



## UNIVERSIDADE DO ESTADO DO PARÁ – UEPA

### CENTRO DE CIÊNCIAS NATURAIS E TECNOLOGIA – CCNT

### CURSO DE BACHARELADO EM ENGENHARIA DE SOFTWARE

*Davi Costa Mendes Da Silva*

*Ryan Souza Santana*

*Prof. Dr. Xxx (Orientador)*

#### **Projeto de Pesquisa:**

Análise Comparativa de LLMs na Elaboração de Resumos de Editais

Públicos: Uma Abordagem com Estratégias de Engenharia de Prompt.

Elaborado em: 27/05/2025

Área de conhecimento: Sistemas de informação

## 1 CONTEXTUALIZAÇÃO

Diante do aumento no uso de modelos de Inteligência Artificial (IA) generativa para tarefas acadêmicas, *chatbots* vêm ganhando espaço no público geral. Por meio do Processamento de Linguagem Natural (*Natural Language Processing* – NLP) e do Aprendizado de Máquina (*Machine Learning* – ML), esses algoritmos são capazes de compreender e gerar respostas contextualizadas ao usuário, promovendo maior eficiência operacional e melhorando sua experiência (Nze, S., 2024). A partir disso, modelos disponíveis no mercado possuem, entre outras funções, o papel de abstrair informações complicadas e ambíguas com o intuito de auxiliar o usuário a um fácil entendimento a partir de uma conversa, similar a de redes sociais.

No entanto, a qualidade e confiabilidade das respostas podem variar significativamente entre diferentes modelos de acordo com o *prompt* empregado. Como comentado por Yampolskiy (2016) em seu artigo *Taxonomy of Pathways to Dangerous Artificial Intelligence* que apresenta a possibilidade de erro em IA no geral. Assim, o intuito deste documento é provar e classificar por meio de critérios de avaliação quais

Modelos de Linguagem Grande (*Large Language Model* - LLM) possuem melhor desempenho e menor quantidade de falhas.

## 2 PROBLEMA DE PESQUISA

- **Pergunta norteadora:** Qual LLM disponível ao público é mais confiável e melhor performa no resumo de editais aplicando diferentes técnicas de engenharia de *prompt* considerando critérios de coerência e integridade da informação?

## 3 HIPÓTESE

- Considerando o avanço das LLMs na produção de texto, pressupõe-se que a ocorrência de variações significativas e erros na geração de conteúdo pelos modelos testados seja baixa. As principais discrepâncias nos resultados tendem a estar associadas às diferenças nos *prompts* utilizados.

## 4 JUSTIFICATIVA

A sumarização automática tem se consolidado como ferramenta essencial para lidar com volumes crescentes de texto, pois permite extrair as informações principais de forma rápida e compreensível (Martins, 2001). Nesse contexto, os LLMs demonstram capacidade notável não apenas para gerar resumos extrativos e abstrativos, mas também para capturar nuances semânticas e gerar textos coesos, o que supera em diversos casos os métodos baseados em regras ou em estatísticas simples (Zhang; Liu; Zhang, 2023). Essa evolução técnica tem viabilizado aplicações em domínios distintos como notícias, relatórios médicos e documentos complexos, apontando para um novo patamar de eficiência e qualidade na geração de resumos automáticos.

Devido essa evolução nos últimos anos, as LLMs vêm sendo amplamente utilizadas em ambientes de pesquisa e ensino, seja para auxiliar na revisão bibliográfica, na redação de artigos ou mesmo no desenvolvimento de pesquisas em disciplinas de ciências exatas e humanas. Dessa forma, estudos comparativos como (Moon, 2025) evidenciam a utilidade dessas plataformas para resolver questões complexas. Esse cenário reforça o protagonismo das LLMs na academia e motiva investigações mais

sistemáticas sobre sua performance em diferentes tarefas de linguagem.

Apesar do entusiasmo, pouco se dedicou a avaliar criticamente a coerência lógica e a integridade factual dos resumos gerados por LLMs em cenários de uso real, isto é, não controlados por métricas acadêmicas ou conjuntos de problemas idealizados. Por exemplo, o estudo (Fabbri *et al.*, 2021) possui um inegável valor acadêmico e serve como uma base para futuros estudos da área, porém foca em materiais padronizados, deixando de lado textos complexos e heterogêneos que a sociedade demanda no dia a dia. Falta, portanto, uma investigação que quantifique e compare erros de omissão, distorção ou introdução de informações imprecisas quando as LLMs são utilizadas por usuários finais em atividades práticas.

