



Universidade Federal de Pernambuco

Centro de Informática

Curso de Engenharia da Computação

**Uma ferramenta interativa para exploração do universo de
Inteligência Artificial Gerativa**

Trabalho de Conclusão de Curso de Graduação

por

Erik de Gouveia Zambom

Orientador: Prof. Filipe Carlos de Albuquerque Calegario

Recife, Outubro / 2023

Erik de Gouveia Zambom

**Uma ferramenta interativa para exploração do universo de Inteligência
Artificial Gerativa**

Monografia apresentada ao Curso de Engenharia da Computação, como requisito parcial para a obtenção do Título de Bacharel em Engenharia da Computação, Centro de Informática da Universidade Federal de Pernambuco.

Orientador: Prof. Filipe Carlos de Albuquerque Calegario

Recife

2023

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Zambom, Erik de Gouveia.

Uma ferramenta interativa para exploração do universo de inteligência artificial gerativa / Erik de Gouveia Zambom. - Recife, 2023.

26 p. : il.

Orientador(a): Filipe Carlos de Albuquerque Calegario

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Informática, Engenharia da Computação - Bacharelado, 2023.

1. IA gerativa. 2. aplicação web. 3. taxonomia. 4. modelos de IA. I. Calegario, Filipe Carlos de Albuquerque. (Orientação). II. Título.

000 CDD (22.ed.)

Agradecimentos

A meus pais, Ana Paula e Ricardo, e à minha irmã, Victória, por todo apoio e por me dar forças para persistir no caminho, enfrentando as diversas dificuldades encontradas.

A Juliano Cezar, pelo companheirismo, pelos conselhos e discussões sobre os mais variados temas, incluindo questões da vida, se tornando meio que um irmão de outra mãe.

A Vitor Matheus e Mariana Cazelli, com quem dividi moradia nos primeiros anos de graduação, sendo alguns dos primeiros vínculos que encontrei nessa terra até então desconhecida.

Ao prof. João Sedraz, por acreditar no meu potencial e conhecimento, mesmo estando no início da graduação, me dando a primeira oportunidade de demonstrar o que era capaz de fazer e, assim, prover um meio de me sustentar por conta própria longe de casa.

A Leandro Marques, Carlos José, Thiago Araújo, prof. Alex Sandro e todo o pessoal da Viitra/V-lab pela oportunidade de desbravar um seguimento da área de desenvolvimento que não havia tido contato até então, pelos novos conhecimentos, pelo aperfeiçoamento dos conhecimentos já existentes e pelos momentos vivenciados (muitos deles bons e alguns nem tanto).

Ao meu orientador, prof. Filipe Calegario, primeiramente por me apresentar esse universo tão interessante que é a área de IA gerativa e por aceitar me acompanhar no desenvolvimento desse trabalho.

À Júlia Cardouzo, minha namorada e parceira, por todo suporte, todo conselho, toda palavra de motivação e pela companhia nos inúmeros sábados enquanto fazíamos nossos projetos. Sinto que o desenvolvimento desse trabalho teria levado mais tempo sem o seu apoio.

Por último, mas não menos importante, meus mais sinceros agradecimentos a todos aqueles que, de alguma forma, passaram pela minha vida, mesmo que essa passagem tenha sido curta. O encontro com vocês certamente deixou resquícios na construção do meu ser que vou levar por toda a vida.

*Em algum lugar, algo incrível
está esperando para ser descoberto.*

Carl Sagan

RESUMO

Este trabalho tem como objetivo apresentar uma ferramenta interativa para permitir a exploração dos diversos modelos de Inteligência Artificial (IA) Gerativa, na forma de uma aplicação na web, juntamente com uma API Rest para consulta e manipulação das informações contidas na ferramenta supracitada, de modo a permitir a constante atualização e incremento da base de conhecimento. Um corpus de mais de 80 modelos de IA gerativa foi revisado para elaborar uma taxonomia descrevendo-os em termos dos seus tipos de dados de entrada e saída, bem como casos de uso e principal objetivo de geração. É disponibilizado uma aplicação web para a visualização e exploração dinâmica e lúdica do corpus de modelos, utilizando uma representação dos modelos como objetos 3D interativos, lançando mão de uma metáfora entre a área de IA gerativa e o universo (<https://project-galactus.onrender.com>).

Palavras-chave: IA gerativa, aplicação web, taxonomia, modelos de IA.

ABSTRACT

This paper presents an interactive tool to allow the exploration of various Generative Artificial Intelligence (AI) models, in the form of a web application along with a Rest API for querying and manipulating the information contained in the tool, allowing the expansion and constant update of the knowledge base. A corpus with more than 80 generative AI models was reviewed to elaborate a taxonomy describing them in terms of it's input/output data, as well as it's use cases and main objective of generation. It's provided a web application with 3D interactive objects to represent these models in order to provide a dynamic and lucid visualization and exploration of the corpus, making use of a metaphor between generative AI area and the universe (<https://project-galactus.onrender.com>).

Keywords: generative AI, web application, taxonomy, AI models.

