



UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CAMPUS REITOR JOÃO DAVID FERREIRA LIMA  
PROGRAMA DE GRADUAÇÃO EM SISTEMAS DE INFORMAÇÃO

Gabriel Avila

**GERAÇÃO AUMENTADA POR RECUPERAÇÃO APLICADA À INTERPRETAÇÃO  
DE METÁFORAS HISTÓRICAS:  
UM PIPELINE COM LLMS E EVIDÊNCIAS TEXTUAIS**

Florianópolis, Santa Catarina – Brasil  
2026

Gabriel Avila

**GERAÇÃO AUMENTADA POR RECUPERAÇÃO APLICADA À INTERPRETAÇÃO  
DE METÁFORAS HISTÓRICAS:  
UM PIPELINE COM LLMS E EVIDÊNCIAS TEXTUAIS**

Trabalho de Conclusão de Curso submetido  
ao Programa de Graduação em Sistemas de  
Informação da Universidade Federal de Santa  
Catarina para a obtenção do Grau de Bacharel em  
Sistemas de Informação.

**Orientador(a):** Rodrigo Bragio Bonaldo, Dr.

Florianópolis, Santa Catarina – Brasil

2026

## RESUMO

A classificação automática de metáforas em contextos históricos representa um desafio relevante no campo do Processamento de Linguagem Natural (PLN), dada a natureza subjetiva, contextual e evolutiva das expressões metafóricas. Enquanto a maior parte dos trabalhos em PLN concentra-se na detecção de metáforas (classificação binária), este projeto parte de um pipeline onde metáforas já foram previamente identificadas, com o objetivo de classificá-las em categorias históricas específicas baseadas no referencial de Fernández Sebastián (2024). Ao fazê-lo, o projeto visa colaborar com a construção e expansão de um dataset diacrônico voltado à análise interpretativa de metáforas políticas e sociais em fontes textuais históricas. O trabalho é dividido em dois módulos principais. O primeiro, denominado *Metaphor Classifier*, realiza a classificação das metáforas detectadas e expande o corpus por meio da recuperação de ocorrências semanticamente similares em obras históricas vetorizadas. Já o segundo módulo, baseado no paradigma *Retrieval-Augmented Generation* (RAG), é responsável por gerar justificativas textuais embasadas para cada classificação atribuída, citando trechos das fontes históricas como evidência e articulando-os conceitualmente às categorias históricas definidas. A metodologia adota uma abordagem mista. A dimensão quantitativa será aplicada na avaliação de classificações automatizadas (via métricas como *precision*, *recall* e *F1-score*, quando pertinente), enquanto a dimensão qualitativa garantirá a coerência conceitual dos pares gerados (*metáfora*, *evidência*, *categoria*), com apoio de uma equipe interdisciplinar de historiadores. Como resultados esperados, prevê-se a entrega de um pipeline funcional e documentado capaz de gerar pelo menos 100 pares validados de metáforas com evidências e classificações históricas, além de justificativas explicativas geradas via LLMs. Este recurso poderá servir de base empírica para estudos historiográficos e fomentar novas hipóteses sobre o uso e transformação de metáforas no discurso político e social ao longo do tempo. O projeto contribui, assim, não apenas com a inovação metodológica na interseção entre história e PLN, mas também com a promoção de maior transparência e rastreabilidade na interpretação automatizada de linguagem metafórica.

**Palavras-chave:** Processamento-de-Linguagem-Natural. Geração-Aumentada-por-Recuperação. Grandes-Modelos-de-Linguagem. Metáforas. Conjuntos-de-Dados-Diacrônicos. Mudança-Semântica. Corpora-Temporais. Linguística-Histórica.

