comunicação	
		3C.3 Divulgar canal de comunicação	
	3D. Realizar consultas aos usuários sobre a	3D.1 Preparar lista de Conjunto de Dados Potenciais para	Art 9. Planilha de identificação da

Fases	Subprocessos	Tarefas	Artefatos
	demanda de dados	Publicação	informação Art 12. Planilha de priorização dos dados
		3D.2 Realizar uma consulta pública	
		3D.3 Divulgar a consulta pública	

Fonte: Elaborado pelo autor.

5.5 Considerações Finais

Por fim, apresentamos nesta seção uma breve análise comparativa entre o processo ODSP e alguns processos relacionados. Especificamente, o Quadro 5.3 relaciona alguns processos e faz uma comparação com o processo ODSP. O referencial comparativo estabelecido é composto de critérios que, de alguma forma, auxiliam os iniciantes na execução das atividades para abertura dos dados, são eles: *Possui notação gráfica? Possui notação de processo? Fornece ferramentas, planilhas ou formulários para auxílio durante o processo? Possui uma versão interativa? Viabiliza uma percepção intuitiva?* A escolha do Brasil como recorte geográfico para seleção dos processos da análise comparativa justifica-se pela proposta se tratar de um processo voltado para instituições brasileiras. Com isso, a análise apresentada a seguir considerou os processos de publicação de Dados Abertos Governamentais nacionais.

Quadro 5.3 - Quadro comparativo dos processos de publicação de Dados Abertos Governamentais nacionais.

	ODSP	Guia de Dados Abertos	Manual dos Dados Abertos	Piece of Cake
Possui notação gráfica?				
Possui notação de processo?				

	ODSP	Guia de Dados Abertos	Manual dos Dados Abertos	Piece of Cake
Fornece ferramentas, planilhas ou formulários para auxílio durante o processo?				
Possui uma versão interativa?				
Viabiliza uma percepção intuitiva?				

Fonte: Elaborado pelo autor.

O ODSP favorece o entendimento das pessoas com pouco conhecimento sobre a notação, por utilizar uma notação gráfica e possuir uma forte abordagem de processo. No entanto, a descrição literal utilizada pelos outros processos, limita a capacidade de expressar adequadamente os processos em outros níveis de complexidade. Tal limitação foi amenizada com a adoção de BPMN no ODSP, oferecendo uma maior expressividade e detalhamento, devido à disponibilidade de elementos gráficos. Além disso, o processo ODSP é o único processo aderente aos conceitos de BPM entre todos.

Em relação às ferramentas, planilhas ou formulários para auxílio durante o processo, o processo ODSP fornece vários documentos editáveis e adaptáveis à instituição, com o objetivo de auxiliar o indivíduo no processo. Enquanto o “Piece of Cake” disponibiliza um único artefato que auxilia no processo, além de se tratar apenas de uma proposta de questionário para identificar a experiência da instituição e equipe em relação a publicação de Dados Abertos, indicando o melhor ponto do processo no qual a instituição deve iniciar. Os outros processos “Guia de Dados Abertos” e o “Manual de Dados Abertos” apresentam desvantagem por fazerem uso de um documento estático PDF, tornando-se uma alternativa limitada. O processo ODSP também oferece uma versão em PDF, motivada em ser uma referência para o usuário situado em algum local que não disponha de Internet. Mas um dos pontos fortes do processo ODSP, é ter uma versão *Web* interativa²⁰, facilitando o acesso e dando celeridade para o entendimento de todo o processo pelo usuário. Diante do exposto, podemos concluir que o processo ODSP traz diversos elementos facilitadores, tornando-se a melhor opção para os órgãos inexperientes na temática de Dados Abertos.

²⁰ <https://sites.google.com/cin.ufpe.br/odsp>

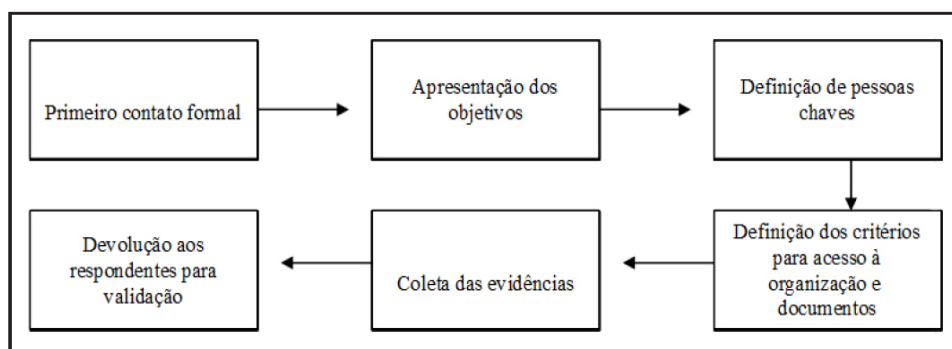
6 Avaliação

Este capítulo apresenta resultados que comprovam, por meio de uma estratégia de avaliação, que o processo proposto é executável e resulta nas saídas esperadas. A avaliação foi realizada por meio de uma investigação empírica do processo dentro de um contexto real. A estratégia de pesquisa consistiu em um estudo de caso no Instituto Federal de Pernambuco (IFPE) - Campus Ipojuca, permitindo pôr em ação a execução do processo. Todo o trabalho foi definido e documentado na forma de um fluxo de trabalho.

6.1 Metodologia para a condução do Estudo de Caso

O estudo de caso teve como objetivo avaliar o funcionamento do processo como um todo, tendo como finalidade testar a nova proposta, caracterizando a pesquisa como descritiva. Após a definição da finalidade do estudo de caso, fez-se necessário adotar o tipo de abordagem mais apropriada para atingir os objetivos da avaliação. Dentre as abordagens tradicionais (Qualitativa e Quantitativa), validamos as evidências obtidas por meio de um *mix* das duas abordagens. Tal estratégia de mesclagem é conhecida como triangulação metodológica ou *mixed-methodology* (FREITAS & JABBOUR, 2011), baseada no uso combinado e sequencial de uma fase de pesquisa quantitativa seguida de uma fase qualitativa, ou vice-versa. Para a condução do estudo de caso na instituição escolhida, foi efetuado um planejamento operacional, que consistiu das seis etapas (Figura 6.1) propostas por FREITAS & JABBOUR (2011):

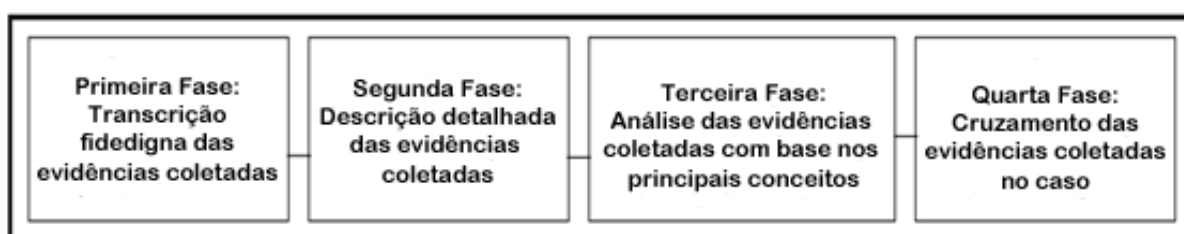
Figura 6.1 - Processo de planejamento da coleta de dados e evidências do estudo de caso.



Fonte: FREITAS & JABBOUR (2011).

Os instrumentos e técnicas de obtenção de dados e evidências utilizados durante a condução do estudo foram questionários, consulta a arquivos, análise de documentos, observação e conversas informais. Cabe ressaltar que o pesquisador em alguns momentos foi colocado dentro do contexto estudado, possibilitando a observação de fatos, comportamentos e objetos ligados ao problema em estudo, mesmo que obtidos informalmente. A última etapa consistiu na análise dos dados e das evidências coletadas, identificando temas e relações, construindo interpretações e gerando novas questões. Em nosso trabalho, utilizamos o roteiro (Figura 6.2) para o processo de análise proposto por FREITAS & JABBOUR (2011).

Figura 6.2 - Processo de análise das evidências coletadas em estudos de casos.



Fonte: FREITAS & JABBOUR (2011).

6.2 Estudo de caso

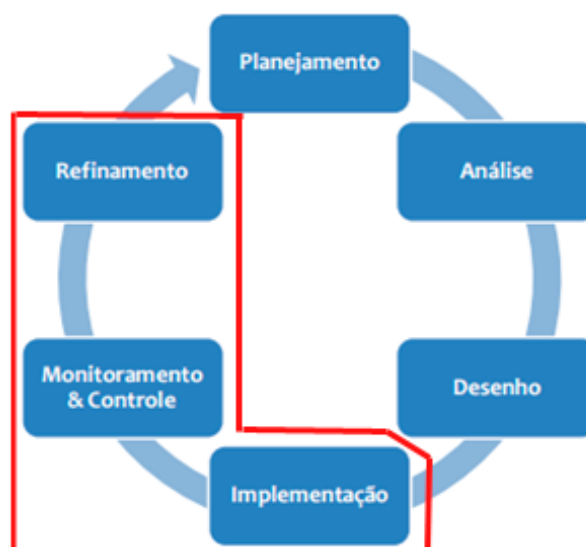
O tipo de estudo utilizado é de caso único no Instituto Federal de Pernambuco (IFPE) - Campus Ipojuca situado na Rodovia PE-60, km 14 - Bairro: Califórnia CEP: 55590-000. As informações dos perfis dos servidores envolvidos no processo estão descritas no quadro 6.1.

Quadro 6.1 - Perfil dos Servidores do IFPE - Campus Ipojuca participantes do estudo de caso.

CARGO	FUNÇÃO	DEPARTAMENTO	CONTATO
Assistente Administrativo	Diretor de Administração e Planejamento	Direção de Administração e Planejamento	dap@ipojuca.ifpe.edu.br
Assistente Administrativo	Coordenador	Divisão de Execução Orçamentária e Financeira (DEOF)	financeiro@ipojuca.ifpe.edu.br

CARGO	FUNÇÃO	DEPARTAMENTO	CONTATO
Administrador	Coordenador	Coordenação de Contratos e Convênios	contratos@ipojuca.ifpe.edu.br
Técnico em Tecnologia da Informação	Pesquisador Institucional	Coordenação de Gestão e Tecnologia da Informação	cgti@ipojuca.ifpe.edu.br
Técnico em Tecnologia da Informação	Coordenador de TI eventual	Coordenação de Gestão e Tecnologia da Informação	cgti@ipojuca.ifpe.edu.br
Coordenador	Coordenador	Assessoria de Comunicação e Eventos	asce@ipojuca.ifpe.edu.br
Jornalista	Jornalista	Assessoria de Comunicação e Eventos	asce@ipojuca.ifpe.edu.br
Assistente Administrativo	Coordenador	Coordenação de Estágios e Egressos	ceeg@ipojuca.ifpe.edu.br

O estudo de caso foi guiado de acordo com as fases (Implementação; Monitoramento e Controle; Refinamento) do ciclo de vida típico para processo (Figura 6.3) apresentado pelo CBOK (2013). Na fase de “Implementação”, aplicamos o processo na organização de acordo com as especificações desenvolvidas na fase “Desenho”. Na fase “Monitoramento e Controle”, extraímos os resultados do processo em comparação ao esperado, além de documentar todas as dificuldades encontradas e as lições aprendidas pela equipe IFPE – Campus Ipojuca. Na fase de “Refinamento”, corrigimos o processo ODSP de acordo com os dados coletados nas evidências do estudo durante a fase de “Monitoramento e Controle”. O Quadro 6.2, mostra as observações pontuadas durante o estudo de caso e as ações tomadas para correção no processo ODSP. Durante aplicação do processo, a equipe produziu diversas atas de reuniões. Na análise das evidências pelo o autor deste trabalho, foram extraídas as observações dessas atas de reuniões.

Figura 6.3 - Fases do Ciclo de vida BPM típico para processos utilizadas no Estudo de Caso

Fonte: Adaptado do CBOK (2013).

Quadro 6.2 - Observações pontuadas e ações tomadas para correção do processo ODSP.

Observações pontuadas	Ações de correção no processo ODSP
Dificuldades na edição dos artefatos, visto que os mesmos estão com permissão de visualização.	Criação de uma seção com tutoriais, incluindo o tutorial de como realizar a cópia do artefato, tornando-o editável.
Artefatos não permitem o inventário das informações com múltiplos formatos.	Adição de colunas para mais formatos no artefato “Identificação da informação”.
Dificuldades na identificação dos formatos dos conjuntos de dados por pessoas leigas.	Associação da extensão do arquivo com seus referidos softwares.
Dificuldades de entender critérios utilizados no artefato de “Análise da Informação” e “Priorização dos Dados.”	Criação de uma seção “Critérios utilizados no processo ODSP”.
Na classificação das informações, surgiram dificuldades em encontrar as categorias desejadas.	Expansão dos níveis elencados do Vocabulário Controlado do Governo Eletrônico (e-sic).
Surgiram dificuldades na utilização de links	Transformação de todos os links externos do

Observações pontuadas	Ações de correção no processo ODSP
externos, visto que são apresentados na descrição das atividades como texto e não como hiperlinks.	modelo em hiperlinks.
Textos das opções de licenças indicadas pelo o processo estavam em inglês.	Alteração das licenças disponíveis em outros idiomas para o Português.
Foi ressaltada a importância do envio do termo de licença para análise da Procuradoria Jurídica.	Adição de uma atividade no subprocesso “1F. Apresentar opções de licenças a serem adotadas”, com intuito de encaminhamento a Procuradoria Jurídica.
E-mail enviado para contato@dados.gov.br com intuito de cadastramento no dados.gov.br, retornando com erro.	Adição de uma atividade no subprocesso “3B. Divulgar dados em meios complementares (Catálogos, FTP, <i>Torrent</i>)”, com intuito de oferecer canais de suporte para quando a atividade “3B.1 Solicitar cadastramento para publicação no Catálogo Nacional” falhar por algum motivo.