LISTA DE FIGURAS

Figura 1	Exemplo de pergunta feita ao Google's Bard para gerar uma lista de modelos de IA gerativa	11
Figura 2	Exemplo de pergunta feita ao ChatGPT para conhecer mais detalhes de um modelo específico (resposta parcial)	12
Figura 3	Exemplo de pergunta feita ao Google's Bard para elaborar uma categorização dos modelos	13
Figura 4	Outro exemplo de pergunta feita ao Google's Bard para elaborar uma categorização dos modelos	13
Figura 5	Diagrama de sequência do Project Galactus	21

SUMÁRIO

1	INTRODUÇÃO	8
1.1	Contexto e Motivação	8
1.2	Objetivos	9
1.2.1	<u>Objetivo Geral</u>	9
1.2.2	<u>Objetivos Específicos</u>	9
1.3	Organização de texto	10
2	MONTAGEM DO CORPUS	11
2.1	Pesquisa	11
2.2	Categorização	12
3	SOLUÇÃO PROPOSTA	18
3.1	Banco de dados e API	18
3.2	Project Galactus	19
4	DISCUSSÃO	22
5	CONCLUSÃO E TRABALHOS FUTUROS	23
	REFERÊNCIAS	24

1 INTRODUÇÃO

Este capítulo descreve o contexto ao redor do tema desenvolvido no trabalho, além de apresentar os objetivos, as contribuições esperadas e as orientações sobre a organização da estrutura deste documento.

1.1 Contexto e Motivação

A área de Inteligência Artificial (IA) gerativa (ou generativa) já está em estudo há algum tempo, mas só ganhou esse nome recentemente. Trata-se de uma forma de inteligência artificial com a qual é possível criar novos conteúdos (texto, imagem, vídeo, código) a partir do treinamento a partir de grandes volumes de dados existentes. Nesse sentido, ela se diferencia da IA discriminativa (ou baseada em classificação) por criar algo totalmente novo e não apenas fazer distinções entre os tipos de entrada. Um dos primeiros registros de IA gerativa foi um *chatbot* baseado em um sistema de regras desenvolvido por Joseph Weizenbaum em 1966, chamado Eliza [1] [2].

No entanto, foi só em meados de 2022 que a área ganhou os holofotes do público, com o lançamento e disponibilização de grandes produtos como Midjourney [3], Stable Diffusion [4] e ChatGPT [5], que deram luz às altas capacidades dos modelos de IA gerativa nos campos de imagem e texto [6]. O grande sucesso e adesão do público a essas ferramentas foi o estopim para que novos avanços na área começassem a ser difundidos. De repente, todos queriam ter o seu próprio "ChatGPT". Não só as chamadas "Big Techs" como Google, Amazon, Meta, passaram a divulgar mais seus trabalhos (que já estavam em desenvolvimento há algum tempo), como também desenvolvedores independentes e pequenas empresas e *startups* começaram a explorar esse "novo" universo de possibilidades.

Nesse cenário, um grande número de aplicações e modelos começaram a ser apresentados a praticamente todo instante, nos mais variados canais e plataformas. Assim, acompanhar os avanços tem ficado cada vez mais complicado, especialmente para pessoas que estão buscando adentrar na área. Dessa forma, surge a necessidade de se ter uma espécie de repositório que centralize esses avanços, de modo que facilite o acesso e a exploração desse conteúdo.

Essa necessidade já é suprida de forma satisfatória na questão das ferramentas

e aplicações com o site "There's an ai for that" [7], que apresenta um compilado das ferramentas existentes desde 2015, sendo atualizado diariamente e com capacidade de se fazer pesquisas e exploração de uma *timeline* dos lançamentos. Todavia, no que tange os modelos de IA gerativa, a base para criação dessas aplicações apresentadas no site, uma ferramenta como essa ainda está em estágios iniciais de maturidade.

O mais próximo que temos de ferramentas como o "There's an AI for that" para os modelos são as *timelines* desenvolvidas por Fabian Mosele [8] e Dr. Alan D Thompson [9]. A primeira é focada em modelos de geração de imagens através de texto, sendo constantemente atualizada. Já a segunda, prioriza os modelos de linguagem e multimodais, dando maior ênfase para os modelos desenvolvidos no ocidente.

Há de se citar também a existência do Hugging Face [10], um repositório e concentrador de modelos, *datasets* e outros dados envolvendo IA. No entanto, seu conteúdo não é focado em IA gerativa e a sua exploração pode acabar sendo confusa.

Assim, uma ferramenta que organize os modelos de IA gerativa, sem restrição de nicho e localidade e que seja constantemente atualizada com os novos avanços, ainda é uma oportunidade a ser explorada.

1.2 Objetivos

1.2.1 Objetivo Geral

O objetivo **geral** deste trabalho é criar uma ferramenta interativa que permita a exploração dos variados modelos de IA gerativa existentes.

1.2.2 Objetivos Específicos

São objetivos **específicos** deste trabalho:

- Analisar, compreender e catalogar diversos modelos de IA gerativa existentes nas mais variadas modalidades de geração.
- Identificar elementos de semelhança entre os modelos para agrupá-los em uma estrutura organizacional.
- Visualizar e explorar a interação com as informações sobre os modelos.
- Elaborar um meio que facilite a atualização da base de dados dos modelos.

1.3 Organização de texto

No **primeiro capítulo**, é feita uma contextualização e apresentação das motivações desse trabalho, bem como uma demonstração dos objetivos, das contribuições esperadas e da estrutura do trabalho.

No **segundo capítulo**, é relatado o processo de coleta de dados e construção do corpus base da aplicação desenvolvida durante o projeto, explicitando como foi feita a divisão e categorização das informações.

A seguir, no **terceiro capítulo**, é descrito o processo de desenvolvimento da aplicação, explicando as decisões tomadas acerca das tecnologias utilizadas e do formato de apresentação dos conteúdos coletados.

Posteriormente, o **quarto capítulo** apresenta uma discussão dos resultados encontrados no desenvolvimento, bem como os desafios e limitações enfrentados durante o processo.

Por fim, o **último capítulo** apresenta os pensamentos finais em relação ao projeto, além de discutir as limitações técnicas e fazer apontamentos para trabalhos futuros.

2 MONTAGEM DO CORPUS

Neste capítulo, é descrito o processo de montagem do corpus de modelos de IA gerativa usado como base para o desenvolvimento do projeto, composto por duas etapas principais, sendo elas **pesquisa** e **categorização**, descrita em mais detalhes a seguir.