## ABSTRACT

The automatic classification of metaphors in historical contexts poses a significant challenge in the field of Natural Language Processing (NLP), due to the subjective, contextual, and evolving nature of metaphorical expressions. While most NLP work focuses on metaphor detection (binary classification), this project begins from a pipeline in which metaphors have already been identified, with the goal of classifying them into specific historical categories based on the framework proposed by Fernández Sebastián (2024). In doing so, the project aims to support the construction and expansion of a diachronic dataset focused on the interpretive analysis of political and social metaphors in historical textual sources. The work is divided into two main modules. The first, called the *Metaphor Classifier*, performs the classification of detected metaphors and expands the corpus by retrieving semantically similar occurrences in vectorized historical works. The second module, based on the *Retrieval-Augmented Generation* (RAG) paradigm, is responsible for generating textual justifications for each assigned classification, quoting excerpts from historical sources as evidence and conceptually linking them to the defined historical categories. The methodology follows a mixed approach. The quantitative dimension will be applied in the evaluation of automated classifications (using metrics such as *precision*, *recall*, and *F1-score*, where applicable), while the qualitative dimension will ensure the conceptual coherence of the generated triplets (*metaphor*, *evidence*, *category*), with support from an interdisciplinary team of historians. The project also adopts principles of modular and iterative development, allowing flexibility in adapting to data and tools delivered by other members of the research group. As expected outcomes, the project aims to deliver a functional and documented pipeline capable of producing at least 100 validated metaphor pairs with historical classifications and supporting justifications generated via LLMs. This resource may serve as an empirical foundation for historiographical studies and stimulate new hypotheses about the use and transformation of metaphors in political and social discourse over time. The project thus contributes not only to methodological innovation at the intersection of history and NLP, but also to promoting greater transparency and traceability in the automated interpretation of metaphorical language.

**Keywords:** Natural Language Processing. Retrieval-Augmented Generation. Large Language Models. Metaphors. Diachronic Datasets. Semantic Change. Temporal Corpora. Historical Linguistics.

## **LISTA DE ABREVIATURAS E SIGLAS**

PLN	Processamento de Linguagem Natural
RAG	Geração Aumentada por Recuperação
LLM	Modelo de Linguagem de Grande Escala
PoC	Prova de Conceito
K-NN	Algoritmo dos K Vizinhos Mais Próximos
TCC	Trabalho de Conclusão de Curso
UFSC	Universidade Federal de Santa Catarina
VLab	Laboratório Virtual da UFSC

## LISTA DE SÍMBOLOS

→	Relação ou mapeamento (ex.: metáfora → evidência)
≥	Maior ou igual que
*F1-score*	Média harmônica entre precisão e revocação
*precision*	Proporção de itens relevantes entre os recuperados
*recall*	Proporção de itens relevantes recuperados entre os disponíveis

## LISTA DE FIGURAS

Figura 1	–	Diagrama de Gantt	14
----------	---	-------------------	----

## LISTA DE TABELAS

Tabela 1	– Entradas externas (pré-requisitos) . . . . .	11
Tabela 3	– Contribuições diretas do autor . . . . .	11
Tabela 5	– Recursos humanos envolvidos . . . . .	15
Tabela 7	– Plano de comunicação . . . . .	16
Tabela 9	– Riscos potenciais do projeto . . . . .	16



## SUMÁRIO

	<b>Lista de Figuras . . . . .</b>	<b>6</b>
	<b>Lista de Tabelas . . . . .</b>	<b>7</b>
	<b>Sumário . . . . .</b>	<b>8</b>
<b>1</b>	<b>PROPOSTA . . . . .</b>	<b>9</b>
1.1	<b>INTRODUÇÃO . . . . .</b>	<b>9</b>
1.2	<b>OBJETIVOS . . . . .</b>	<b>9</b>
1.2.1	<b>Objetivo Geral . . . . .</b>	<b>9</b>
1.2.2	<b>Objetivos Específicos . . . . .</b>	<b>10</b>
1.2.3	<b>Escopo do Trabalho . . . . .</b>	<b>10</b>
1.2.3.1	1 · Entradas externas (pré-requisitos) . . . . .	10
1.2.3.2	2 · Contribuições diretas do autor . . . . .	10
1.2.3.3	3 · Fora do escopo . . . . .	10
1.2.3.4	4 · Premissas & restrições . . . . .	11
1.2.3.5	5 · Cronograma técnico resumido . . . . .	12
1.3	<b>MÉTODO DE PESQUISA . . . . .</b>	<b>13</b>
1.4	<b>CRONOGRAMA . . . . .</b>	<b>13</b>
1.5	<b>CUSTOS . . . . .</b>	<b>14</b>
1.6	<b>RECURSOS HUMANOS . . . . .</b>	<b>15</b>
1.7	<b>COMUNICAÇÃO . . . . .</b>	<b>15</b>
1.8	<b>RISCOS . . . . .</b>	<b>15</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>17</b>