Fonte: Elaborado pelo autor.

Os vários questionamentos presentes nas observações da equipe foram relacionados ao desconhecimento dos critérios utilizados nos artefatos de “Análise da Informação” e “Priorização dos Dados”. Além do mais, tais artefatos ofereceram uma maior dificuldade em relação aos outros na adequação da planilha de abertura dos dados do “Guía para la apertura de datos em colombia” para o contexto do processo ODSP. Por esta razão, o processo ganhou uma seção “Critérios utilizados no processo ODSP”. Os critérios utilizados nos artefatos de “Análise da Informação” e “Priorização dos Dados” estão explicitados nos apêndices A e B, respectivamente.

6.2.1 Execução do Processo

Esta seção permite uma visão do caminho e de todas as inter-relações vivenciadas pela a equipe durante o processo de abertura de dados, evidenciando as saídas de cada etapa e

trazendo uma discussão a cerca dos resultados. Cabe ressaltar que foi aplicado um questionário para os envolvidos no processo, com objetivo de levantar da equipe o conhecimento de opiniões, crenças, sentimentos, interesses, expectativas ou situações vivenciadas em relação à temática, antes e após a execução do processo. O questionário utilizou a escala Likert de cinco níveis (Muito baixo, Baixo, Médio, Alto e Muito alto), considerando as seguintes perguntas:

1. *Avalie seu nível de conhecimento sobre Dados Abertos ou Dados Abertos Conectados?*
2. *Avalie seu nível de conhecimento sobre a legislação em vigor que institui a Política de Dados Abertos do Poder Executivo federal. (Decreto 8.777, de 11 de maio de 2016)?*
3. *Avalie seu nível de entendimento sobre a necessidade de se ter um Plano de Dados Abertos (PDA) na instituição?*
4. *Avalie seu nível de conhecimento sobre a legislação em vigor que institui a Política de Governança Digital no âmbito dos órgãos e das entidades da administração pública federal direta, autárquica e fundacional. (Decreto 8.638, de 11 de maio de 2016)?*
5. *Avalie seu nível de conhecimento sobre a legislação em vigor sobre acesso a informação (Lei 12.527, de 18 de novembro de 2011)?*

Em seguida, foi possível transformar as opções da escala em valores numéricos assumindo valores entre 1 a 5. Os resultados obtidos consideraram a média das respostas de todos respondentes para cada questão e a média geral para todas as questões. As evidências (Atas de reunião, Artefatos e Planilhas) do estudo de caso estão disponíveis em um repositório na Web²¹.

As atividades dos subprocessos “Identificar partes interessadas” e “Definir profissionais a serem envolvidos” foram conduzidas pelo o autor. No processo, o pesquisador executou apenas o papel de “Promotor da Difusão da Inovação”, levando conhecimento para os tomadores de decisão e demonstrando a necessidade e os benefícios que os Dados Abertos poderiam trazer para a instituição. Tais ações foram de responsabilidade do pesquisador porque coincidiram com as três etapas iniciais para condução do estudo de caso, elencadas na Figura 6.1: Primeiro contato formal; Apresentação dos objetivos e Definição de pessoas

²¹ <https://sites.google.com/cin.ufpe.br/odsp>

chaves. Após a decisão positiva em favor da execução do processo, o processo foi conduzido pela equipe da instituição.

Na reunião de apresentação da iniciativa vislumbramos e comprovamos a assertiva exposta por OLIVEIRA et al (2016), na qual as pessoas do setor de Tecnologia da Informação estão mais interessadas nos dados brutos e nas questões técnicas, enquanto os tomadores de decisões (Diretor) vislumbravam as análises dos dados graficamente. A exploração das visualizações gráficas de outras iniciativas similares foi essencial para a formação de opinião a favor da iniciativa. O trabalho de convencimento feito pelo “Promotor da Difusão da Inovação” foi facilitado com os *Early Adopters*, perfil indicado pelo processo e presente na equipe escolhida. SCHULER (2002) defende que em um processo de mudança, o sucesso depende de uma base contando com o apoio de aliados naturais chamados pelo o autor de “*Early Adopters*”. Um perfil de *stakeholder* que está propenso desde o início a adotar a inovação, e devido sua credibilidade na instituição, ao adotar o que estiver sendo proposto como mudança, muitos irão segui-lo. A consequência dessa atividade de identificar e convidar “*Early Adopters*” para compor a equipe teve efeito durante a execução do processo, visto que houve duas adesões de pessoas para compor voluntariamente a iniciativa.

Durante o subprocesso “Identificar os dados que serão abertos” foram identificados 17 (Dezessete) conjuntos de dados passíveis de abertura. Cabe ressaltar que as informações foram catalogadas no artefato “Planilha de Identificação da Informação²²”. O levantamento pela equipe consistiu na combinação de duas estratégias de levantamento das informações, uma foi obter dados com os setores e a outra obter lista de sistemas. A estratégia proposta pelo processo, incluía também a obtenção da lista de dados pedidos pelo Serviço de Informação ao Cidadão (SIC). Neste caso, tal estratégia foi excluída pelo Campus, porque o SIC está localizado na Reitoria. Tal fato valida o paralelismo na representação das estratégias no processo ODSP, tornando-as atividades opcionais e independentes uma da outra. O levantamento resultou em um total de 58,82% dos dados no formato XLS e 41,18% disponíveis em banco de dados relacional. Mais da metade em um formato estruturado, um contraditório com as análises das informações disponíveis nos portais. O relatado pela equipe é que algumas dessas informações são publicadas no formato PDF nos portais para garantir a fidedignidade das informações.

²² <https://sites.google.com/cin.ufpe.br/odsp>

Uma das modificações realizadas na “Planilha de Identificação da Informação” em relação a versão colombiana, foi a alteração da coluna classificação das Categorias, a qual representa os aspectos ou temas correlacionados à informação. Utilizamos como categorias, exclusivamente, o primeiro nível do Vocabulário Controlado do Governo Eletrônico (VCGE) elencado na Figura 6.4, definidos no Padrão de Interoperabilidade do Governo Eletrônico (e-Ping) e no Padrão de Metadados do Governo Eletrônico (e-PMG). Dos conjuntos de dados, 11 (Onze) foram classificados na categoria “Educação”, 6 (Seis) em “Economia e Finanças” e 1 (Um) em “Governo e Política”. No entanto, tal classificação pode não estar representada corretamente, visto que a equipe sentiu dificuldades em utilizar apenas o primeiro nível do VCGE na classificação. Por outro lado, elencar tais informações pelo o VCGE facilitará o processo de catalogação no portal dados.gov.br.

Figura 6.4 - Categorias do Vocabulário Controlado do Governo Eletrônico (VCGE).



Fonte: <http://vocab.e.gov.br/2011/03/vcge>

Nas atividades que compõem “Analisar o nível de sigilo dos dados a serem abertos”, a atividade “Identificar instrumentos regulatórios referente às informações da instituição” não foi realizada pela a equipe por 2 (dois) motivos. Primeiramente, por não haver nenhum documento institucional interno referente a classificação das informações, e segundo pelo o artefato da atividade posterior a essa estar de acordo com os critérios estabelecidos pela Lei de Acesso a Informação (LAI). Então, o filtro das informações de acordo com os instrumentos regulatórios foi realizado pela equipe no artefato de “Análise da Informação”²³. Enxergamos

²³ <https://sites.google.com/cin.ufpe.br/odsp>

tal artefato como um facilitador, visto que o conhecimento da equipe em relação à LAI é baixo, conforme questionário aplicado antes da execução do processo. Com o auxílio do artefato “Análise da Informação”, 6 (Seis) dos conjuntos de dados foram identificados como contendo informações pessoais. Dos 6 (seis), 4 (Quatro) foram classificados em parcialmente publicáveis, ou seja, apresentaram informações consideradas sigilosas de acordo com a legislação. No entanto, existe a possibilidade de anonimizar essas informações, tornando o conjunto de dados publicável. Os outros 2 (dois) foram descartados para publicação.

Dos conjuntos de dados restantes para publicação, após a análise do nível de sigilo dos dados, 9 (nove) foram priorizados para publicação em curto prazo, 2 (dois) em médio prazo e 4 (quatro) em longo prazo. A equipe utilizou o artefato de “Priorização dos Dados”²⁴ de acordo com dois parâmetros: Impacto e Dificuldade. Os critérios relacionados ao impacto de cada conjunto de dados, foram: i) Grau de Relevância para o cidadão; ii) Atendimento a instrumentos regulatórios e iii) Alinhamento perante o Planejamento Estratégico Institucional (PEI) e os planejamentos setoriais. Os critérios relacionados a dificuldade foram: i) Esforço requerido para publicar; ii) Gestão do conjunto de informações e sistemas do órgão e iii) Fonte de dados. Os critérios e escalas utilizados nos artefatos de “Priorização dos Dados” estão explicitados no Apêndice B. A Tabela 6.1 demonstra os resultados para “Impacto” e “Dificuldade” para cada conjunto de dados em uma escala de 1 a 4.

Tabela 6.1 - Resultados da análise para “Impacto” e “Dificuldade” dos conjuntos de dados.

Conjunto de Dados	Total de Impacto	Total de Dificuldade	Nível de Prioridade do Dados
Alunos Matriculados	3,30	1,98	Curto Prazo
Contratos de TI	3,30	0,99	Curto Prazo
Atas de Registro de Preço	2,64	1,98	Curto Prazo
Contratos ativos	2,64	1,98	Curto Prazo
Projetos de Extensão	3,30	0,99	Curto Prazo
Projetos de Pesquisa	3,30	0,99	Curto Prazo
Estudantes Egressos	3,30	0,99	Curto Prazo

²⁴ <https://sites.google.com/cin.ufpe.br/odsp>

Evasão	3,30	0,99	Curto Prazo
Consumo de Energia	1,98	0,99	Médio Prazo
Consumo de Água	1,98	0,99	Médio Prazo
Lista de Funcionários Terceirizados	3,30	0,99	Curto Prazo
Lista de Computadores	2,31	2,97	Longo Prazo
Lista de Multimídia	2,31	2,97	Longo Prazo
Lista de Lousa Interativa	2,31	2,97	Longo Prazo
Lista de Laboratórios	2,31	2,97	Longo Prazo

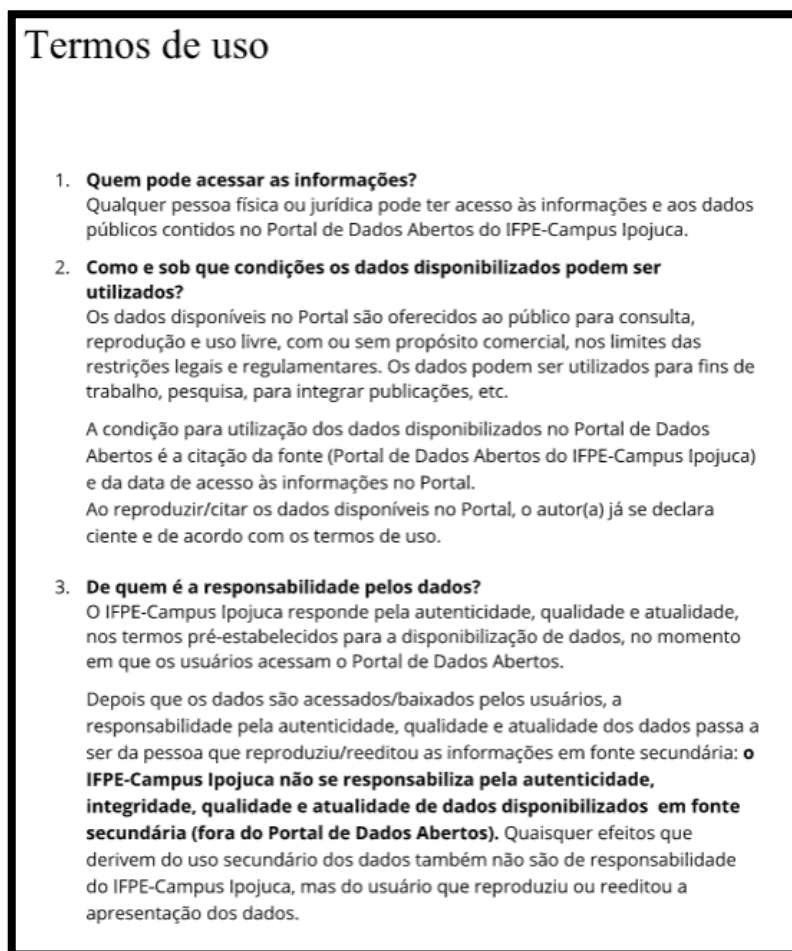
Fonte: Elaborado pelo o autor.

No subprocesso “Apresentar opções de licenças”, as opções de licenças existentes prevista na atividade “Identificar as licenças existentes”, traziam consigo textos em inglês, dificultando o entendimento do público alvo pretendido pela equipe. Sendo assim, a equipe criou a sua própria licença. Cabe ressaltar que a licença desenvolvida para o Campus Ipojuca foi adaptada da Licença de Dados Abertos do Governo de São Paulo²⁵. A equipe aprovou o termo de licença criado e optou por utilizar a mesma licença para todos os conjuntos de dados publicados. Por se tratar de um documento de cunho jurídico, a equipe sugeriu encaminhar a licença para a Procuradoria Jurídica da instituição. A Figura 6.5 demonstra o Termo de Uso²⁶ criado para o Portal de Dados Abertos do IFPE – Campus Ipojuca.