2.1 Pesquisa

A montagem do corpus começou por uma pesquisa exploratória dos modelos de IA gerativa existentes atualmente, com a finalidade de coletar informações como uma breve descrição do modelo, o tipo dos artefatos de entrada e saída e links para artigos e páginas de implementação ou exibição dos referidos modelos.

A priori, foi montada uma lista dos modelos a partir de consultas em repositórios [11] [12] [13] [14] e artigos [15] [16], além de perguntas à ferramentas como ChatGPT [5] e Google's Bard [17] como demonstrado em Figura 1.



Figura 1: Exemplo de pergunta feita ao Google's Bard para gerar uma lista de modelos de IA gerativa

A seguir, foi-se explorado os artigos base dos modelos, bem como suas páginas de apresentação e implementação de código (quando disponíveis) para a coleta das informações. Nessa fase, também se deu o uso do ChatGPT [5] e do Google's Bard [17] para auxiliar no processo de entendimento mais detalhado de alguns dos modelos, como demonstrado em Figura 2.

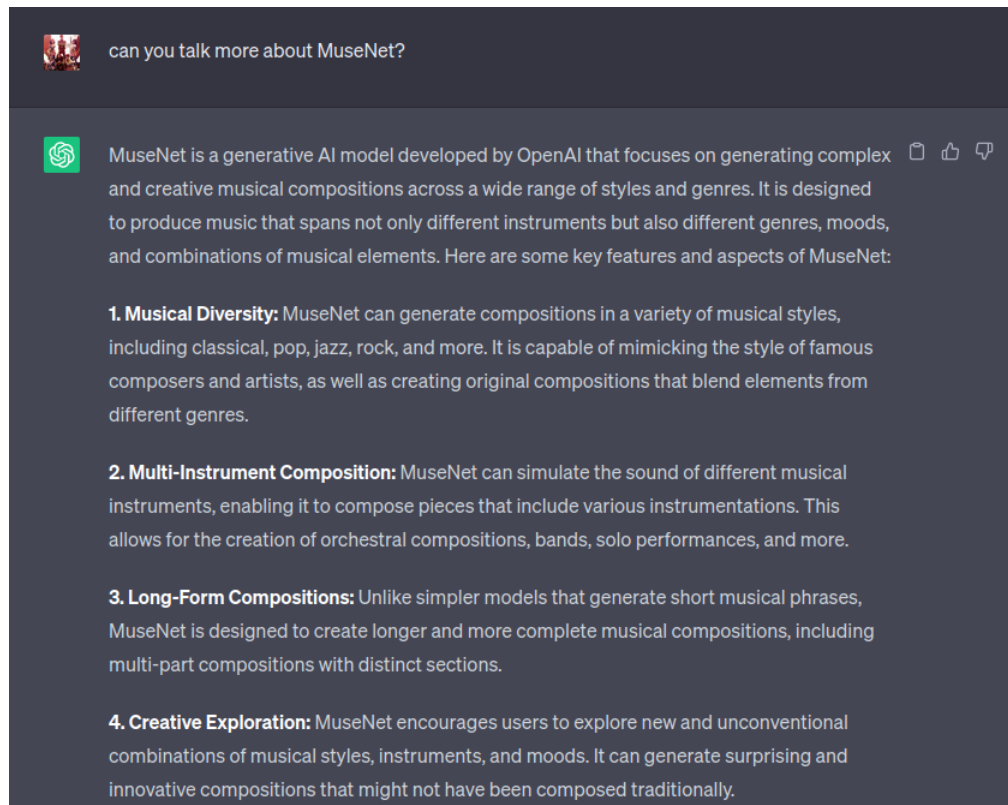


Figura 2: Exemplo de pergunta feita ao ChatGPT para conhecer mais detalhes de um modelo específico (resposta parcial)

2.2 Categorização

Com os modelos listados juntamente com suas informações, foi-se pensado uma forma de agrupar modelos semelhantes sob um mesmo rótulo. Aqui, também foi feita uma consulta ao ChatGPT [5] e ao Google's Bard [17] sobre maneiras de se categorizar os modelos, exemplificado em Figura 3 e Figura 4.

Assim, baseado nas conversas supracitadas, na análise dos modelos e em taxonomias apresentadas em outros trabalhos [15] [16], foi-se elaborado uma categorização em dois níveis: o primeiro nível, mais macro, considerando majoritariamente o tipo de artefato gerado na saída do modelo, sendo dividido nas categorias **Texto**, **Imagem**, **3D**, **Audio** e **Video**; o segundo nível, sendo uma subdivisão das categorias do primeiro nível, onde foram considerados o tipo do artefato de entrada, o propósito principal do modelo quanto aos seus casos de uso e as diferentes subcategorizações do tipo do artefato de saída.

É necessário salientar, no entanto, que um mesmo modelo tem potencial para pertencer a mais de uma categoria e/ou subcategoria. Nesse sentido, foi-se escolhido registrar o modelo apenas uma vez, na categoria/subcategoria que melhor representa suas