# 1 PROPOSTA

## 1.1 INTRODUÇÃO

### Introdução

A classificação automática de metáforas constitui um desafio clássico no campo do processamento de linguagem natural (PLN), devido à natureza abstrata e dependente de contexto das expressões metafóricas. Diferentemente das abordagens tradicionais que buscam identificar a presença ou ausência de metáforas, este trabalho recebe metáforas já detectadas em um pipeline prévio e propõe classificá-las em categorias específicas baseadas no referencial histórico proposto por Fernández Sebastián (2024). Esta classificação, sustentada teoricamente pelo estudo qualitativo e erudito das fontes históricas, visa compreender mais profundamente os padrões e o desenvolvimento das metáforas ao longo do tempo. Essa tarefa ocorre no contexto da construção de um dataset diacrônico — um recurso ainda relativamente raro em PLN — cuja curadoria está sendo conduzida por outros membros da equipe. O presente trabalho contribui com sua expansão e validação por meio da classificação automatizada das metáforas e da recuperação de evidências históricas.

Além disso, o projeto contempla dois módulos complementares: o *Metaphor Classifier*, que classifica e expande o conjunto de metáforas por meio da recuperação de trechos semanticamente semelhantes nas fontes históricas; e o *RAG* (Retrieval-Augmented Generation), responsável por gerar justificativas textuais embasadas, citando passagens das obras consultadas. Este segundo módulo combina capacidades avançadas dos modelos grandes de linguagem (LLMs) com mecanismos de recuperação automatizada de contextos históricos e literários relevantes. Espera-se, assim, contribuir para uma compreensão mais fundamentada e transparente de metáforas em corpora históricos, oferecendo suporte empírico para hipóteses historiográficas. O *RAG* cumpre a função central do trabalho de prova historiográfico, atuando em sua função referencial (LEWIS *et al.*, 2020).

## 1.2 OBJETIVOS

### Objetivos

#### 1.2.1 Objetivo Geral

Desenvolver e integrar dois módulos: o *Metaphor Classifier*, responsável por atribuir categorias históricas a metáforas previamente detectadas e expandir o conjunto de exemplos por meio de recuperação semântica de ocorrências análogas; e o *RAG*, voltado à geração de justificativas textuais embasadas, citando e recuperando evidências históricas relevantes para cada classificação.

## 1.2.2 Objetivos Específicos

### 1. Construção de regras de categorização

Desenvolver regras interpretáveis com base em um documento de categorias e exemplos elaborado por historiadores, de modo a possibilitar a atribuição automática ou semi-automática de categorias históricas a metáforas detectadas.

### 2. Metaphor Classifier — Recuperação de evidências históricas

Dado um conjunto de metáforas detectadas, atribua categorias históricas e recupere metáforas similares em contextos históricos, compondo pares metáfora ↔ e-evidência textual.

### 3. RAG — Geração de justificativas explicativas

Desenvolver um segundo módulo capaz de gerar respostas textuais explicativas para cada metáfora classificada, citando a evidência textual recuperada e a justificando com base na categoria histórica atribuída.

### 4. Avaliação funcional do pipeline

Avaliando-os com métricas formais (como F1-score ou precisão, se aplicável) ou, se simbólicos, com checklist qualitativo.

### 5. Documentação e replicabilidade prática

Manter documentação clara e progressiva no repositório do projeto, incluindo instruções de uso, descrição das etapas e exemplos de entrada e saída para cada módulo.

### 6. Flexibilidade de escopo e adaptação contínua

Conduzir o desenvolvimento sob abordagem iterativa e adaptativa, permitindo ajustes de rota conforme os dados e insumos forem disponibilizados por terceiros, respeitando os marcos definidos no cronograma.