²⁵ <http://www.governoaberto.sp.gov.br/regras-de-utilizacao/>

²⁶ <https://sites.google.com/cin.ufpe.br/odsp>

Figura 6.5 - Termos de uso para o Portal de Dados Abertos do IFPE – Campus Ipojuca.

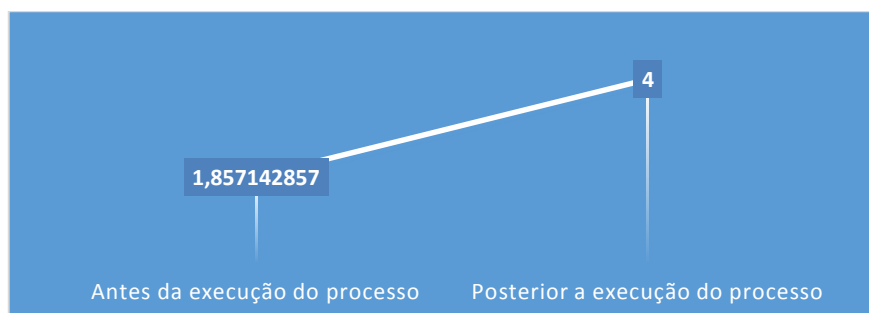


Fonte: IFPE – Campus Ipojuca.

No subprocesso “Elaborar um plano de ações para publicação dos dados”, iniciaram uma proposta de Plano de Dados Abertos de acordo com o Manual para elaboração de PDA²⁷. Mas a proposta do PDA foi abandonada para ganhar celeridade no estudo de caso, mesmo assim essa curta vivência fez a equipe entender a necessidade de um documento orientador na publicação de Dados Abertos. É importante ressaltar que o entendimento da equipe a respeito da necessidade de um PDA na organização estava em um nível de muito baixo para baixo antes da execução do processo. Porém, após execução do processo, passou para um nível mais alto (Figura 6.6), considerando uma escala de 1 a 5. Tal resultado foi mensurado através da questão 3 (três) do questionário descrito no início da sessão.

²⁷http://www.planejamento.gov.br/secretarias/upload/arquivo/governoaberto/copy_of_manual_elaboracao_plano_dados_abertos.pdf

Figura 6.6 - Entendimento da equipe a respeito da necessidade de um PDA na organização.



Fonte: Elaborado pelo autor.

No planejamento para execução das fases de criação e publicação, foi utilizado um único conjunto de dados referente aos contratos ativos do IFPE – Campus Ipojuca (Figura 6.7). Em seguida, a equipe preencheu a planilha de “Descrição de Metadados”²⁸ com os metadados referente ao conjunto de dados escolhido e criou um “Dicionário de Dados”²⁹ para o conjunto de dados com um vocabulário padrão estabelecido pelo o IFPE – Campus Ipojuca.

Figura 6.7 - Conjunto de dados em seu formato inicial (XLSX).

TABELA DE CONTRATOS VIGENTES DO IFPE CAMPUS IPOJUCA 2016								Data da última atualização :		09/11/2016	
SITUAÇÃO	ORIGEM	CONTRATO	PROCESSO	EMPRESA	CNPJ / CPF	OBJETO	VALIDADE	VALOR CONTRATO	FISCAIS	SITUAÇÃO	OBSERVAÇÃO
ATIVO	SRP	001/2016	23296.00042.2016-71	PETROCARD	08.201.107/0001-76	COMBUSTÍVEL	26/01/2017	R\$ 83.385,13	Adelino Paulino	Contrato em Renovação	IMPOSSIBILIDADE DE RENOVACÃO
ATIVO	Pregão	002/2014	23296.001681.2014-64	S. S. EMPREENDIMENTOS E SERVIÇOS LTDA	03.153.145/0001-28	Motorista	07/02/2017	R\$ 117.534,48	Adelino Paulino	Contrato em Renovação	MEMORANDO SUGERINDO RENOVACÃO
ATIVO	SRP	003/2013	23296.003737.2013-34	EMBRATEL (CLARO)	33.530.486/0001-29	Telefonia Fixa	14/02/2017	R\$ 8.610,00	José Mário	Contrato em andamento	Não será renovado.
ATIVO	Inexatibilidade	016/2014	23296.000817.2014-19	COMPESA	09.769.035.0001-64	Água encanada	28/02/2017	R\$ 37.500,00	Kethuly Albuquerque	Contrato em andamento	AGUARDA RESPOSTA DA COMPESA
ATIVO	CONCORRÊNCIA 001/2015	007/2016	23296.006443.2015-26	ANDRE CLEDSON OLIVEIRA MELO - ME	11.717.380/0001-04	CANTINA	14/03/2017	R\$ 6.262,56	RAIANE DOS SANTOS	Contrato em Renovação	AGUARDA RESPOSTA DA EMPRESA
ATIVO	SRP	002/2016	23296.016876.2015-91	TRIPS PASSAGENS E TURISMO LTDA	00.013.698/0001-80	PASSAGENS	18/03/2017	R\$ 50.306,60	CRISLÂNDIA	Contrato em Renovação	1º ANO DE CONTRATO
ATIVO	SRP	005/2013	23296.005300.2012-12	S. S. EMPREENDIMENTOS E SERVIÇOS LTDA	03.153.145/0001-28	Motorista	01/04/2017	R\$ 117.534,48	Marlon Péricles	Contrato em andamento	PRORROGADO
ATIVO	Inexatibilidade	009/2016	23296.001518.2016-63	IMPRESA NACIONAL	04.196.645/0001-00	publicação no D.O.U.	25/04/2017	R\$ 50.000,00	JANE KELLY	Contrato em andamento	PRORROGADO
ATIVO	Inexatibilidade	010/2016	23296.001521.2016-67	EBC	09.168.704/0001-42	Distribuição de Publicidade Legal	16/05/2017	R\$ 13.423,20	JANE KELLY	Contrato em andamento	1º ANO DE CONTRATO
ATIVO	SRP 000/2014 REITORIA	000/2016	23296.005162.2014-75	MAPFRE VIDA S.A.	02.443.932/0408-90	SEGURO DE ALUNO	20/05/2017	R\$ 20.531,96	VILMAR	Contrato em andamento	1º ANO DE CONTRATO
ATIVO	SRP	003/2013	23296.001604.2012-23	VIVO (Telefônica) S.A.	02.443.932/0408-90	Internet móvel	10/05/2017	R\$ 3.330,00	José Mário	Contrato em andamento	
ATIVO	SRP	002/2013	23296.00125.2012-74	TNL PCS S/A	04.164.616/0001-53	Telefonia móvel	03/06/2017	R\$ 37.128,48	José Mário	Contrato em Renovação	
ATIVO	Pregão	006/2012	23296.002334.2012-37	INSTRUCON COM. E SERV. LTDA ME	00.948.860.0001-30	Manutenção	02/07/2017	R\$ 80.350,00	RODRIGO EMANUEL	Contrato em andamento	AUTORIZADA A PRORROGAÇÃO
ATIVO	Dispensa	018/2013	23296.005465.2013-15	CELPE	10.835.332/0001-00	Energia Elétrica	02/08/2017	R\$ 300.000,00	RODRIGO EMANUEL	Contrato em andamento	
ATIVO	Pregão nº 11/2013	015/2014	23296.003163.2014-61	SERVAL	03.822.268.0001-05	Jardineiro	01/09/2017	R\$ 21.415,32	Aurelio Filgueiras	Contrato em andamento	
ATIVO	Pregão nº 11/2013	017/2014	23296.010154.2014-41	SERVAL	03.822.268.0001-05	Pedreiro	04/09/2017	R\$ 27.390,60	Aurelio Filgueiras	Contrato em andamento	
ATIVO	SRP	012/2015	23296.006442.2015-81	GUARDSECURE	42.035.097/0002-07	Vigilância	16/10/2017	R\$ 554.308,32	Carina Souto	Contrato em andamento	
ATIVO	Pregão	013/2013	23296.011381.2013-71	SERVAL PORTUARIA	03.822.268/0001-05	Portaria	21/10/2017	R\$ 12.412,32	Adelino Paulino	Contrato em andamento	
ATIVO	Pregão	013/2016	23296.003471.2016-72	WORLDNET TELECOM COMERCIO E	05.773.360/0001-00	Internet e Dados	01/11/2017	R\$ 26.160,00	Leonardo Maria José	Contrato em andamento	
ATIVO	Pregão	019/2016	23296.005343.2016-63	SURA	33.065.639/0001-27	SEGURO DE VEÍCULOS	01/11/2017	R\$ 6.381,35	Adelino Paulino	Contrato em andamento	1º ANO DE CONTRATO
ATIVO	Pregão	012/2016	23296.01557.2016-19	TICKET	03.506.307/0001-57	Manutenção de frota	01/11/2017	R\$ 33.323,00	Adelino Paulino	Contrato em andamento	1º ANO DE CONTRATO
ATIVO	SRP	014/2016	23296.003471.2016-72	SERVAL	03.822.268.0001-05	Conservação e Limpeza	01/12/2017	R\$ 383.681,16	DANILO ÁVILA	Contrato em andamento	1º ANO DE CONTRATO
ATIVO	SRP	015/2015	23296.01543.2015-88	ART-JET	05.556.967/0001-78	Reparação	17/12/2016	R\$ 21.516,00	Leandro	Contrato em andamento	DEROILDA EMITIDA
ATIVO	COMODATO	012/2015	23296.013435.2015-36	PETROBRAS	33.000.167/1111-08	COMODATO	16/02/2019	R\$	ENIO CAMILO	Contrato em andamento	
ATIVO	Inexatibilidade	014/2015	23296.017127.2015-81	CORREIOS	32.028.316/0021-57	CORRESPONDENCIA	17/12/2020	R\$ 6.000,00	JANE KELLY	Contrato em andamento	
	EM PRORROGAÇÃO										
	PRORROGADOS										

Fonte: IFPE – Campus Ipojuca

A Figura 6.8 mostra o conjunto de dados no formato CSV após os subprocessos de conversão, higienização e anonimização dos dados. As atividades dessas etapas foram realizadas manualmente pela a equipe de TI utilizando a ferramenta Microsoft Excel³⁰.

²⁸ <https://sites.google.com/cin.ufpe.br/odsp>

²⁹ <https://sites.google.com/cin.ufpe.br/odsp>

³⁰ <https://products.office.com/pt-br/excel>

Figura 6.8 - Captura da tela referente ao arquivo CSV do conjunto de dados.

	SITUACAO;ORIGEM;CONTRATO;PROCESSO;EMPRESA;CNPJ;OBJETO;VALIDADE;V
1	ATIVO;PREGAO SRP;001/2016;23296.000142.2016-71;PETROCARD;08.201.
2	ATIVO;PREGAO TRADICIONAL;002/2014;23296.001681.2014-64;S. S. EMP
3	ATIVO;PREGAO SRP;009/2013;23296.003737.2013-34;EMBRATEL (CLARO);
4	ATIVO;INEXIBILIDADE;016/2014;23296.000817.2014-19;COMPESA;09.769
5	ATIVO;CONCORRENCIA ;007/2016;23296.006443.2015-26;ANDRE CLEDSON
6	ATIVO;PREGAO SRP;002/2016;23296.016876.2015-91;TRIPS PASSAGENS E
7	ATIVO;PREGAO SRP;005/2013;23296.005900.2012-12;S.S. EMPREENDIMEN
8	ATIVO;INEXIBILIDADE;009/2016;23296.001518.2016-63;IMPrensa NACIO
9	ATIVO;INEXIBILIDADE;010/2016;23296.001521.2016-87;EBC;09.168.704
10	ATIVO;PREGAO SRP;008/2016;23296.005162.2014-75;MAFFRE VIDA S.A.;
11	ATIVO;PREGAO SRP;003/2013;23296.001604.2012-23;VIVO (TELEFONICA)
12	ATIVO;PREGAO SRP;002/2013;23296.009125.2012-74;TNL PCS S/A;04.16
13	ATIVO;PREGAO TRADICIONAL;006/2012;23296.002334.2012-97;INSTRUCON
14	ATIVO;DISPENSA;018/2013;23296.005465.2013-15;CELPE;10.835.932/00
15	ATIVO;PREGAO TRADICIONAL;015/2014;23296.009183.2014-61;SERVAL;03
16	ATIVO;PREGAO TRADICIONAL;017/2014;23296.010154.2014-41;SERVAL;03
17	ATIVO;PREGAO SRP;012/2015;23296.006442.2015-81;GUARDSECURE;42.03
18	ATIVO;PREGAO TRADICIONAL;019/2013;23296.011981.2013-71;SERVAL PO
19	ATIVO;PREGAO TRADICIONAL;013/2016;23296.003471.2016-72;WORLDNET
20	ATIVO;PREGAO TRADICIONAL;011/2016;23296.005343.2016-63;SURA;33.0
21	ATIVO;PREGAO TRADICIONAL;012/2016;23296.001557.2016-19;TICKET;03
22	ATIVO;PREGAO SRP;014/2016;23296.003471.2016-72;SERVAL;03.822.268
23	ATIVO;PREGAO SRP;015/2015;23296.017543.2015-88;ART-JET;05.556.96
24	ATIVO;COMODATO;013/2015;23296.013435.2015-36;PETROBRAS;33.000.16
25	ATIVO;INEXIBILIDADE;014/2015;23296.017127.2015-81;CORREIOS;32.02
26	

Fonte: IFPE – Campus Ipojuca

O conjunto de dados, obtido ao final da fase de criação, foi armazenado em um repositório local e catalogado no Portal Brasileiro de Dados Abertos (Figura 6.9 e Figura 6.10). Cabe ressaltar que o departamento de TI enfrentou problemas no acesso do portal dados.gov.br. No envio do e-mail indicado no processo, o mesmo retornava com erro. O problema foi resolvido com abertura de chamado na Central de Serviço e Suporte do SISP (Sistema de Administração dos Recursos de Tecnologia da Informação)³¹. O SISP é um sistema do Ministério do Planejamento e possui uma equipe especializada para auxílio no processo de abertura dos dados. Tal atividade de suporte foi incorporada no processo para casos similares que venham acontecer futuramente em outras instituições.