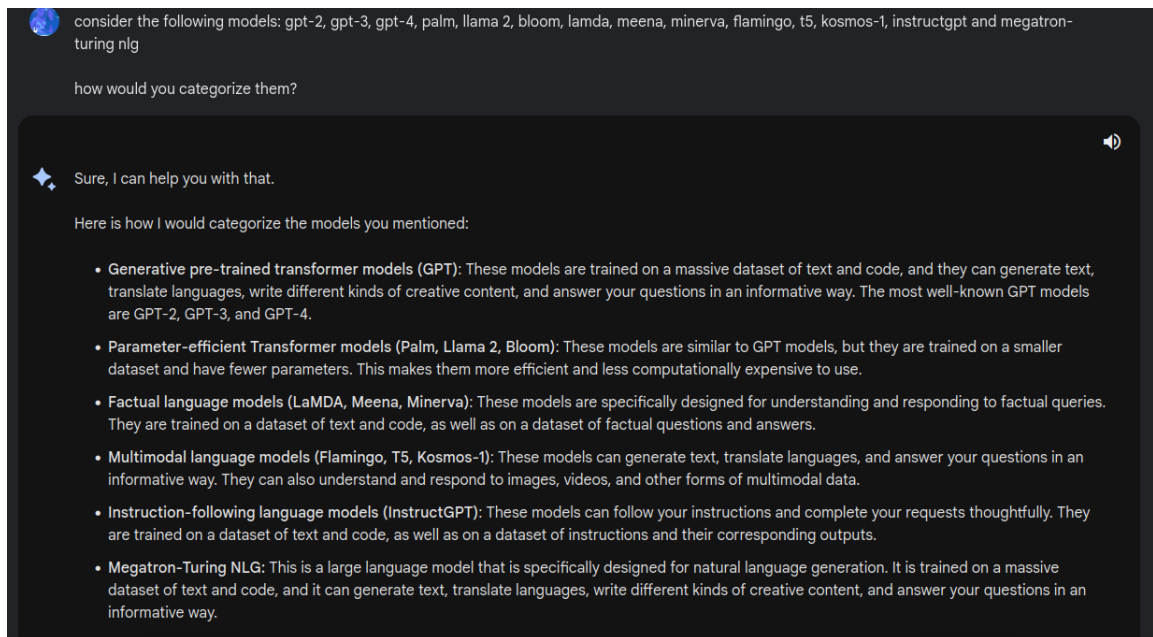


Figura 3: Exemplo de pergunta feita ao Google's Bard para elaborar uma categorização dos modelos

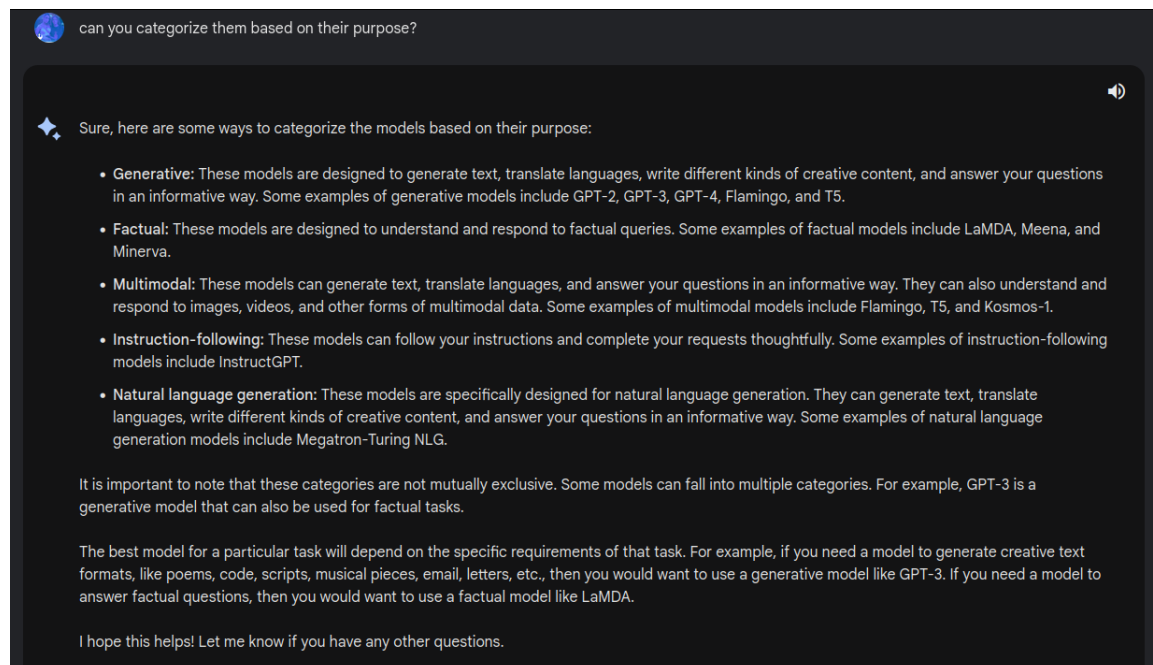


Figura 4: Outro exemplo de pergunta feita ao Google's Bard para elaborar uma categorização dos modelos

características e casos de uso.

Uma planilha com a lista de modelos e sua categorização é disponibilizada em Generative AI Models e um detalhamento das categorias (e respectivas subcategorias) é apresentado abaixo.

Texto

Nessa categoria, encontram-se a maior parte dos modelos cujo artefato de saída é representado de forma textual, podendo ser desde uma frase de resposta a uma pergunta até trechos de código nas mais variadas linguagens. Apresentam-se como suas subcategorias:

- **Criativo:** modelos treinados com um grande *dataset* de textos e códigos, capazes de gerar diversos tipos de conteúdos textuais criativos;
- **Raciocínio:** muitos dos modelos aqui podem ser enquadrar também na subcategoria **criativo**, todavia seu enfoque é em raciocinar sobre o artefato de entrada e, a partir disso, gerar uma saída;
- **Saúde/Ciência:** modelos que produzem artefatos de texto em viés científico ou na área de saúde, capaz de gerar conselhos médicos, explicar artigos, etc;
- **Legendagem/Descrição:** incorpora os modelos cujo principal objetivo é apresentar uma descrição ou legenda para o artefato de entrada, sendo esse artefato uma imagem, um áudio, um vídeo;
- **Código:** modelos com capacidade de interpretar, corrigir e gerar blocos de código nas mais variadas linguagens;
- **Instrução:** modelos mais focados em seguir um conjunto de instruções para completar tarefas;
- **Jogos:** modelos capazes de gerar conteúdos de jogos, desde personagens até níveis inteiros.