## 1.2.3 Escopo do Trabalho

### 1.2.3.1 1 · Entradas externas (pré-requisitos)

### 1.2.3.2 2 · Contribuições diretas do autor

### 1.2.3.3 3 · Fora do escopo

- Detecção inicial de metáforas (Classificação Binária)
- Dataset-A (pré-existente)
- OCR, interface web

Tabela 1 – Entradas externas (pré-requisitos)

Origem	Entregável	Data prevista	Formato	Observação
Eduardo	Embeddings vetoriais das obras históricas	31 jul 2025	Parquet / FAISS / Chroma	Recuperação semântica
Franciele	Documento de categorias (3 exemplos cada)	15 ago 2025	Markdown	Base p/ regras de categorização

Fonte: Elaboração própria.

Tabela 3 – Contribuições diretas do autor

Módulo	Descrição resumida	Critério de aceite
<b>Metaphor Classifier</b>	Classifica metáforas detectadas e recupera expressões análogas em obras históricas.	$\geq 100$ pares ( <i>metáfora, evidência, categoria</i> ) gerados; execução end-to-end comprovada
<b>QuantEval</b>	Avalia a qualidade e coerência dos pares por checklist ou métricas formais.	Relatório técnico em Markdown com amostra comentada ou validação formal
<b>RAG</b>	Gera justificativas textuais baseadas na evidência recuperada.	Executável em $\geq 10$ metáforas distintas com justificativas coerentes e citação da fonte
<b>Docs</b>	Documentação viva (README, exemplos, instruções) atualizada a cada entrega.	Cada etapa entregue contém documentação correspondente no repositório

Fonte: Elaboração própria.

- LLM fine-tuned usada no RAG (fornecida por terceiros)

#### 1.2.3.4 4 · Premissas & restrições

- VLab-UFSC disponível com GPU/CPU até **julho/2026**.
- Textos históricos em domínio público; nenhuma licença adicional é necessária.
- Volume de dados modesto (< 10 GB); notebook pessoal e Google Drive são suficientes como redundância.

- Se embeddings ou categorias atrasarem, o escopo será reduzido ou o autor assumirá tarefas extras.

#### 1.2.3.5 5 · Cronograma técnico resumido

- **mar–abr 2025**

- Leitura de fundamentos sobre RAG
- PoC inicial com prompt fixo
- Leitura dos capítulos 4–7 de Fernández Sebastián
- Estudo do estado da arte sobre PLN e metáforas
- Base estruturada inicial do *Metaphor Classifier*

- **jul–ago 2025**

- Recebimento dos embeddings (Eduardo)
- Recebimento do documento de categorias (Franciele)
- Implementação do *Metaphor Classifier*

- **set–out 2025**

- Expansão do Dataset-B com recuperação de metáforas similares
- Ajuste das regras de categorização com base nas novas evidências
- Avaliação funcional do *Metaphor Classifier* (QuantEval)
- Desenvolvimento do módulo *RAG* (geração de justificativas textuais)
- Redação do Projeto II

- **dez 2025 – fev 2026**

- Curadoria de logs e análise de falhas
- Organização de outputs e documentação técnica
- Planejamento da redação final do TCC

- **mar–jun 2026**

- Versão revisada(foco em precisão) dos Scripts
- Redação da versão completa do TCC final

- **jun 2026**

- Margem para retorno da banca antes da defesa
- Preparação da defesa pública (slides, ensaios)

- **jul 2026**

- Ajustes finais e formatação ABNT
- Depósito do TCC na BU/UFSC

### 1.3 MÉTODO DE PESQUISA

#### Método de Pesquisa

Adotar-se-á uma abordagem mista, combinando métodos quantitativos e qualitativos. A dimensão **quantitativa** será utilizada sempre que o processo de classificação empregar modelos supervisionados, com aplicação de métricas como *F1-score*, *precision* e cobertura. Caso o classificador opere de forma simbólica, a avaliação seguirá uma abordagem qualitativa baseada em checklist, com inspeção manual da correspondência entre metáfora, categoria atribuída e evidência textual recuperada.

A dimensão **qualitativa** estará presente tanto na definição das categorias quanto na validação da coerência histórico-conceitual dos pares gerados (*metáfora* ↔ *evidência*). Nessa etapa, as classificações e justificativas produzidas pelos módulos serão analisadas à luz do referencial de Fernández Sebastián, com apoio da equipe de historiadores do projeto.