³¹ <https://c3s.sisp.gov.br/cau/>

Figura 6.9 - Conjunto de dados do IFPE - Campus Ipojuca catalogado no portal dados.gov.br.



Fonte: dados.gov.br

Figura 6.10 - Conjunto de dados do IFPE - Campus Ipojuca visualizado no portal dados.gov.br.

URL: http://200.133.17.107/dadosabertosipojuca/contratosativos_ifpeipojuca_2016.csv
Arquivo CSV referente ao Conjunto de Dados "Lista de Contratos Ativos do IFPE - Campus Ipojuca".

Explorador de Dados

Embutir

Add Filter

Grid Graph Map 25 records 1 - 25 Search data Go Filters

_id	SITUAC...	ORIGEM	CONTR...	PROCE...	EMPRES...	CNPJ	OBJETO	VALIDA...	VALOR...	FISCAL...
1	ATIVO	PREGA...	001/2016	23296.0...	PETRO...	08.201.1...	COMBU...	2017-01	R\$ 83.9...	ADEILD...
2	ATIVO	PREGA...	002/2014	23296.0...	S. S. EM...	03.159.1...	MOTORI...	2017-07	R\$ 117.5...	ADEILD...
3	ATIVO	PREGA...	009/2013	23296.0...	EMBRAT...	33.530.4...	TELEFO...	2017-02	R\$ 8.61...	JOSE M...
4	ATIVO	INEXIBI...	016/2014	23296.0...	COMPESA	09.769.0...	AGUA E...	2017-02	R\$ 37.5...	KETHUL...
5	ATIVO	CONCO...	007/2016	23296.0...	ANDRE...	11.717.3...	CANTINA	2017-03	R\$ 6.26...	RAIANE...
6	ATIVO	PRFCA	007/2016	23296.0...	TRIPS P	08.015.8...	PASSAG	2017-03	R\$ 40.3...	CRISI A

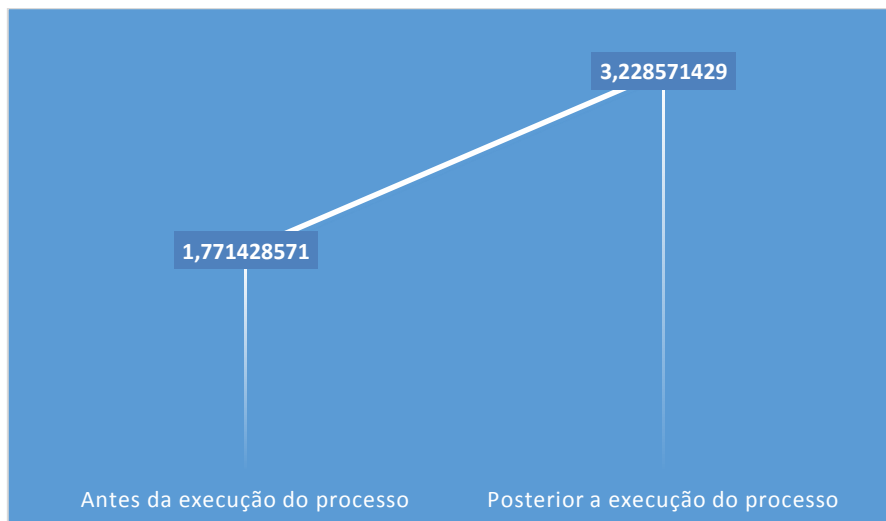
Fonte: dados.gov.br

6.3 Considerações Finais

Finalmente, com relação aos objetivos propostos pelo processo ODSP, o desempenho do processo foi satisfatório, visto que as observações pontuadas não impediram a execução do processo, apenas retardando o mesmo em alguns momentos. O processo foi executado dentro do esperado, tendo em curto prazo (15 dias), um resultado prático de abertura de dados à disposição da sociedade. Além de elevar a maturidade da equipe e, conseqüentemente, a maturidade da instituição. A Figura 6.11 demonstra um aumento de 29% em uma escala de 0 a 5 para o conhecimento da equipe em relação à temática de Dados Abertos. Este índice de

conhecimento foi mensurado considerando a média das respostas de todos respondentes para todas questões do questionário descrito no início da sessão anterior.

Figura 6.11 - Índice de conhecimento da equipe em relação a temática de Dados Abertos.



Fonte: Elaborado pelo o autor:

É importante ressaltar que as observações da equipe foram essenciais para a melhoria do processo e que o feedback sempre será necessário, sendo capturado continuamente. Como visto no capítulo anterior, estamos falando de um processo que estará sempre em revisão e atualização, tornando o feedback um insumo fundamental para o seu aperfeiçoamento.

7 Considerações Finais

7.1 Contribuições

O trabalho desenvolvido nesta dissertação apresentou como principal contribuição uma proposta de processo para Publicação de Dados Abertos na Rede Federal de Educação Profissional, Científica e Tecnológica, baseado nos conceitos de BPM. Do desenvolvimento até os resultados, outras contribuições foram geradas, incluindo: O levantamento do estado da arte, demonstrando que ainda existem limitações nas abordagens existentes para publicação de Dados Abertos; A disponibilização de um processo focado em instituições educacionais, onde não foram identificados processos voltados ao contexto de publicação dos Dados Educacionais; Do ponto de vista computacional, um processo que permite a produção dos dados até a sua disponibilização com qualidade, permitindo uma melhor análise e compreensão dos dados, gerando novos conhecimentos e novos negócios. O impacto é muito mais extenso do que o produto gerado. O fomento para tornar os dados públicos, não só das esferas governamentais, mas também das iniciativas privadas, traz possibilidades infinitas e oportunidades ainda desconhecidas. Nesse contexto, este trabalho pode ser visto tanto como um instrumento de fomento, mas também como uma forma de praticar a responsabilidade social. A vida de um cidadão pode mudar pelo valor agregado a partir dos resultados das decisões guiadas por dados sobre domínios diversos (Educação, Transporte, Meio ambiente, Saúde, dentre outros). Por um lado, o governo em qualquer esfera (Municipal, Estadual e Federal) pode fazer um melhor uso do dinheiro público, oferecendo melhor qualidade de vida para os cidadãos, e, por outro lado, os dados permitirão que profissionais prevejam catástrofes, rastreiem epidemias, achem a cura de doenças. A partir disso, temos que o impacto de um trabalho dessa natureza é realmente imensurável.

Desviando a atenção dessa visão global e focando nos objetivos do Mestrado Profissional, podemos citar a contribuição para o desenvolvimento profissional do autor, passando pelo processo de pesquisa e finalizando com uma maior gama de conhecimentos relacionados à Ciência da Computação, com ênfase nas temáticas de Dados Abertos e Gerenciamento de Processos de Negócios. Além do mais, a soma desse conhecimento em favor da empresa, torna-se um diferencial no desenvolvimento das próprias empresas e, conseqüentemente, para o desenvolvimento de cada região na qual a mesma esteja inserida. O

Mestrado Profissional tem o objetivo da interação da academia com a instituição através do aluno. A interação com o Instituto Federal de Pernambuco (IFPE) gerou discussões em relação ao objeto de estudo da pesquisa (Dados Abertos), culminando em um resultado prático de abertura de dados e na iminência de uma proposta para criação de uma Comissão de Dados Abertos para trabalhar a temática em 2017 na instituição em nível estadual. Além do mais, a contribuição se estende para toda Rede Federal como um instrumento que auxilia no atendimento aos instrumentos regulatórios frente à temática, visto que fomentar a disponibilização e o uso de dados abertos é um dos objetivos estratégicos elencados na Estratégia de Governança Digital, instituída pelo Decreto nº 8.638, de 15 de janeiro de 2016, juntamente com o decreto lei nº 8.777, que institui a Política de Dados Abertos do Poder Executivo Federal.

O processo para Publicação de Dados Abertos na Rede Federal de Educação Profissional, Científica e Tecnológica, denominado ODSP (OPEN DATA STARTER PROCESS), por ter sido baseado nos conceitos de BPM, propõe alguns passos essenciais para instituições imaturas, não limitando a criatividade profissional dos envolvidos. Além disso, configura-se como um instrumento que determina um planejamento metódico para um processo de abertura de dados. Dessa forma, ele harmoniza as áreas envolvidas, facilitando e acelerando a implantação do processo, além da possibilidade de ser um rito único e de referência para as instituições. Um dos pontos fortes do processo ODSP, é ter uma versão *Web* interativa³², facilitando o acesso e dando celeridade para o entendimento de todo o processo. Por fim, avaliamos os objetivos propostos pelo processo em um ambiente real, além de contribuir com um conjunto de dados abertos para a sociedade.

7.2 Limitações

A limitação da pesquisa foi a estratégia de avaliação ter utilizado o estudo de caso do tipo único, ou seja, testada apenas em uma instituição. Os estudos de casos múltiplos demandam maiores recursos e tempo por parte do pesquisador. Mas fica o convite para que o processo seja aplicado em outras instituições, proporcionando o *feedback* necessário para o seu aperfeiçoamento.

³² <https://sites.google.com/cin.ufpe.br/odsp>

7.3 Trabalhos futuros

Como trabalhos futuros destacamos:

1. Aperfeiçoar o processo ODSP, mediante a execução de outros estudos de casos reais. Podendo ser expandido para outros contextos distintos para qual foi proposto (Rede Federal).
2. Avaliar aceitação e uso do processo ODSP utilizando o Modelo de Aceitação de Tecnologia (*Technology Acceptance Model* – TAM³³).
3. O processo ODSP tem foco na abertura de dados. Com isso surge a necessidade de mais um nível para auxiliar as instituições no processo de conexão dos dados.
4. Desenvolvimento de ferramentas e softwares para automação do processo proposto.
5. Submeter trabalhos dessa pesquisa em eventos da área, bem como publicar em periódicos.

³³ https://en.wikipedia.org/wiki/Technology_acceptance_model

REFERÊNCIAS

- ALCANTARA, Williams et al. Desafios no uso de Dados Abertos Conectados na Educação Brasileira. In: Anais do DesafiE-4º Workshop de Desafios da Computação Aplicada à Educação. CSBC. 2015.
- ANTONIOU, Grigoris; HARMELEN, Frank Van. A semantic web primer 2nd ed. MIT press, 2008.
- ARAÚJO, Lucas de Ramos; DE SOUZA, Jairo Francisco. AUMENTANDO A TRANSPARÊNCIA DO GOVERNO POR MEIO DA TRANSFORMAÇÃO DE DADOS GOVERNAMENTAIS ABERTOS EM DADOS LIGADOS. Revista Eletrônica de Sistemas de Informação, v. 10, n. 1, 2011.
- AUER, S. et al. Managing the Life-Cycle of Linked Data with the LOD2 Stack. The Semantic Web - Lecture Notes in Computer Science, v. 7650, n. 257943, p. 16, 2012.
- ÁVILA, T. J. T. Uma proposta de modelo de processo para publicação de Dados Abertos Conectados Governamentais. 2015. Mestrado em Modelagem Computacional do Conhecimento – Universidade Federal de Alagoas.
- BANCO MUNDIAL, Open Government Data Toolkit, 2016.
- BALDAM, Roquemar de Lima et al. Gerenciamento de processos de negócios: BPM– Business Process Management. São Paulo: Érica, 2007.
- BAUER, Florian; KALTENBÖCK, Martin. Linked open data: The essentials. Edition mono/monochrom, Vienna, 2011.
- BIZER, Christian; HEATH, Tom; BERNERS-LEE, Tim. Linked data-the story so far. Semantic Services, Interoperability and Web Applications: Emerging Concepts, p. 205-227, 2009.
- BRASIL. Cartilha Técnica para Publicação de Dados Abertos no Brasil. [S.l.], 2011. 10 p. Disponível em: <<http://dados.gov.br/cartilha-publicacao-dados-abertos/>>.
- BRASIL. Constituição da República Federativa do Brasil de 1988. 1988. Disponível em: <https://www.planalto.gov.br/>. Acesso em: 02 jan. 2017.
- BRASIL. Decreto No 8.638, de 15 de Janeiro de 2016. 2016. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/D8638.htm. Acesso em: 02 jan. 2017.
- BRASIL. Decreto No 8.777, de 11 de Maio de 2016. 2016. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2016/Decreto/D8777.htm. Acesso em: 02 jan. 2017.
- BRASIL. Decreto No 11.892, de 29 de Dezembro de 2008. 2008. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/lei/11892.htm. Acesso em: 05 jan. 2017.

BRASIL. Lei complementar nº101, de 4 de maio de 2000. Lei de Responsabilidade Fiscal. 2000. Disponível em: <https://www.planalto.gov.br/>. Acesso em: 02 jan. 2017.

BRASIL. Lei complementar nº131, de 27 de maio de 2009. Lei da Transparência. 2009. Disponível em: <https://www.planalto.gov.br/>. Acesso em: 02 jan. 2017.

BRASIL. Lei No 12.527, de 18 de Novembro de 2011. Lei de Acesso a Informação. 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2011-2014/2011/lei/l12527.htm.

BRASIL. Kit de dados abertos. [S.l.], 2014. 3–5 p. Disponível em: <<http://kit.dados.gov.br>>.