Imagem

Aqui se enquadram os modelos capazes de gerar imagens estáticas como artefato de saída, aceitando *prompts* de texto, áudios, outras imagens, etc. Suas subcategorias são:

- **Texto para imagem:** nessa categoria estão agrupados os modelos cujo principal objetivo é gerar imagens a partir de descrições textuais, apesar de muitos deles também ser capaz de realizar alterações em imagens passadas como artefato de entrada;
- **Áudio para imagem:** modelos que geram imagens a partir de descrições via áudio;
- **Edição:** modelos especializados em realizar modificações e uma imagem passada como artefato de entrada, geralmente baseada em uma descrição textual da alteração;
- **Imagem multiplano:** modelos capazes de gerar imagens com múltiplos planos, ou seja, imagens com sensação de profundidade;
- **Sketch:** modelos que trabalham com *sketchs*, seja utilizando-os como artefato de entrada para transformá-los em imagens finais ou simplificando uma imagem trabalhada a um *sketch*;
- **Vectorização:** modelos capazes de gerar imagens vetorizadas, que podem ser altamente escaláveis;
- **Imagens com texto:** modelos feitos para criar imagens contendo um elemento textual, de forma que o texto fique natural e legível na imagem.

Audio

Essa categoria traz os modelos especializados em gerar arquivos de áudio, podendo esse áudio ser a fala de uma pessoa, uma conversa ou uma música. É dividida em:

- **Geração de fala:** modelos que são capazes de converter textos em áudio;
- **Saúde/Ciência:** modelos que produzem artefatos de áudio com viés científico ou na área de saúde;
- **Geração de música:** modelos que possibilitam a criação de faixas de áudio contendo sons de instrumentos, dos mais simples aos mais complexos, além de potencialmente a voz, obedecendo conceitos musicais como ritmo, harmonização, etc;
- **Conversão de voz:** modelos especializados em transformar a voz reproduzida em um áudio para outra, como se emulasse a fala de outra pessoa;

- **Continuação de discurso:** modelos que geram áudios complementares ao passado como artefato de entrada, mantendo as características presentes no mesmo.

Video

Agrupam-se nessa categoria os modelos que tem por finalidade gerar um artefato de vídeo, seja pela edição de um já existente ou através de uma descrição textual do desejado, entre outros. Organizam-se da seguinte forma:

- **Geração:** modelos que visam a geração de vídeos a partir, principalmente, de uma descrição textual;
- **Edição:** modelos eficazes em realizar alterações em vídeos existentes, com a inclusão de elementos ou a mudança de estilo do mesmo;
- **Reconstrução:** modelos capazes de reconstruir vídeos a partir da sua representação latente reduzida;
- **Clipe de música:** modelos especializados na geração de clipes musicais;
- **Pesquisa de cena:** modelos capazes de fazer um recorte de um vídeo mais longo que represente um descrição textual passada como artefato de entrada;
- **Aumento de qualidade/FPS:** modelos que visam aprimorar a qualidade de um vídeo passado como artefato de entrada, seja por meio de aumento da resolução, da geração de *frames* complementares ou outros meios.

3D

Os modelos enquadrados nessa categoria são aqueles destinados à geração de objetos 3D, indo desde de texturas até cenas completas, separados nas seguintes subcategorias:

- **Texto para modelo:** modelos capazes de gerar modelos de objetos 3D a partir de descrições textuais;
- **Imagem para modelo:** modelos que geram modelos de objetos 3D a partir de imagens de referência;
- **Nuvem de pontos para modelo:** modelos capazes de transcrever uma nuvem de pontos para um modelo de objeto 3D;

- **Geração de cena:** modelos que possibilitam a geração de uma cena 3D completa a partir de imagens de referência ou descrições textuais;
- **Texto para nuvem de pontos:** modelos que geram uma nuvem de pontos que representa um objeto 3D a partir de uma descrição textual;
- **Imagem para humano:** modelos capazes de gerar um modelo 3D de um humano baseado em imagens de referência;
- **Saúde/Ciência:** modelos que produzem artefatos 3D com viés científico ou na área de saúde.

3 SOLUÇÃO PROPOSTA

Este capítulo discorre sobre a elaboração da ferramenta desenvolvida para elaboração do corpus coletado, sendo o produto final deste trabalho. É dividido em duas seções, onde a primeira discute sobre a escolha do banco de dados e da construção da api e a segunda aborda o desenvolvimento da ferramenta em si.

3.1 Banco de dados e API

Com o corpus (base de dados) estabelecido, foi o momento de elaborar a ferramenta de exploração. Como estabelecido previamente, a ferramenta precisa ser de fácil acesso e exploração, além de ser capaz de mostrar os modelos de forma organizada e ter a possibilidade de ser atualizada constantemente de maneira simples.

Esses dois últimos requisitos são facilmente cumpridos com o uso de um banco de dados para armazenamento das informações dos modelos e uma API REST [18] para realizações de operações nesses dados, permitindo ler, adicionar, modificar e excluir quaisquer informação com rapidez e simplicidade.

Como os dados do corpus apresentam uma estrutura bem definida e visando garantir a manutenção dessa estruturação, foi-se escolhido um banco de dados relacional para se armazenar essas informações, utilizando o PostgreSQL [19] como sistema gerenciador de banco de dados. O fato da maior parte das informações serem textuais e a possibilidade existencial de um sistema de pesquisa nessas informações poderia justificar a escolha de um sistema de banco de dados não relacional, no entanto, a garantia da manutenção da estruturação dos dados se apresentam como uma característica mais vantajosa que que uma potencial melhora de performance nas buscas.