Trata-se de um estudo descritivo-exploratório, cujo objetivo principal é mapear e caracterizar padrões metafóricos diacrônicos com base em fontes textuais históricas. Não se busca inicialmente estabelecer relações de causalidade, mas sim levantar regularidades e hipóteses interpretativas que possam ser testadas posteriormente em estudos historiográficos mais amplos.

A pesquisa é de natureza aplicada, voltada ao desenvolvimento de módulos computacionais que integrarão um pipeline mais amplo. Dentre esses módulos, destacam-se o *Metaphor Classifier*, responsável pela categorização e expansão do conjunto de metáforas, e o *RAG*, voltado à geração textual explicativa com embasamento rastreável.

Serão utilizadas técnicas modernas de PLN, como *Large Language Models* (LLMs), vetorização semântica e algoritmos de vizinhança (K-NN). As ferramentas específicas — como bibliotecas de vetorização, índices semânticos ou modelos de geração — serão escolhidas conforme os requisitos práticos da implementação.

A execução ocorrerá em ambiente computacional misto, utilizando notebook pessoal e infraestrutura da UFSC, como o VLab da UFSC, bem como o repositório colaborativo IAeHistoriaUFSC, no qual o progresso do projeto será mantido de forma incremental e documentada.

### 1.4 CRONOGRAMA

#### Cronograma

**Nota:** O cronograma abaixo parte do pressuposto de que certos insumos técnicos — como embeddings vetoriais das obras históricas e etapas de pré-processamento — estão sendo desenvolvidos paralelamente por integrantes da equipe técnica (especialmente Maiko Nunes, Giovanna Ramalho

e Eduardo Goulart). O autor assume a integração desses artefatos nos módulos sob sua responsabilidade, sem duplicar esforços já em curso.