BRASIL. Manual para Elaboração de Plano de Dados Abertos. [S.l.], 2014. v. 7, 38 p. Disponível em: <http://www.planejamento.gov.br/secretarias/upload/Arquivos/governo_aberto/manual_elaboracao_plano_dados_abertos.pdf>.

BRASIL. Plano de Dados Abertos - Ministério do Planejamento, Orçamento e Gestão do Brasil. [S.l.], 2014. 1–15 p. Disponível em: <http://www.planejamento.gov.br/secretarias/upload/Arquivos/governo_aberto/plano_dados_abertos.pdf>.

BREITMAN, Karin et al. Open government data in Brazil. *IEEE Intelligent Systems*, v. 27, n. 3, p. 45-49, 2012.

BRITO, Kellyton dos Santos et al. Is Brazilian Open Government Data Actually Open Data?: An Analysis of the Current Scenario. *International Journal of E-Planning Research (IJEPR)*, v. 4, n. 2, p. 57-73, 2015.

CARNEIRO, Margareth Fabíola dos Santos. Fundamentos do Gerenciamento de Programas e Projetos. PMA professional management, 2010.

CBOK, BPM. Guia para o Gerenciamento de Processos de Negócio Corpo Comum de Conhecimento. Association of Business Process Management Professionals. ABPMP BPM CBOK, v. 3. 2013.

CGI. TIC Governo Eletrônico 2013 - Pesquisa sobre o uso das tecnologias da informação e comunicação no setor público brasileiro. São Paulo, 2013.

CHILE. Norma Técnica para Publicación de Datos Abiertos en Chile. [S.l.], 2013. 1–28 p. Disponível em: <<http://instituciones.gobiernoabierto.cl/NormaTecnicaPublicacionDatosChilev2-1.pdf>>.

COLOMBIA. Guía para la apertura de datos en Colombia. [S.l.], 2012. 67 p. Disponível em: <<https://goo.gl/gtLuMq>>.

COMSODE. Documents of Practice for Methodology for publishing datasets as open data - COMSODE. [S.l.], 2014. 1–58 p. Disponível em: <http://www.comsode.eu/wp-content/uploads/Annex1_D5.1-Documentation_of_practices.pdf>.

COMSODE. Methodology for publishing datasets as open data - COMSODE. [S.l.], 2014. 1–31 p. Disponível em: <<http://www.comsode.eu/index.php/deliverables/>>.

CONSOLI, S. et al. Geolinked Open Data for the Municipality of Catania. *Proceedings of the 4th International Conference on Web Intelligence, Mining and Semantics (WIMS14)*, p. 58, 2014.

- DE FARIA CORDEIRO, Kelli et al. Use of linked data in the design of information infrastructure for collaborative emergency management system. In: Computer Supported Cooperative Work in Design (CSCWD), 2011 15th International Conference on. IEEE, 2011. p. 764-771.
- DING, L. et al. TWC LOGD: A Portal for Linked Open Government Data Ecosystems. Journal of Web Semantics, Elsevier B.V., v. 9, n. 3, p. 325–333, 2011. ISSN 15708268.
- EAVES, David. The three laws of open government data. Eaves.ca, v. 30, 2009.
- ECUADOR. Guia de Política Pública de Datos Abiertos. [S.l.], 2014. 21 p. Disponível em: <<http://www.gobiernoelectronico.gob.ec/wp-content/uploads/2014/12/GPP-DA-v01-20141128-SNAP-SGE.pdf>>.
- ELMASRI, Ramez; NAVATHE, Shamkant B.; DE OLIVEIRA MORAIS, Rinaldo. Sistemas de banco de dados. 2005.
- FREITAS, Wesley RS; JABBOUR, Charbel JC. Utilizando estudo de caso (s) como estratégia de Pesquisa qualitativa: boas práticas e sugestões. Estudo & Debate, v. 18, n. 2, 2011.
- GALIOTOU, E.; FRAGKOU, P. Applying Linked Data Technologies to Greek Open Government Data: A Case Study. Procedia - Social and Behavioral Sciences, v. 73, p. 479–486, 2013. ISSN 18770428. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S187704281300373X>>.
- GOUVEIA, Juliana Maria Arruda de Andrade. Um estudo sobre melhoria de processos de suporte de TI no Centro de Informática da UFPE. 2011.
- HOXHA, Julia; BRAHAJ, Armand. Open government data on the web: A semantic approach. In: Emerging Intelligent Data and Web Technologies (EIDWT), 2011 International Conference on. IEEE, 2011. p. 107-113.
- HYLAND, B.; WOOD, D. The Joy of Data - A Cookbook for Publishing Linked Government Data on the Web. In: Linking Government Data. [S.l.: s.n.], 2011. p. 3–25.
- ISOTANI, Seiji; BITTENCOURT, Ig Ibert. Dados Abertos Conectados. Novatec Editora, 2015.
- JANOWICZ, Krzysztof et al. Five stars of linked data vocabulary use. Semantic Web, v. 5, n. 3, p. 173-176, 2014.
- KALAMPOKIS, Evangelos; TAMBOURIS, Efthimios; TARABANIS, Konstantinos. A classification scheme for open government data: towards linking decentralised data. International Journal of Web Engineering and Technology, v. 6, n. 3, p. 266-285, 2011.
- LÓSCIO, B. F.; OLIVEIRA, M. I. S. ; BITTENCOURT, I. I. . Publicação e Consumo de Dados na Web: Conceitos e Desafios. In: Carmem Hara; Fabio Porto, Eduardo Ogasawara. (Org.). TÓPICOS EM GERENCIAMENTO DE DADOS E INFORMAÇÕES 2015. 1ed.Porto Alegre: Sociedade Brasileira de Computação, 2015, v. , p. 39-69.
- LUNA, Alexandre. ManGve: Implantando governança ágil: uma visão crítica, uma abordagem prática / Alexandre Luna. – Rio de Janeiro: Brasport, 2011.

MACHADO, Alexandre Lopes; DE OLIVEIRA, Jose Maria Parente. DIGO: An open data architecture for e-government. In: 2011 IEEE 15th International Enterprise Distributed Object Computing Conference Workshops. IEEE, 2011. p. 448-456.

MARIANO, Isadora Cidade. Melhoria de Processos pelo BPM: aplicação no setor público. 2012.

MEDEIROS, Simone Assis; MAGALHÃES, Roberto; PEREIRA, José Roberto. Lei de Acesso à Informação: em busca da transparência e do combate à corrupção. Informação & Informação, v. 19, n. 1, p. 55-75, 2013.

MENDONÇA, R. R. d. et al. LOP - Capturing and Linking Open Provenance on LOD Cycle. In: Proceedings of the Fifth Workshop on Semantic Web Information Management - SWIM '13. ACM Press, 2013. p. 1-8. Disponível em: <<http://dl.acm.org/citation.cfm?id=2484712.2484715>>.

MODESTO, Lisandro Rogério; FERNEDA, Edberto; SANT'ANA, Ricardo Cezar Gonçalves. Representação e persistência de recursos informacionais digitais gerados dinamicamente em sítios oficiais do Governo Federal. Brazilian Journal of Information Science, v. 8, n. 1-2, 2014.

MPOG. GesPública – Guia de Gestão de Processos. 2009. Disponível em: http://gestao.planejamento.gov.br/gespublica/sites/default/files/documentos/guia_de_gestao_d_e_processos_de_governo_0.pdf. Acesso em: 02 jan. 2017.

MPOG. SLTI – Guia de Boas Práticas em Contratação de Soluções de Tecnologia da Informação Versão 2.0. 2014. Disponível em: <https://goo.gl/LvTMOl>. Acesso em: 02 jan. 2017.

NAÇÕES UNIDAS. Guidelines on Open Government Data for Citizen Engagement. 2013. Disponível em: <https://goo.gl/BwwXts>

OKF. Guia de Dados Abertos. 2015. Disponível em: <<http://opendatahandbook.org/guide/ptBR>>.

OLIVEIRA, M; OLIVEIRA, L; LIMA, G; LÓSCIO B. Enabling a Unified View of Open Data Catalogs. In Proceedings of the 18th International Conference on Enterprise Information Systems. 2016. ISBN 978-989-758-187-8, pages 230-239. DOI: 10.5220/0005835202300239

PINHEIRO, R. Flávio. Fundamentos no Gerenciamento de Serviço de TI com base na ITIL V3. São Paulo: TI exames, 2010.

PEDROSO, Louise; TANAKA, Asterio; CAPPELLI, Claudia. A Lei de Acesso à Informação brasileira e os desafios tecnológicos dos dados abertos governamentais. IX Simpósio Brasileiro de Sistemas de Informação, 2013.

ROGERS, E. M. Diffusion of innovations 5th Edition. New York: The Free Press, 2003.

SÃO PAULO; REINO UNIDO. Guia de Dados Abertos. 2016. Disponível em: <<http://www.governoaberto.sp.gov.br/wp-content/uploads/2016/02/Book-Web-Guia-de-Dados-Abertos.pdf>>

SÃO PAULO; REINO UNIDO. Guia de Web Semântica. 2016. Disponível em: <<http://www.governoaberto.sp.gov.br/wp-content/uploads/2016/05/Book-Web-Semantica.pdf>>

SÃO PAULO; REINO UNIDO. Guia do Modelo de Maturidade de Dados Abertos. 2016. Disponível em: <<http://www6.governoaberto.sp.gov.br/wp-content/uploads/2016/02/Book-Web-Modelo-de-Maturidade-de-Dados-Abertos.pdf>>

SÃO PAULO; REINO UNIDO. Guia de Incentivo ao Re uso de Dados Abertos. 2016. Disponível em: <<http://www.governoaberto.sp.gov.br/wp-content/uploads/2016/09/Guia-de-Incentivo-ao-Re-uso-de-Dados-Abertos.pdf>>

SAYOGO, Djoko Sigit; PARDO, Theresa A.; COOK, Meghan. A framework for benchmarking open government data efforts. In: 2014 47th Hawaii International Conference on System Sciences. IEEE, 2014. p. 1896-1905.

SHADBOLT, Nigel; O'HARA, Kieron. Linked data in government. IEEE Internet Computing, v. 17, n. 4, p. 72-77, 2013.

SCHULER, A. J. Planning Change Management Communication. WHA Quality Center, v. R 2-12, p. 1-3, 2002.

SIX MINDS. Ferramentas BPMS. 2015. Seminário de Gestão de Processo do CIN –UFPE. Disponível: <<https://goo.gl/XMSzdI>>.

SUNLIGHT FOUNDATION. Guidelines for Open Data Policies. 2014. Disponível em: <http://sunlightf.wpengine.com/wp-content/uploads/2016/09/OpenDataGuidelines_v3.pdf>

TRIBUNAL DE CONTAS DA UNIÃO. Cartilha 5 Motivos para a abertura de dados na administração pública federal. 2015. Disponível em: <<http://portal.tcu.gov.br/lumis/portal/file/fileDownload.jsp?fileId=8A8182A24F0A728E014F0B366F2E2A40>>

UNDERDAHL, Brian. Business process management for dummies. IBM Limited Edition, 2011.

URUGUAY. Guía rápida de publicación em datos.gub.uy. Montevideo, 2012. 17 p. Disponível em: <http://www.agesic.gub.uy/innovaportal/file/2478/1/guia_publicacion_datos_abiertos.pdf>.

VILLAZON-TERRAZAS, B. et al. Methodological guidelines for publishing government ´ linked data. Linking Government Data, p. 27-49, 2011.

W3C. Best Practices for Publishing Linked Data. 2014. Acessado em 12/10/2016. Disponível em: <<http://www.w3.org/TR/ld-bp/>>.

W3C BRASIL. Manual dos Dados Abertos: desenvolvedores. [S.I.:s.n.], 2011.

W3C BRASIL. Manual dos Dados Abertos: governo. [S.I.:s.n.], 2011.

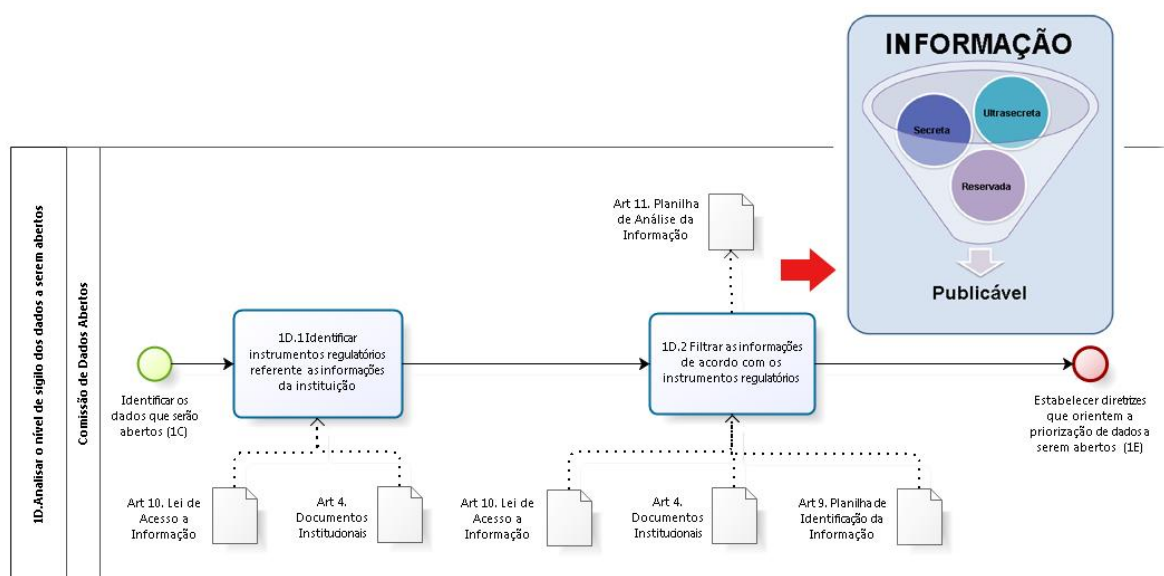
WONDERLICH, John. Ten principles for opening up government information. Washington, DC: Sunlight Foundation. August, v. 11, p. 2010, 2010.