Já para a API REST foi-se escolhido a construção utilizando NodeJS [20] e Express [21], dado a sua facilidade e rapidez, tanto na construção quanto na manutenção da mesma, além da simplicidade de conexão e gerenciamento do banco de dados através de pacotes como o Sequelize [22].

A escolha de desenvolver uma API REST em detrimento à outras possibilidades arquiteturais, como cliente-servidor, se deu majoritariamente por conta da capacidade de se acoplar outros serviços à aplicação, além de permitir que outras ferramentas utilizem a mesma base de dados sem que seja necessário um grande esforço na sua integração.

O acesso à API REST desenvolvida, bem como a documentação dos seus *endpoints*, pode ser feito através do link <http://project-galactus-api.onrender.com>.

3.2 Project Galactus

O requisito de fácil acesso e exploração fica por conta do método utilizado para, de fato, mostrar os dados, ou seja, o *frontend* da aplicação. Nesse ponto, várias abordagens poderiam ser utilizadas como gráficos, tabelas e até mesmo linhas do tempo como visto nos trabalhos citados anteriormente. Todavia, todas essas abordagens podem se tornar maçantes e de difícil visualização e entendimento dado ao grande volume de informação, podendo não passar de forma muito efetiva as relações existentes nas informações.

Com isso em mente, uma nova análise nos dados coletados foi feita, observando principalmente a forma como os dados estavam estruturados e suas relações. Nisso, chegou-se à uma interessante conclusão: a área de IA gerativa se organiza, e potencialmente se comporta, de forma semelhante ao próprio universo!

Pensando na estrutura básica do universo, nota-se que ele é majoritariamente composto por galáxias, onde cada galáxia é composta por um conjunto de sistemas estelares/planetários. Esses sistemas, por sua vez, são formados por estrelas orbitadas por planetas, que podem ou não conter satélites os orbitando. Além disso, muitos desses elementos apresentam semelhanças, mas dificilmente são encontrados dois elementos iguais um ao outro.

Como demonstrado anteriormente, a área da IA gerativa pode ser dividida em categorias (galáxias) que podem ser subdivididas em áreas menores (sistemas estelares) nas quais se agrupam os modelos (planetas). Indo além, é possível observar as ferramentas e aplicações geradas a partir dos modelos como os satélites desses modelos representados pelos planetas.

Outro ponto a ser observado é que, assim como o próprio universo, onde novos planetas, estrelas e até mesmo galáxias podem ser descobertos a cada instante [23], a área de IA gerativa também está em constante expansão, com novos modelos, aplicações e ferramentas surgindo a cada instante, como já foi citado.

Com essa metáfora em mente, foi desenvolvido o **Project Galactus**, uma aplicação web que apresenta os dados coletados no corpus como objetos 3D representando elementos astronômicos como planetas, estrelas e galáxias.

Para o desenvolvimento foi escolhido a biblioteca React [24], utilizando Three.js [25] e React Three Fiber [26] para a renderização de objetos 3D no navegador, sendo parte da base de codificação dos elementos 3D retirados de tutoriais [27] [28] e outros códigos [29].

A aplicação é composta por três páginas principais. Na primeira é apresentado as categorias estabelecidas no corpus, visualmente representadas por galáxias. Navegando para a segunda página através da interação com uma das galáxias, tem-se as subcategorias da categoria escolhida, sendo representadas por estrelas. Por fim, a terceira página, acessada de forma semelhante à segunda, agora por meio da interação com uma das estrelas, apresenta os modelos presentes na categorização navegada, representados por planetas.

Ao acessar cada página, é feita uma requisição GET para a api para recuperar as informações a serem exibidas. Além disso, na primeira vez que um elemento (galáxia, estrela ou planeta) é apresentado em tela, são gerados parâmetros para a sua exibição, como sua cor, posição na tela, entre outros. Esses parâmetros são então enviados para a api via requisição PUT para serem armazenados no banco de dados, possibilitando sua recuperação posterior e a consistência das visualizações.

Um diagrama representando o fluxo básico de funcionamento descrito acima pode ser visto em Figura 5.

O acesso à aplicação pode ser feito em <https://project-galactus.onrender.com> e seu código encontra-se disponível em <https://github.com/Zambom/project-galactus>.