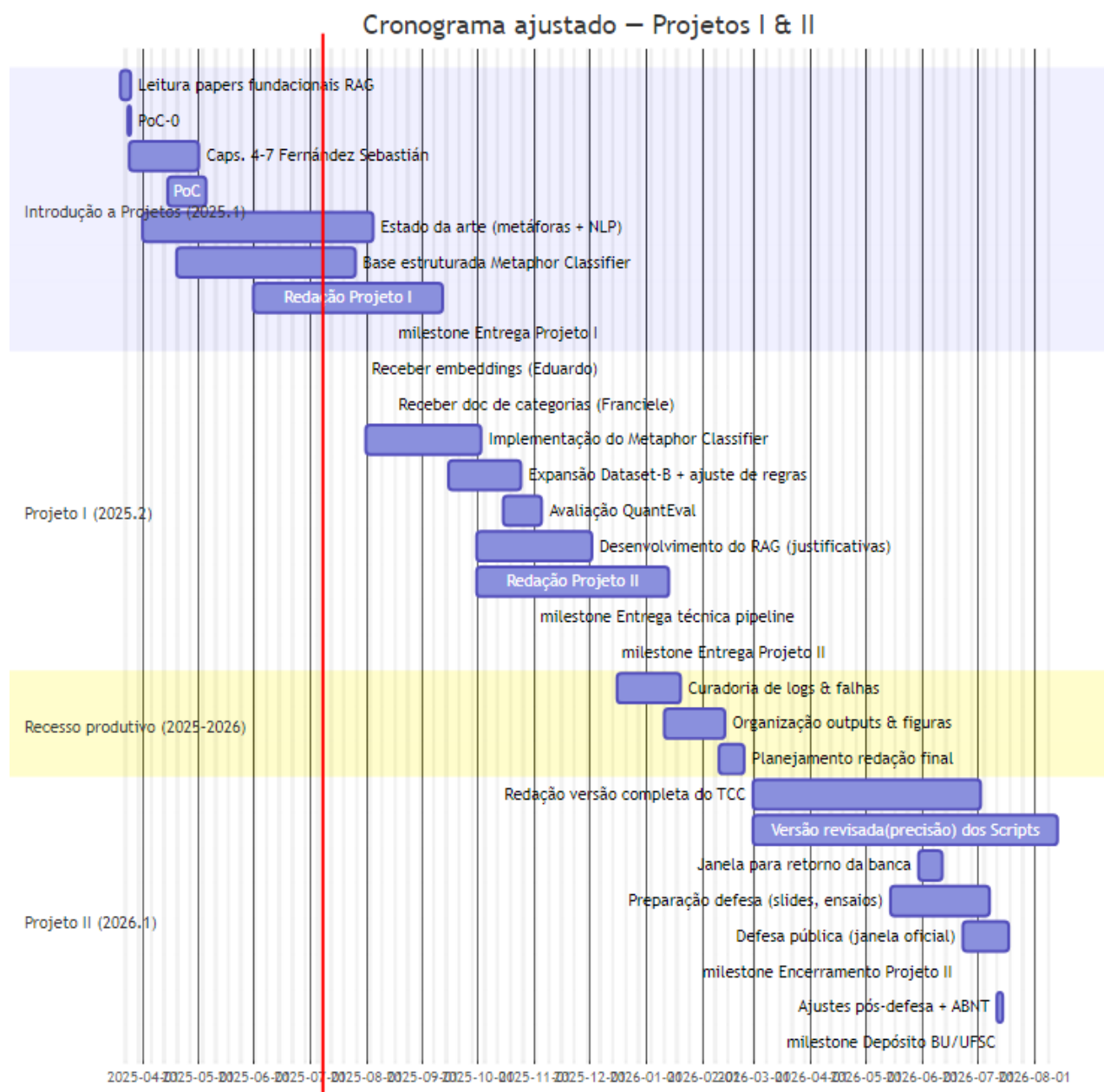


Figura 1 – Diagrama de Gantt

## 1.5 CUSTOS

### Custos

Não há custos previstos para a execução deste trabalho, uma vez que todos os recursos necessários — incluindo ferramentas de processamento de linguagem natural, repositórios digitais, infraestrutura computacional (como o VLab da UFSC) e canais

de comunicação — são gratuitos, institucionais ou já disponíveis para os membros da equipe.

## 1.6 RECURSOS HUMANOS

### Recursos Humanos

Tabela 5 – Recursos humanos envolvidos

Nome	Função
Rodrigo Bragio Bonaldo	Orientador
Jean Carlo Rossa Hauck	Coorientador/Responsável
Chiru Stefan	Senior Developer
Patricia Biral Varela	Pesquisadora em História
Maiko Ademir Nunes	Desenvolvedor Junior
Giovanna Ramalho	Engenharia de Software
Eduardo Peres Luckner Goulart	Desenvolvedor Junior
Franciele Dias da Silva	Pesquisadora em História com função organizacional
Mateus Freitas Borsatti	Pesquisador em História
Letícia Portella Milan	Pesquisadora em História
Éric Gabriel Kundlatsch	Pesquisador em História
Isabella Stersa de Oliveira	Pesquisadora em História
Gilson Mateus Pinto Júnior	Pesquisador em História
Leonardo Nogueira Aucar	Pesquisador em História
Gabriela Graudenz Muller	Linguística
Willian Meurer Welter	Pesquisador em História
Danillo Melo da Fonseca	Pesquisador em História
Alysson Julio Risso da Silva	Pesquisador em História
Marcel Vieira Silva	Pesquisador em História
Thamiris Fátima dos Santos	Desenvolvedor Front-End
Renata da Conceição Aquino da Silva	Pesquisadora em História
Fernanda Lyrio Heinzelmann	Psicologia
Breno Ampáro	Pesquisador em História
Gabriel Choucair Garcia	Pesquisador em História

Fonte: Elaboração própria.

## 1.7 COMUNICAÇÃO

### Comunicação

## 1.8 RISCOS

### Riscos



Tabela 7 – Plano de comunicação

O que será comunicado	Por quem	Para quem	Meio	Frequência
Progresso técnico (pipeline, classificador, RAG)	Autor	Equipe técnica	Reunião Google Meet	Ter 14 h - 16 h semanal
Discussões conceituais sobre categorias	Autor	Equipe de História	Reunião Google Meet	Ter 16h30 - 17h30 semanal
Dúvidas de código/ferramentas	Autor	Stefan (co-orientador técnico)	WhatsApp ou reunião	Sob demanda / semanal
Entregas parciais	Autor	Orientador(es)	GitHub / reunião	Ao final de cada etapa
Integrações entre módulos	Equipe técnica	Integrador designado	WhatsApp / Discord	Quando necessário
Avisos gerais	Autor	Todos	Grupo “IA e História” (WhatsApp)	Assíncrono contínuo

Fonte: Elaboração própria.