XAVIER, Maria Betânia Gonçalves et al. Mensuração da maturidade da governança de TI na Administração Direta Federal Brasileira. 2010.

APÊNDICE A – Critérios utilizados para análise do nível de sigilo dos dados a serem abertos.

O subprocesso “Analisar o nível de sigilo dos dados a serem abertos” utiliza o auxílio do artefato da “Planilha de Análise da Informação”, artefato inspirado na ferramenta do Governo Colombiano. A classificação da informação em relação ao sigilo utiliza o método de separação (Figura A.1), no qual é um filtro que utiliza diferentes critérios baseados na Lei de Acesso a Informação do País.

Figura A.1 - Subprocesso “Analisar o nível de sigilo dos dados a serem abertos”



Fonte: Elaborado pelo autor.

O artefato foi adaptado para nossa realidade com as alterações dos critérios de acordo com o Artigo 23 da Lei nº 12.527, de 18 de novembro de 2011.

As informações poderão ser classificadas em:

Dados e informações publicáveis: São dados e informações que deverá está disponível para qualquer pessoa.

Dados e informações não publicáveis: São dados e informações cujo acesso, a divulgação e o tratamento de informação ficarão restritos a pessoas que tenham necessidade de conhecê-la de acordo com a legislação.

Dados e informações parcialmente publicáveis: São dados e informações que apresentam em suas partes informações consideradas sigilosas de acordo com a legislação. No entanto existe a possibilidade de anonimizar essas informações, tornando o conjunto de dados publicável.

APÊNDICE B – Critérios utilizados para a priorização de dados a serem abertos

Baseado na metodologia da ferramenta do governo colombiano, utilizamos “Impacto” e “Dificuldade” como parâmetro de priorização dos dados. O impacto é a medida dos benefícios que cada conjunto de dados trará com sua publicação. A dificuldade é a medida de esforço no processo de abrir os dados. A concepção de “Dificuldade” foi alterada em nosso modelo, visto que na ferramenta colombiana, ela é vista como a medida de facilidade na abertura dos dados.

Ressaltamos que cada parâmetro é composto de critérios. A tabela de critérios B.1 abaixo foi a base utilizada para avaliação das informações que foram inventariadas. Cada instituição poderá definir seus critérios, no entanto a base de critérios utilizada no artefato teve como referência o Kit de Dados Abertos (BRASIL, 2014).

Tabela B.1 - Tabela de critérios utilizados nos parâmetros “Impacto” e “Dificuldade” para priorização dos dados.

Critério	Escala	Peso	Descrição	Qualificação
IMPACTO				
Grau de Relevância para o cidadão.	1-4	0,33	O grau de relevância para o cidadão, observando-se especialmente as demandas encaminhadas via solicitações de acesso à informação, dentre as quais, as realizadas a partir de junho de 2015 e que não contenham informações restritas (pessoais ou sigilosas), podem ser consultadas pela busca de pedidos e respostas do sistema e-SIC. Os temas mais procurados também podem ser levantados a partir de estatísticas da ouvidoria do órgão, bem como dos setores e dos serviços mais procurados em seus sítios eletrônicos. Adicionalmente, a organização pode também realizar um encontro com os principais consumidores de seus dados (ex: organizações da sociedade civil, ONGs, pesquisadores, outros órgãos públicos) para entender aonde concentram-se suas necessidades.	(1) Não foi identificada demanda por informações dessa categoria. (2) Demandada por outros órgãos. (3) Demandada por desenvolvedores. (4) Grande demanda por cidadãos, através de consultas na instituição.

Atendimento a instrumentos regulatórios.	1-4	0,33	A informação vem atender os normativos legais e os compromissos formalmente assumidos pelo órgão, inclusive perante organismos internacionais (ex: OGP).	(1) Não tem relação com nenhum instrumento regulatório ou compromisso do órgão. (4) Atende os normativos legais e os compromissos do órgão.
Alinhamento perante o Planejamento Estratégico Institucional (PEI) e os planejamentos setoriais.	1-4	0,33	A informação está alinhada perante o Planejamento Estratégico Institucional (PEI) e os planejamentos setoriais, bem como os relacionados às áreas de tecnologia da informação.	(1) Não existe alinhamento (2) Alinhada parcialmente com alguns planejamentos setoriais (3) Alinhada perante ao Planejamento Estratégico Institucional (PEI) (4) Alinhada totalmente com os planejamentos institucionais.
DIFICULDADE				
Esforço requerido para publicar	1 – 4	0,33	Nível de maturidade e organização das informações e dados existentes. Dificuldade da instituição na publicação da referida informação.	(1) Não requer esforço de desenvolvimento, possui as ferramentas necessárias para extração da

				informação da base de dados. (4) Requer um alto esforço, incluindo processos de Desenvolvimento, ETL e Outros.
Gestão do conjunto de informações e sistemas do órgão.	1 – 4	0,33	O conjunto de informações e sistemas sob a gestão do órgão, em especial, se houver, sistemas estruturadores que são de uso obrigatório transversal para os órgãos da APF;	(1) Dados sob gestão do próprio órgão. (2) Dados sob gestão de outros órgãos da APF. (3) Dados sob gestão de terceiros. (4) Gestão dos dados desconhecida.
Fonte de dados	1 - 4	0,33	Dificuldade dos sistemas de informações para a extração da informação, devido a sua funcionalidade de operação.	(1) Se encontra um servidor com uma base de dados de relatórios e séries históricas. (4) Se encontra em servidor em produção com operações diárias.

Cada informação recebe uma pontuação para impacto e para dificuldade. A pontuação é calculada utilizando a formula matemática da Figura B.1:

Figura B.1 – Formula Matemática de Priorização.

$$\begin{array}{c}
 \text{Valor } CT_1 \\
 \vdots \\
 \text{Valor } CT_m
 \end{array}
 =
 \begin{array}{c}
 \text{Peso } C_1 \times \text{Pontuação } C_1CT_1 \\
 + \dots + \\
 \text{Peso } C_n \times \text{Pontuação } C_nCT_1 \\
 \vdots \\
 \text{Peso } C_1 \times \text{Pontuação } C_1CT_m \\
 + \dots + \\
 \text{Peso } C_n \times \text{Pontuação } C_nCT_m
 \end{array}$$

Fonte: Adaptada de COLOMBIA, 2012.

- CT_x é o critério total “x”
- C_y é o critério de priorização “y”
- C_yCT_x é o critério de priorização “y” aplicado ao critério total “x”
- Os critérios vão desde $x=1\dots n$ e $y=1\dots m$

As pontuações dos parâmetros irão compor uma matriz (Figura B.2) na qual o valor de impacto representará o eixo “Y” e a dificuldade o eixo “X”. De acordo com alocação dos valores nos quadrantes, a informação será priorizada em curto, médio ou longo prazo.

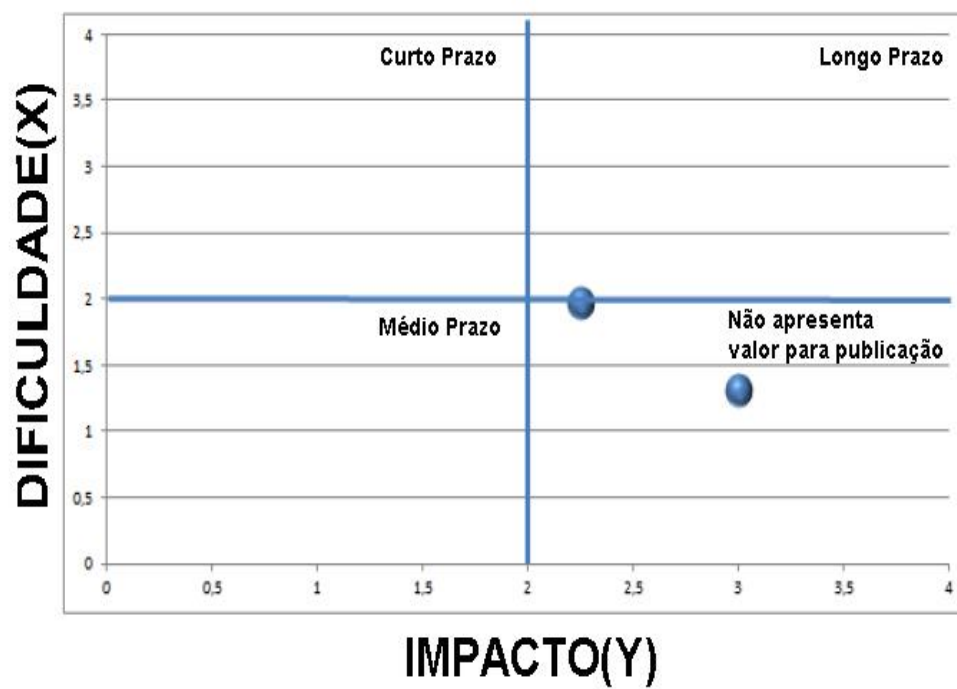
Curto Prazo: São os dados cuja publicação tem um grande impacto e uma dificuldade de implantação relativamente baixa.

Médio Prazo: Prazo aplicado devido à baixa complexidade na implantação e o baixo impacto dos dados.

Longo Prazo: Dados que apresentam um alto impacto e uma alta complexidade para sua publicação.

Não apresenta valor para publicação: Representa os dados com um impacto baixo e uma elevada dificuldade.

Figura B.2 - Matriz de Priorização.



Fonte: Adaptada de COLOMBIA, 2012.

ANEXO A – Recomendações para publicação de Dados Abertos e Dados Abertos Conectados.

Todos os elementos do Anexo A é oriundo do trabalho “ÁVILA, T. J. T. Uma proposta de modelo de processo para publicação de Dados Abertos Conectados Governamentais. 2015. Mestrado em Modelagem Computacional do Conhecimento – Universidade Federal de Alagoas.”

Considerando os documentos apresentados na Tabela A.1 e as melhores práticas elencadas no documento “Best Practices for Publishing Linked Data (BPLD)”³⁴, ÁVILA (2015) extraiu recomendações para publicação de Dados Abertos e Dados Abertos Conectados conforme os gráficos constantes neste anexo.

Tabela A.1 – Processos de publicação de dados abertos analisados

Código	Origem	Documentos (processos) Avaliados	Tipo
P1	Brasil	Manual para Elaboração de Plano de Dados Abertos; Plano de Dados Abertos - Ministério do Planejamento, Orçamento e Gestão do Brasil; Kit de Dados Abertos; Cartilha Técnica para Publicação de Dados Abertos no Brasil v1.0 (BRASIL, 2014)	Dados Abertos Governamentais
P2	Chile	Norma Técnica para Publicación de Datos Abiertos en Chile (CHILE, 2013).	Dados Abertos Governamentais
P3	Colômbia	Guía para la Apertura de Datos em Colombia (COLOMBIA, 2012)	Dados Abertos Governamentais
P4	Equador	Guía de Política Pública de Datos Abiertos – Ecuador (ECUADOR, 2014).	Dados Abertos Governamentais
P5	Uruguai	Guía rápida de publicación em datos.gub.uy	Dados Abertos Governamentais

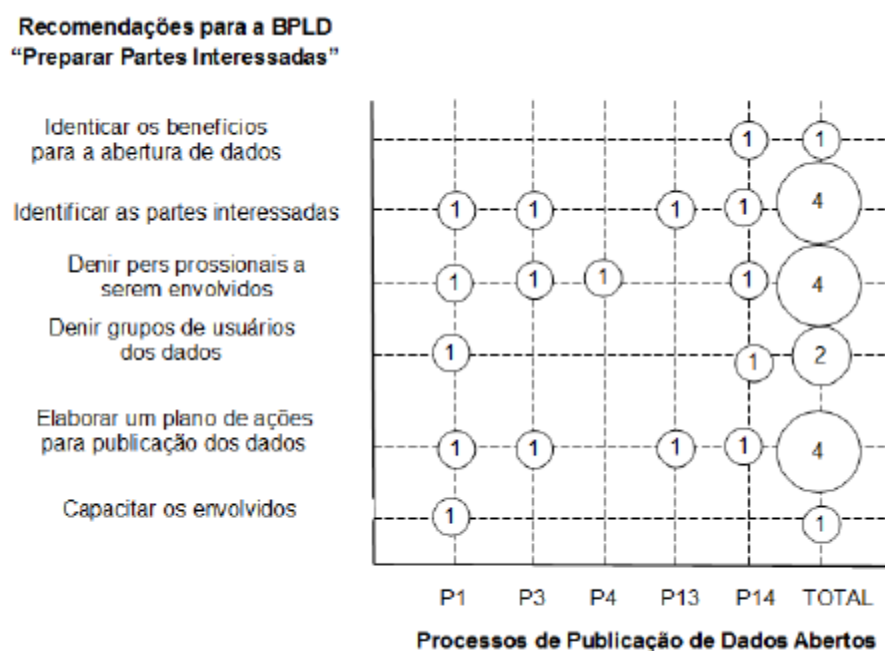
³⁴ <https://www.w3.org/TR/ld-bp/>

Código	Origem	Documentos (processos) Avaliados	Tipo
		(URUGUAY, 2012)	
P6	Itália	Geolinked Open Data for the Municipality of Catania (CONSOLI et al., 2014)	Dados Abertos Conectados Governamentais
P7	Internacional	LOP – Capturing and Linking Open Provenance on LOD Cycle (MENDONÇA et al., 2013)	Dados Abertos Conectados
P8	Internacional	TWC LOGD: A portal for linked open government data ecosystems (DING et al., 2011)	Dados Abertos Conectados Governamentais.
P9	Internacional	Linked Open Data: The Essentials - A Quick Start Guide for Decision Makers (BAUER; KALTENBÖCK, 2011)	Dados Abertos Conectados
P10	Internacional	The Joy of Data - A Cookbook for Publishing Linked Government Data on the web (HYLAND; WOOD, 2011)	Dados Abertos Conectados
P11	Grécia	Applying Linked Data Technologies to Greek Open Government Data: A Case Study (GALIOTOU; FRAGKOU, 2013)	Dados Abertos Conectados Governamentais
P12	Internacional	Managing the Life-Cycle of Linked Data with the LOD2 Stack (AUER et al., 2012)	Dados Abertos Conectados
P13	Espanha	Methodological Guidelines for Publishing Government Linked Data (VILLAZÓN-TERRAZAS et al., 2011)	Dados Abertos Conectados Governamentais