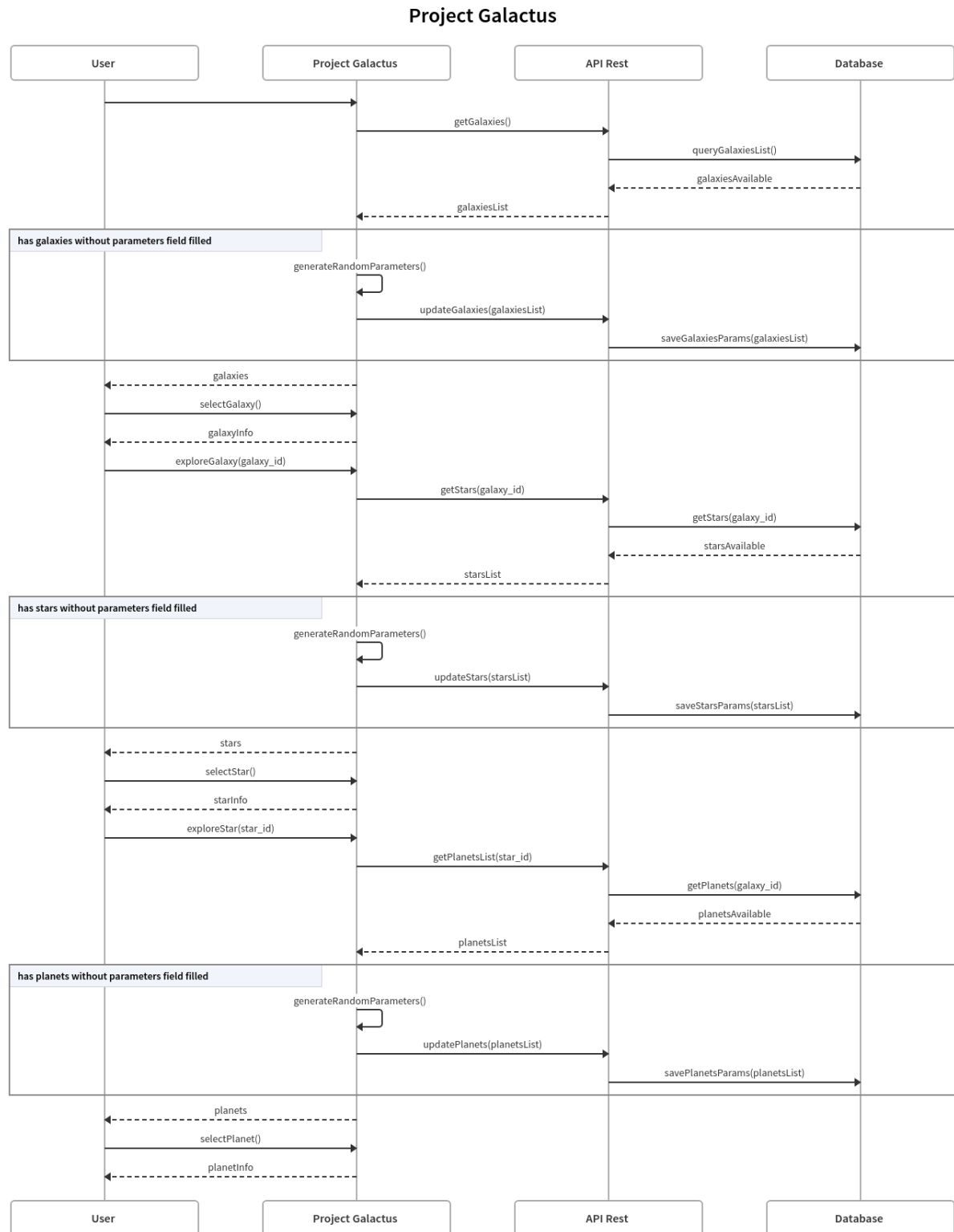


Figura 5: Diagrama de sequência do Project Galactus

4 DISCUSSÃO

Como estabelecido nos capítulos iniciais deste trabalho, o objetivo principal era criar uma ferramenta interativa que possibilitaria a exploração dos modelos de IA gerativa disponíveis atualmente. Nesse sentido, a solução apresentada tem potencial para atingir esse propósito de forma satisfatória ao agrupar e organizar diversos modelos em um único local, lançando mão de uma metáfora para representação da taxonomia do conteúdo.

Outra contribuição relevante é a própria API REST utilizada para fornecimento dos dados da aplicação, que permite o desenvolvimento de novas ferramentas de visualização, além de se mostrar eficiente em outro objetivo informado: a questão de manter as informações atualizadas e com a possibilidade de expandi-las.

Além disso, o estudo realizado sobre a classificação dos modelos pode servir de base para futuras discussões visando o melhor entendimento sobre como a área de IA gerativa se organiza e qual a melhor forma de classificar seus elementos.

Por outro lado, a aplicação apresenta algumas limitações técnicas que ainda precisam ser contornadas. A principal dessas limitações é o número de elementos 3D que a ferramenta consegue mostrar na tela, por conta do alto nível de processamento necessário para *renderizar* os objetos. Além disso, garantir que os elementos não ficariam sobrepostos uns aos outros sem que os cálculos aumentassem a complexidade da aplicação também foi um desafio.

O próprio processo de desenvolvimento demonstrou algumas complicações. Dado o pouco conhecimento das tecnologias que envolvem o desenvolvimento de páginas com visualizações de objetos 3D, boa parte do tempo disponível foi investido em seu aprendizado. Dessa maneira, alguns aspectos precisaram ser negligenciados, como a estética visual dos elementos *html* da página, tendo em vista que seu aspecto mais importante era a funcionalidade de mostrar a informação e/ou navegar na ferramenta.

Essas dificuldades de desenvolvimento atrasaram o cronograma tomando o tempo que seria necessário para a etapa de testes e avaliações públicas. No entanto, a ferramenta está disponível para acesso geral, de modo que testes possam ser realizados em trabalhos futuros.

5 CONCLUSÃO E TRABALHOS FUTUROS

Este trabalho apresenta uma aplicação web que, fazendo uso de uma metáfora entre a área da IA gerativa e o universo, apresenta uma forma de exploração interativa e lúdica dos modelos de IA gerativa reunidos em um corpus composto por 86 modelos, sendo possível adicionar informações à esse corpus por meio de uma API REST desenvolvida em conjunto, de modo a ser uma boa fonte de consulta e aprendizado da área de IA gerativa.

Também demonstra um estudo sobre a taxonomia dos modelos de IA gerativa, apresentando uma forma de classificar e organizar esses modelos, podendo servir de base para futuras discussões sobre o tema, em busca de maior entendimento sobre a melhor forma de fazer essa organização.

Como trabalhos futuros incluem-se o aumento e atualização das informações da base de dados, além da otimização do *frontend* da aplicação, de modo a permitir a renderização de um maior número de elementos 3D na tela, além de aprimorar a sua utilização em dispositivos móveis.