Os editais públicos são instrumentos centrais na comunicação entre órgãos públicos e a sociedade, especialmente no que diz respeito à oferta de oportunidades acadêmicas, científicas e sociais. Eles orientam desde processos seletivos até a concessão de bolsas e financiamentos, sendo, portanto, fundamentais para garantir o acesso a oportunidades voltadas à educação, à pesquisa e à inovação. Contudo, sua linguagem técnica e estrutura complexa dificultam a compreensão por parte de grande parte da população, limitando o alcance dessas oportunidades. Nesse contexto, utilizar LLMs para gerar resumos acessíveis desses documentos pode contribuir significativamente para ampliar o entendimento e promover inclusão. Além disso, sua estrutura formal, repleta de datas, requisitos e critérios, torna-os um caso de uso ideal para testar a consistência dos resumos gerados. Assim, os editais não apenas representam um desafio técnico interessante para avaliação de LLMs, como também constituem uma aplicação com potencial direto de impacto social e acadêmico.

Diante desse panorama, esta pesquisa vem preencher a lacuna existente ao comparar diretamente, em um ambiente de uso prático, quatro LLMs consagradas e disponíveis gratuitamente (ChatGPT, Gemini, DeepSeek e Copilot) na tarefa de sumarização abstrativa de editais públicos. Será avaliado não apenas o grau de coerência e integridade informacional de cada modelo, mas também o impacto das variações dos *prompts* utilizados nas saídas geradas. Ao adotar uma metodologia inspirada em trabalhos comparativos anteriores (Tabosa *et al.*, 2020) e em estudos de avaliação de sumarização como (Fabbri *et al.*, 2021), este projeto almeja identificar o modelo/*prompt* de melhor desempenho e oferecer diretrizes práticas para a comunidade acadêmica e gestora de políticas públicas.

## 5 OBJETIVOS

O objetivo dessa pesquisa é determinar, dentre 4 IAs generativas, ChatGPT, Gemini, DeepSeek e Copilot, qual performa melhor no resumo abstrativo de editais. Para tanto, os seguintes objetivos específicos deverão ser atingidos:

- Definição das melhores técnicas para construção dos *prompts* para uso na criação dos resumos;
- Avaliação dos resultados da interação cada um dos modelos com cada *prompt* construído e seleção os melhores resumos;
- Comparaçāo dos melhores resumos e seleção da melhor opção de modelo e modelo/*prompt* para sumarização.

## 6 REFERENCIAL TEÓRICO

Para embasar esta pesquisa, foram estudados casos relacionados na literatura existente, focando em sumarização automática, comparação de LLMs, estudos sobre técnicas de *Prompt* e avaliação de resumos abstratos para a definição dos métodos e critérios que tomaremos de base ao longo da pesquisa.

### 6.1 Sumarização automática: conceito geral e diferenciação de técnicas

A sumarização automática é o processo computacional de condensar textos mantendo suas informações mais relevantes pela identificação do conteúdo essencial a partir de um conjunto maior de dados (Martins *et al.*, 2001). As abordagens de sumarização dividem-se principalmente entre extrativas e abstrativas. A sumarização extrativa seleciona diretamente frases do texto original, enquanto a abstrativa consiste em compreender o texto original e recontá-lo em menos palavras. Embora os métodos extrativos sejam mais simples e seguros em termos de fidelidade factual, os abstrativos, por imitarem o comportamento humano, são preferidos em contextos que exigem legibilidade e fluência, como documentos mais complexos .