Código	Origem	Documentos (processos) Avaliados	Tipo
P14	União Europeia	Methodology for publishing datasets as open data - COMSODE; Documents of Practice for Methodology for publishing datasets as open data - COMSODE (COMSODE, 2014a; COMSODE, 2014a)	Dados Abertos Governamentais
P15	Internacional	Open Data Handbook (OKF, 2015)	Dados Abertos

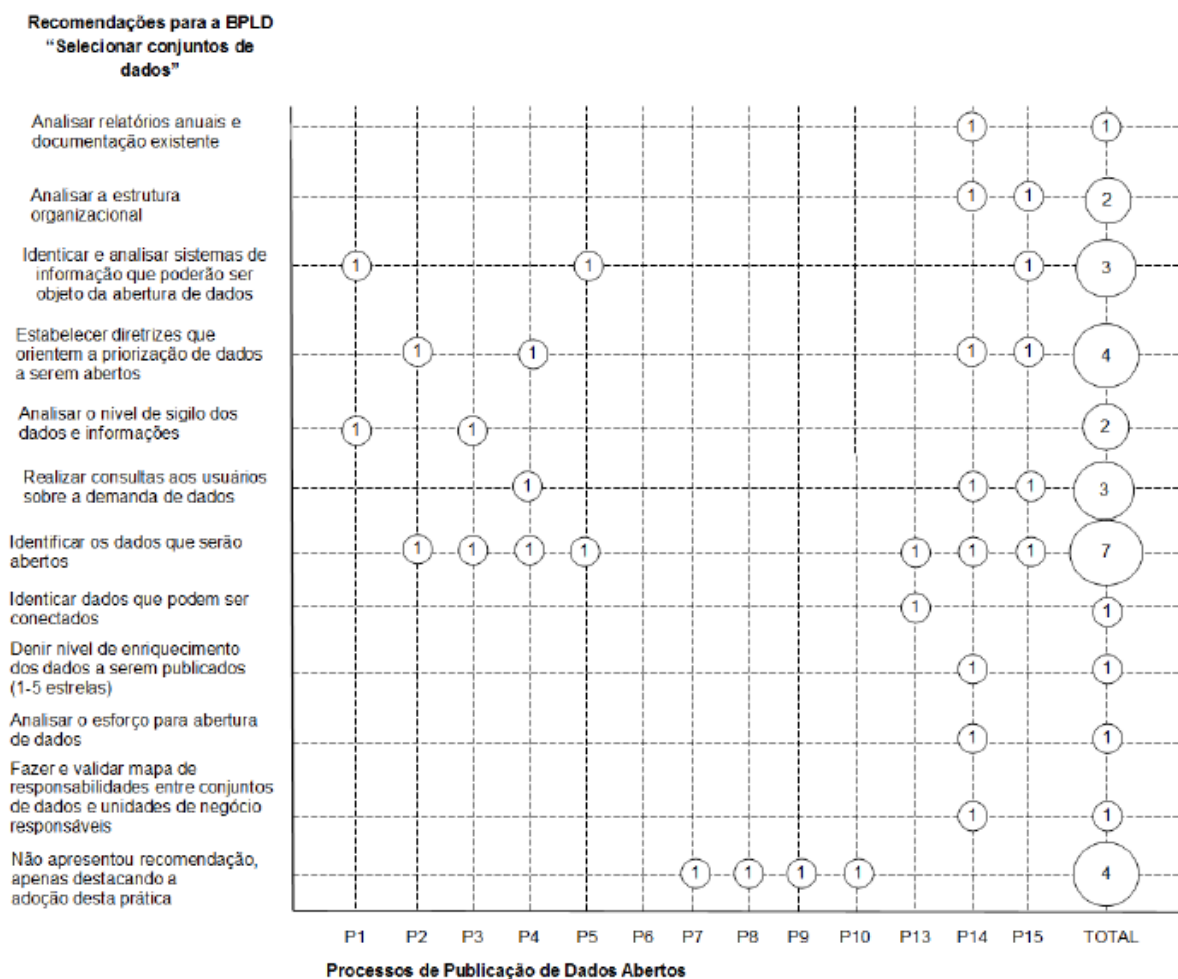
Fonte: ÁVILA, 2015.

Figura A.1 – Identificação de recomendações para a “Preparar Partes Interessadas” nos processos de publicação de dados abertos analisados



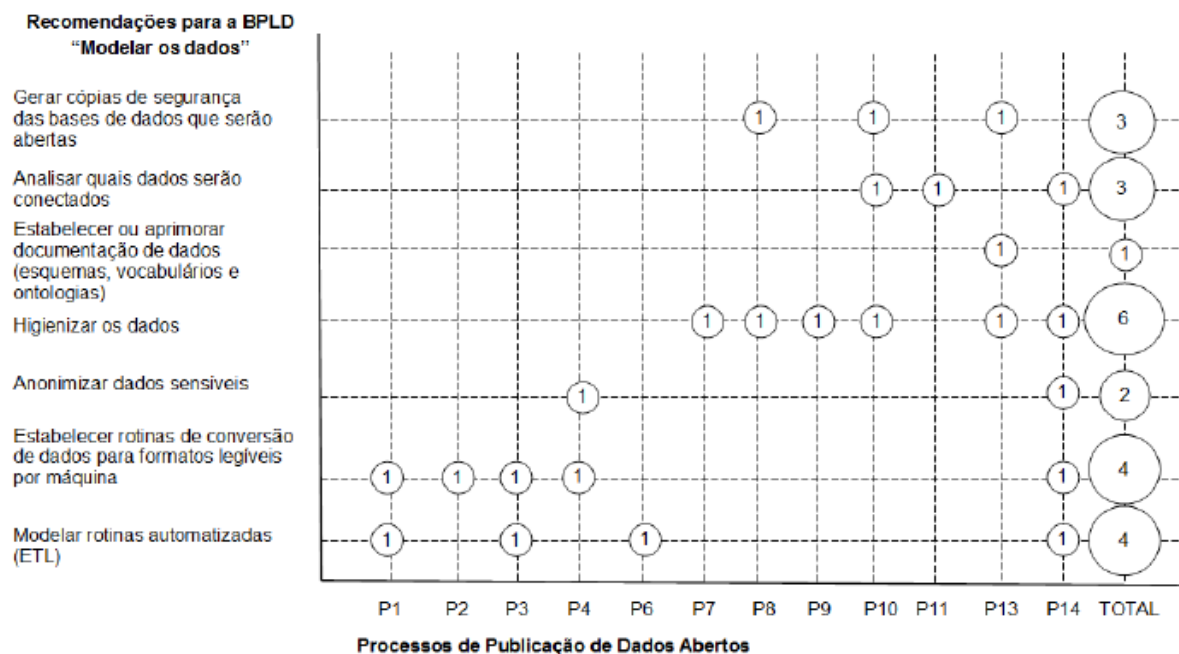
Fonte: ÁVILA, 2015.

Figura A.2 - Identificação de recomendações para a BPLD "Selecionar Conjuntos de Dados" nos processos de publicação de dados analisados



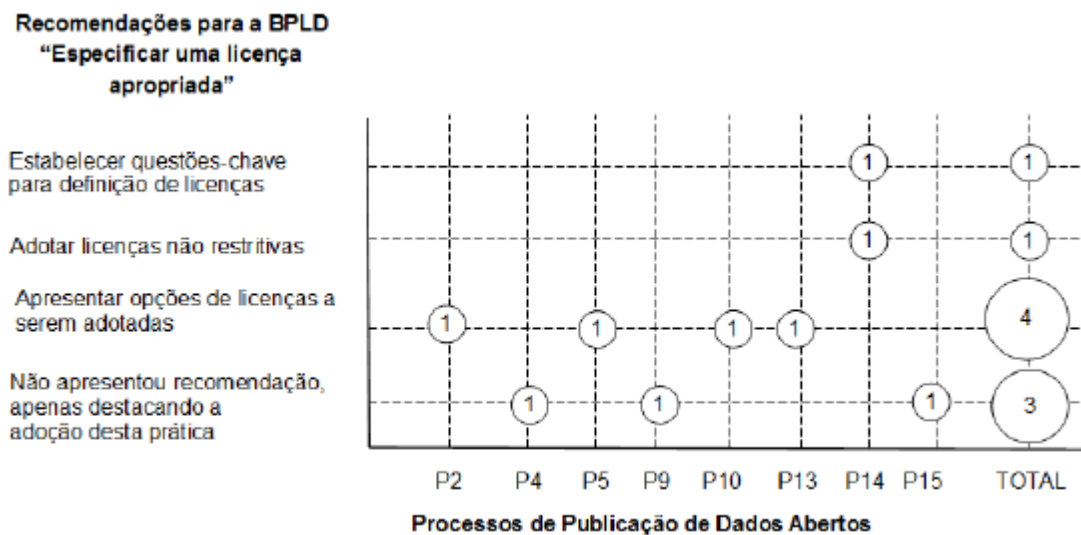
Fonte: ÁVILA, 2015.

Figura A.3 - Identificação de recomendações para a BPLD "Modelagem dos Dados" nos processos de publicação de dados abertos analisados



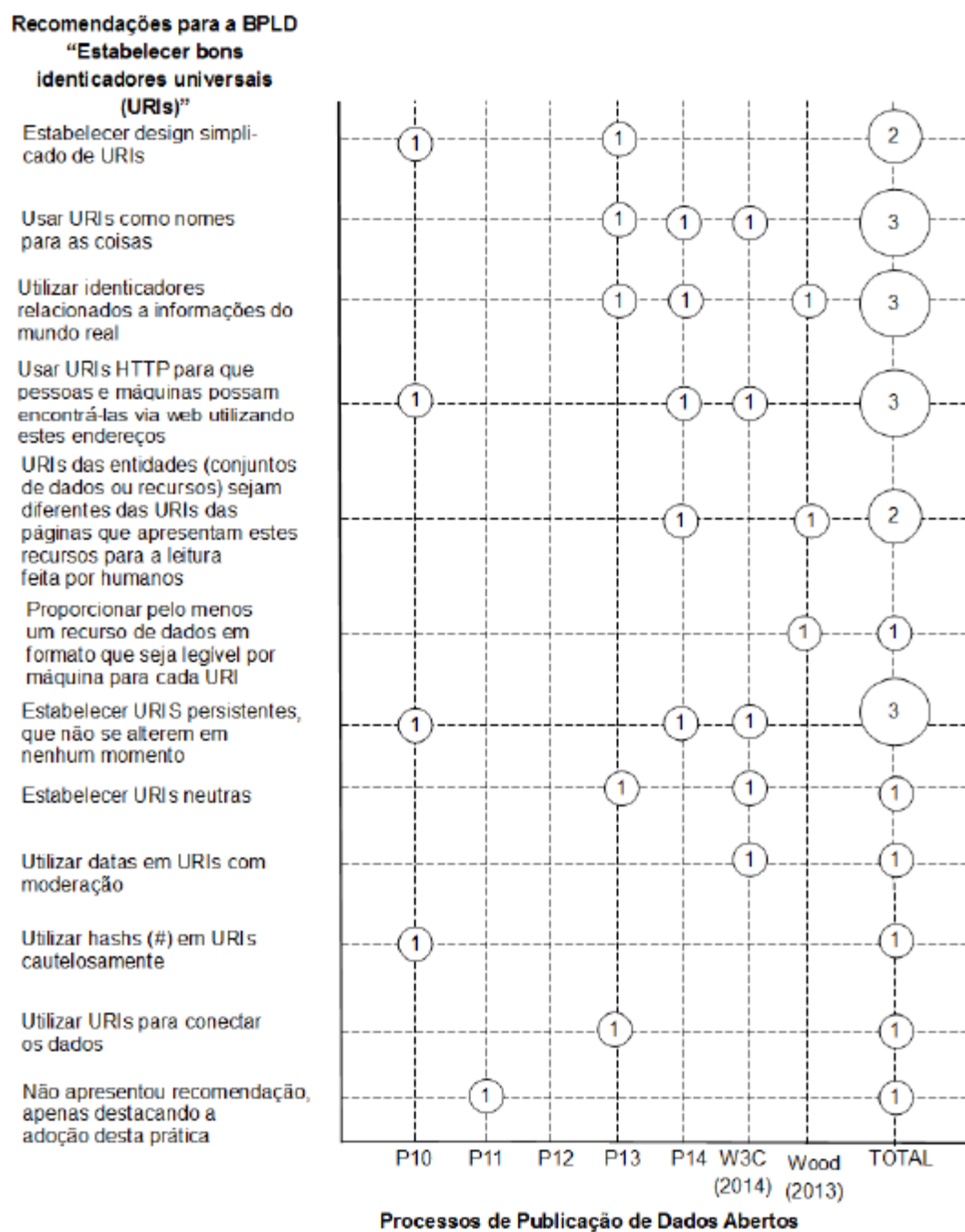
Fonte: ÁVILA, 2015.