Além disso, outras aplicações podem ser desenvolvidas para mostrar as informações da base de dados em diferentes estilos de visualização, bem como a adição de novas funcionalidades como um sistema de busca e filtragem dos modelos.

O incremento da visualização para considerar também as aplicações de cada modelo é outra questão a ser observada em contribuições futuras.

Por fim, é observado um potencial em abordar a metáfora do universo para criação de ferramentas de exploração em outras áreas, não só de Inteligência Artificial, como também de outros ramos da tecnologia, dado que muitos desses conhecimentos podem ser organizados em uma estrutura semelhante ao apresentado.

REFERÊNCIAS

- [1] Towards AI. *The Generative AI Revolution: Exploring the Current Landscape*. Available at: <<https://pub.towardsai.net/the-generative-ai-revolution-exploring-the-current-landscape-4b89998fcc5f>>. Accessed in: 22/08/2023.
- [2] Josh Fruhlinger. *O que é a IA generativa?* Available at: <<https://itforum.com.br/computerworld/o-que-ia-generativa/>>. Accessed in: 11/09/2023.
- [3] MIDJOURNEY. *Midjourney*. Available at: <<https://www.midjourney.com/>>. Accessed in: 22/08/2023.
- [4] COMPVIS. *Stable Diffusion*. Available at: <<https://github.com/CompVis/stable-diffusion>>. Accessed in: 22/08/2023.
- [5] OPENAI. *ChatGPT*. Available at: <<https://chat.openai.com>>. Accessed in: 22/08/2023.
- [6] Elad Gil. *Early days of AI*. Available at: <<https://blog.eladgil.com/p/early-days-of-ai>>. Accessed in: 22/08/2023.
- [7] ANDREI. *There's an Ai for that*. Available at: <<https://theresanaiforthat.com/>>. Accessed in: 22/08/2023.
- [8] MOSELE, F. *AI Timeline*. Available at: <<https://www.fabianmosele.com/ai-timeline>>. Accessed in: 22/08/2023.
- [9] THOMPSON, A. D. *LLM timeline*. Available at: <<https://lifearchitect.ai/timeline>>. Accessed in: 22/08/2023.
- [10] Hugging Face. *Hugging Face – The AI community building the future*. Available at: <<https://huggingface.co/>>. Accessed in: 11/09/2023.
- [11] CALEGARIO, F. *Awesome Generative AI*. Available at: <<https://github.com/flipecalegario/awesome-generative-ai>>. Accessed in: 24/08/2023.

- [12] MICROSOFT. *unilm*. Available at: <<https://github.com/microsoft/unilm>>. Accessed in: 24/08/2023.
- [13] VAERENBERGH, S. V. *Awesome Generative AI*. Available at: <<https://github.com/steven2358/awesome-generative-ai>>. Accessed in: 24/08/2023.
- [14] ZHOU, Y. *Awesome Text-to-Image*. Available at: <<https://github.com/Yutong-Zhou-cv/Awesome-Text-to-Image>>. Accessed in: 24/08/2023.
- [15] BANDI, A.; ADAPA, P. V. S. R.; KUCHI, Y. E. V. P. K. The power of generative ai: A review of requirements, models, input–output formats, evaluation metrics, and challenges. *Future Internet*, MDPI, v. 15, n. 8, p. 260, 2023.
- [16] GOZALO-BRIZUELA, R.; GARRIDO-MERCHÁN, E. C. A survey of generative ai applications. *arXiv preprint arXiv:2306.02781*, 2023.
- [17] GOOGLE. *Google Bard*. Available at: <<https://bard.google.com/>>. Accessed in: 23/08/2023.
- [18] HAT, R. *O que é API REST?* Available at: <<https://www.redhat.com/pt-br/topics/api/what-is-a-rest-api>>. Accessed in: 26/08/2023.
- [19] GROUP, P. G. D. *PostgreSQL: The World's Most Advanced Open Source Relational Database*. Available at: <<https://www.postgresql.org/>>. Accessed in: 26/08/2023.
- [20] FOUNDATION, O. *Node.js® is an open-source, cross-platform JavaScript runtime environment*. Available at: <<https://nodejs.org/en>>. Accessed in: 26/08/2023.
- [21] FOUNDATION, O. *Express Fast, unopinionated, minimalist web framework for Node.js*. Available at: <<https://expressjs.com/>>. Accessed in: 26/08/2023.
- [22] CONTRIBUTORS, S. *Sequelize*. Available at: <<https://sequelize.org/>>. Accessed in: 26/08/2023.
- [23] NASA. *Exoplanets Exploration - Discoveries Dashboard*. Available at: <<https://exoplanets.nasa.gov/discovery/discoveries-dashboard/>>. Accessed in: 13/09/2023.

- [24] SOURCE, M. O. *React*. Available at: [<https://react.dev/>](https://react.dev/). Accessed in: 26/08/2023.
- [25] MRDOOB. *Three.js*. Available at: [<https://threejs.org/>](https://threejs.org/). Accessed in: 26/08/2023.
- [26] POIMANDRES. *React-Three Fiber*. Available at: <https://github.com/pmndrs/react-three-fiber>. Accessed in: 26/08/2023.
- [27] SIMON, B. *Learn WebGL with Three.js*. Available at: <https://threejs-journey.com/>. Accessed in: 26/08/2023.
- [28] ARTIUKH, Y. *Creating sun with noise and Three.js 2*. Available at: <https://www.youtube.com/live/3krH52AhPqk?si=ucqjSm-DCvUDiVdx>. Accessed in: 26/08/2023.
- [29] COLORDODGE. *Procedural Planet*. Available at: <https://github.com/colordodge/ProceduralPlanet/tree/master>. Accessed in: 26/08/2023.