Tabela 9 – Riscos potenciais do projeto

Risco	Probabilidade	Impacto	Prioridade	Estratégia de resposta	Ações de prevenção
Integração complexa entre etapas do pipeline	Baixa	Alta	Média	Modularizar componentes	Revisões técnicas semanais com o Senior Pilotos
Ingestão do dataset diacrônico demanda maior esforço	Alta	Alta	Alta	Reduzir escopo ou redefinir metas	Incrementais

Fonte: Elaboração própria.

## REFERÊNCIAS

ARMASELU, Florentina *et al.* Multilingual Word Embedding and Linguistic Linked Open Data for Tracing Semantic Change. **Rasprave Instituta za hrvatski jezik i jezikoslovlje**, v. 50, n. 2, p. 219–244, 2024. DOI: 10.31724/rihjj.50.2.1.

BONALDO, Rodrigo Bragio. As palavras e os tokens: Projeção vetorial aplicada ao estudo da semântica dos tempos históricos. **Revista de Teoria da História**, v. 27, n. 1, p. 7–50, 2024. DOI: 10.5216/rth.v27i1.79370.

CHOI, Minjin *et al.* MeLBERT: Metaphor Detection via Contextualized Late Interaction Using Metaphorical Identification Theories. *In: PROCEEDINGS of the Annual Meeting of the Association for Computational Linguistics*. [S.l.: s.n.], 2021. arXiv: 2104.13615. Disponível em: <<https://arxiv.org/abs/2104.13615>>.

DANKIN, Lena; BAR, Kfir; DERSHOWITZ, Nachum. Can Yes-No Question-Answering Models be Useful for Few-Shot Metaphor Detection? *In: PROCEEDINGS of the 3rd Workshop on Figurative Language Processing (FLP)*. Abu Dhabi, United Arab Emirates: Association for Computational Linguistics, 2022. P. 125–130. DOI: 10.18653/v1/2022.flp-1.17.

FERNÁNDEZ-SEBASTIÁN, Javier. **Key Metaphors for History: Mirrors of Time**. Abingdon, UK: Taylor & Francis, 2024. ISBN 978-1-138-35446-3.

LEWIS, Patrick *et al.* Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. **arXiv preprint**, 2020. arXiv: 2005.11401 [cs.CL]. Disponível em: <<https://arxiv.org/abs/2005.11401>>. Citado na p. 9.

LIN, Yujie *et al.* A Dual-Perspective Metaphor Detection Framework Using Large Language Models. **arXiv preprint**, 2024. arXiv: 2412.17332 [cs.CL]. Disponível em: <<https://arxiv.org/abs/2412.17332>>.

MONTANELLI, Marco; PERITI, Sara. A Survey on Contextualised Semantic Shift Detection. **arXiv preprint**, 2023. arXiv: 2304.01666 [cs.CL]. Disponível em: <<https://arxiv.org/abs/2304.01666>>.

PRYSTAWSKI, Ben *et al.* Psychologically-Informed Chain-of-Thought Prompts for Metaphor Understanding in Large Language Models. **arXiv preprint**, 2023. arXiv: 2209.08141 [cs.CL]. Disponível em: <<https://arxiv.org/abs/2209.08141>>.

VASWANI, Ashish *et al.* Attention Is All You Need. **arXiv preprint**, 2017. arXiv: 1706.03762 [cs.CL]. Disponível em: <<https://arxiv.org/abs/1706.03762>>.

XIONG, Ruibin *et al.* On Layer Normalization in the Transformer Architecture. **arXiv preprint**, 2020. arXiv: 2002.04745 [cs.LG]. Disponível em: <<https://arxiv.org/abs/2002.04745>>.