Figura A.4 - Identificação de recomendações para a BPLD "Especificar uma licença apropriada" nos processos de publicação de dados abertos analisados



Fonte: ÁVILA, 2015.

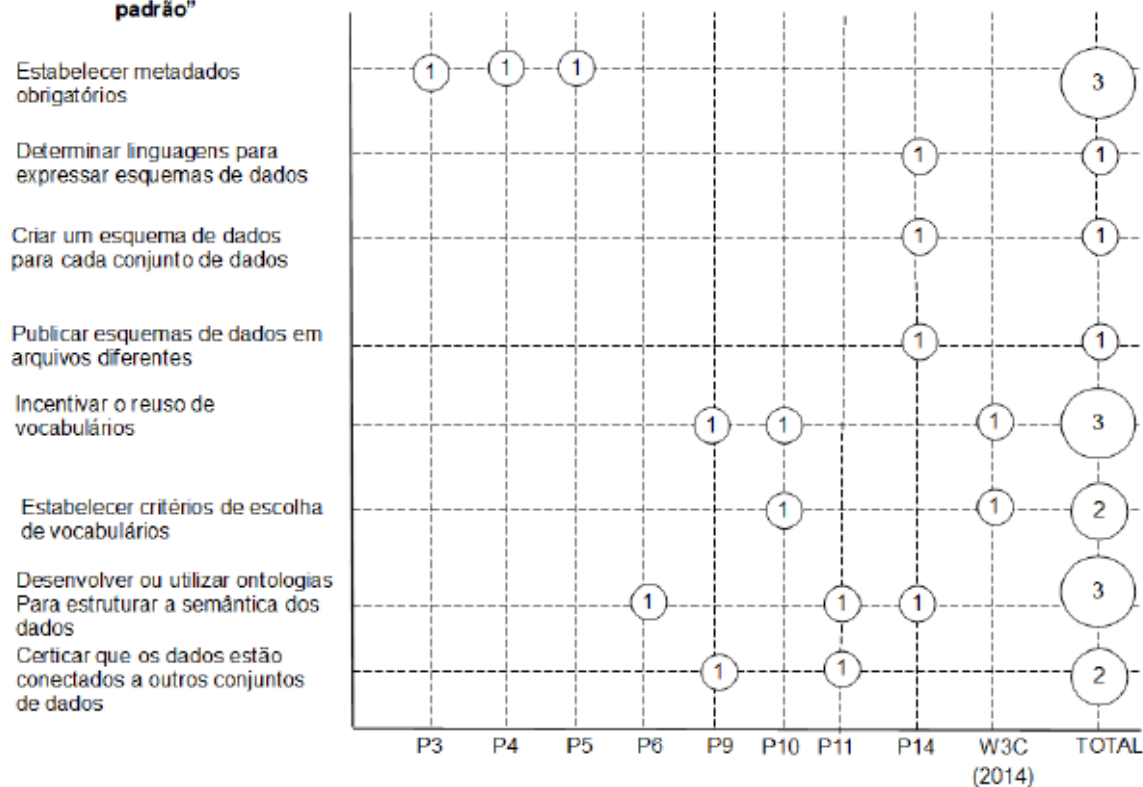
Figura A.5 - Identificação de recomendações para a BPLD "Estabelecer bons identificadores universais (URIs)" nos processos de publicação de dados abertos analisados



Fonte: ÁVILA, 2015

Figura A.6 - Identificação de recomendações para a BPLD "Utilização de vocabulário padrão" nos processos de publicação de dados abertos analisados

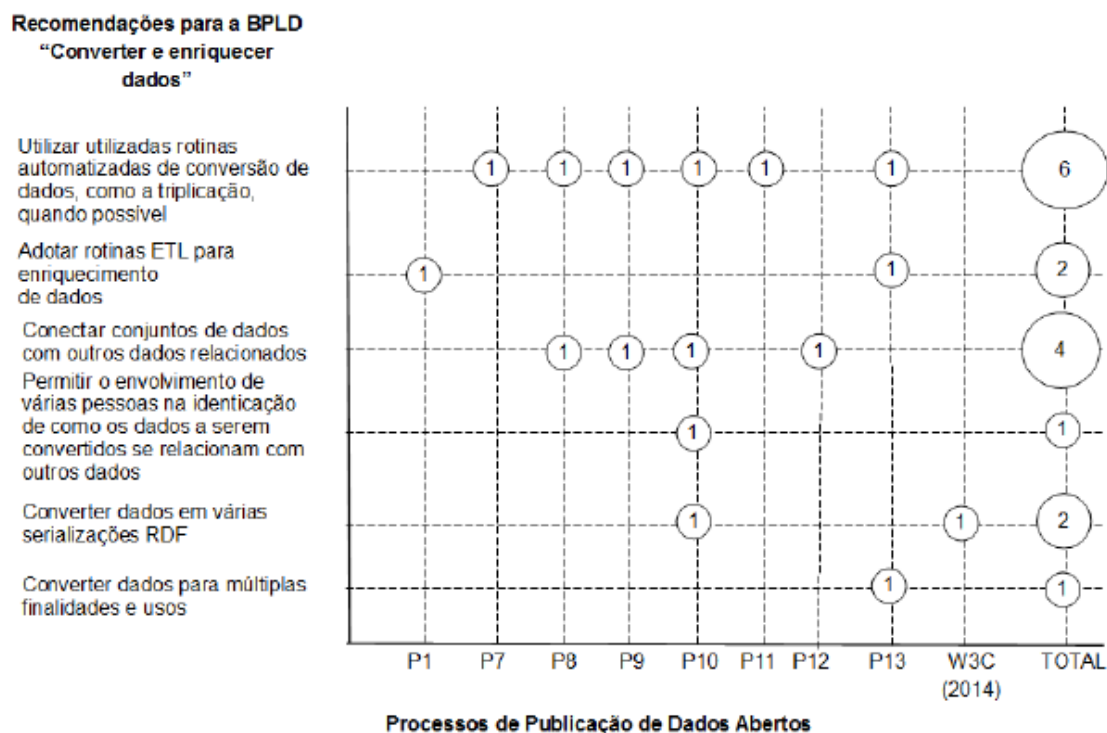
Recomendações para a BPLD
"Utilização de vocabulários
padrão"



Processos de Publicação de Dados Abertos

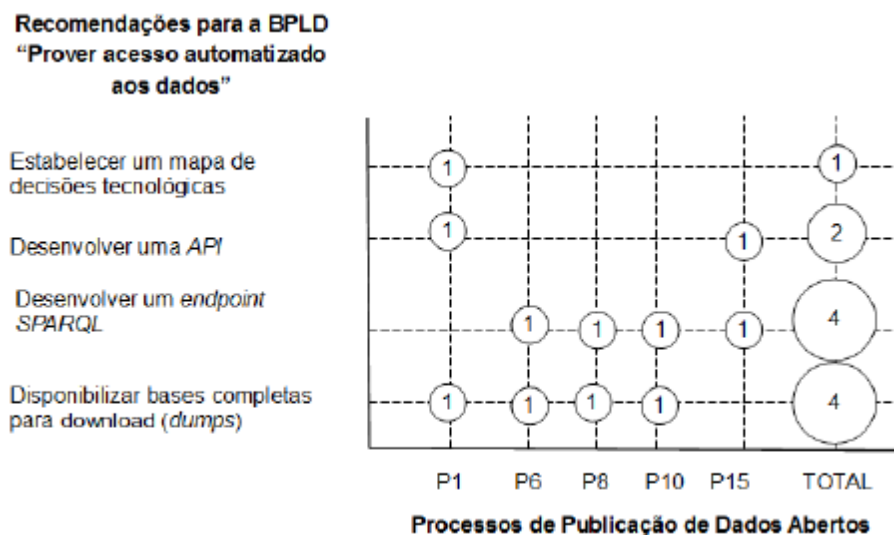
Fonte: ÁVILA, 2015

Figura A.7 - Identificação de recomendações para a BPLD "Converter e enriquecer dados" nos processos de publicação de dados abertos analisados



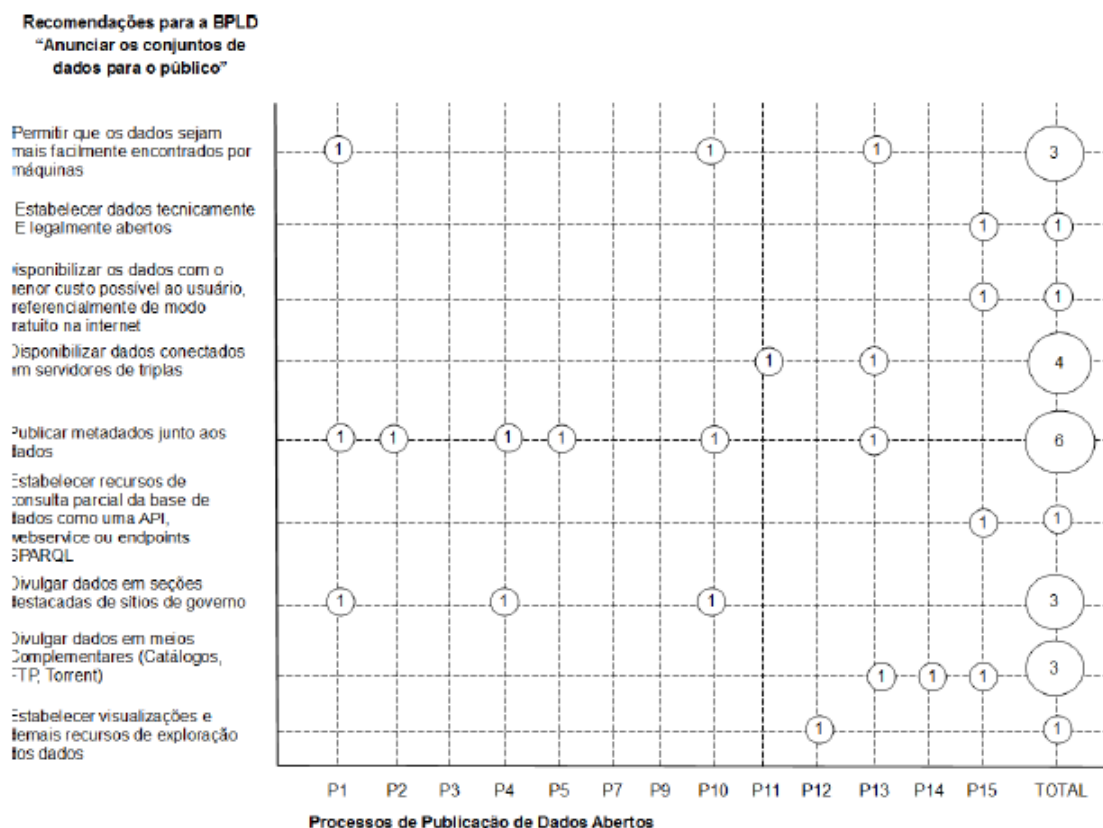
Fonte: ÁVILA, 2015

Figura A.8 - Identificação de recomendações para a BPLD "Prover acesso automatizado aos dados" nos processos de publicação de dados abertos analisados



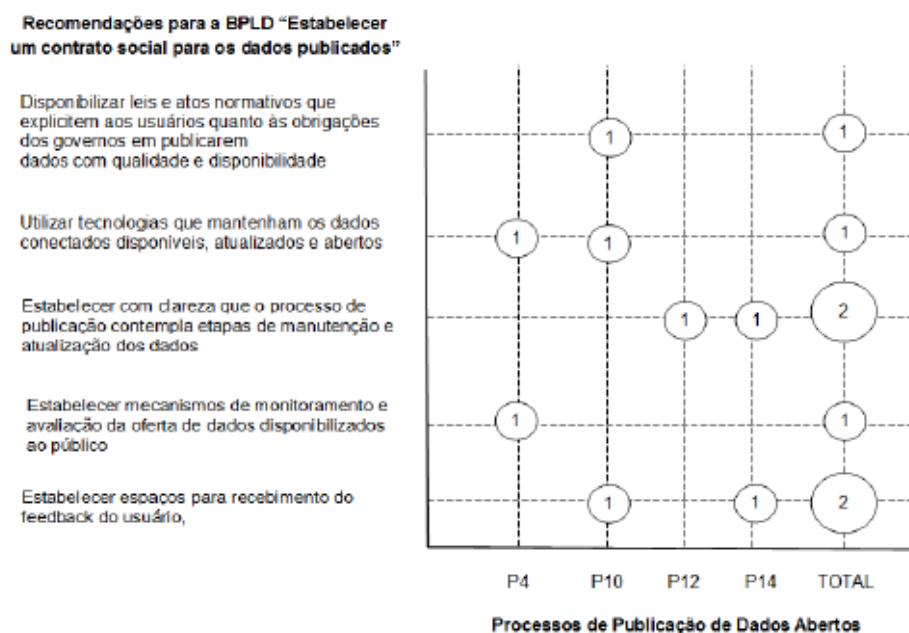
Fonte: ÁVILA, 2015

Figura A.9 - Identificação de recomendações para a BPLD "Anunciar os conjuntos de dados para o público" nos processos de publicação de dados abertos analisados



Fonte: ÁVILA, 2015

Figura A.10 - Identificação de recomendações para a BPLD "Estabelecer um contrato social para os dados publicados" nos processos de publicação de dados abertos analisados



Fonte: ÁVILA, 2015