### 6.2 Large Language Models (LLMs) e sua aplicāo em sumarizaçāo

As LLMs representam um marco no processamento de linguagem natural ao demonstrar desempenho robusto em tarefas como tradução, geração de texto e,

especialmente, sumarização. Segundo (Zhang; Liu; Zhang, 2023), essas redes profundas pré-treinadas são capazes de produzir resumos comparáveis aos gerados por humanos, inclusive superando métodos supervisionados em avaliações qualitativas. Embora originalmente treinadas para diversas finalidades, essas LLMs têm demonstrado versatilidade ao lidar com resumos tanto extrativos quanto abstrativos, tornando-se uma alternativa viável e acessível para usuários não especialistas. No entanto, diferentes modelos apresentam níveis variados de fidelidade e coerência nos resumos gerados, especialmente quando se varia a estrutura do *prompt* ou o domínio textual. Inspirados por metodologias comparativas como as aplicadas em (Moon, 2025) , este projeto seguirá abordagem semelhante, mas com foco em resumos de editais públicos e uso prático direto das plataformas LLM, sem API ou pré-processamento adicional, o que simula um cenário realista de usuário final.

### **6.3 Engenharia de *prompt*: conceito e estratégias para garantir a integridade da informação**

A engenharia de *prompt* consiste na formulação estratégica de instruções fornecidas às LLMs para orientar suas respostas. Conforme (Son; Won; Lee, 2025), técnicas como *Chain-of-Thought*, onde o modelo é instruído a percorrer um raciocínio passo a passo, antes de dar a resposta final, ou que apliquem comandos “*step by step*” são capazes de melhorar substancialmente a precisão e a consistência das saídas geradas, especialmente quando se deseja factualidade. Se observou em (Zhang; Liu; Zhang, 2023) o uso de *prompts* explicativos com exemplos de estruturas e instruções explícitas para atingir maior fidelidade informacional, além disso, uma técnica de um resumo misto (Extrativo-Abstrativo), instruindo a fabricação de uma sumarização extrativo e em seguida uma abstrativa baseada no resumo anterior. Esta pesquisa pretende aplicar as técnicas comentadas mantendo os *prompts* simplificados, juntamente com um exemplo sem uso explícito de técnicas, simulando o comportamento de usuários comuns que não têm domínio da construção de *prompts*, o que permite avaliar a robustez e desempenho real dos modelos.

#### **6.4 Avaliação de resumos abstrativos: limites dos métodos automáticos e a escolha por avaliação humana**

A avaliação de resumos abstrativos representa um desafio metodológico significativo. Métricas automáticas como ROUGE, embora amplamente adotadas, apresentam baixa correlação com julgamentos humanos em tarefas que exigem interpretação semântica e verificação factual (Kryscinski *et al.*, 2019). A situação se agrava quando se utiliza LLMs como avaliadores, pois estudos como (Shen *et al.*, 2023) demonstram que esses modelos, embora úteis em avaliações superficiais, mostram inconsistência ao julgar critérios como coerência, fluência e factualidade em resumos de alta qualidade, além de serem sensíveis a variações mínimas. Em virtude dessas limitações, este trabalho utilizará avaliadores humanos e escalas do tipo Likert para mensurar critérios de avaliação baseados nos critérios empregados em (Fabbri *et al.*, 2021) (Tabosa *et al.*, 2020):

- **Coerência:** A qualidade coletiva de todas as frases. O resumo deve ser bem estruturado e bem organizado. O resumo não deve ser apenas um amontoado de informações relacionadas, mas deve construir, frase por frase, um corpo coerente de informações sobre um tópico.
- **Consistência:** O alinhamento factual entre o resumo e a fonte resumida. Um resumo factualmente consistente contém apenas declarações que são implícitas no documento de origem. Esse critério avalia se o resumo apresenta alucinações dos modelos.
- **Fluência:** A qualidade de frases individuais. Avalia a existência de erros gramaticais ou estruturais, as frases no resumo não devem apresentar problemas de formatação, erros de capitalização ou frases claramente agramaticais que dificultem a leitura do texto.
- **Relevância:** Seleção de conteúdo importante da fonte. O resumo deve incluir apenas informações importantes do documento de origem e não omitir informações cruciais. Esse critério avalia se houve devaneios.

## 7 METODOLOGIA

### 7.1 Coleta dos editais que serão usados

Serão escolhidos 5 a 10 editais representativos de diferentes áreas (concursos, licitações, bolsas, ingressos), abrangendo o máximo de variações que compartilhem a estrutura de um edital, armazenados e catalogados.

### 7.2 Construção dos *prompts* que serão utilizados pelos modelos

Para a construção dos *prompts* definimos que todos iriam instruir que o resumo seja estruturado em tópicos definidos (por exemplo título, resumo geral, cargos e vagas, requisitos, critérios, datas, taxas e valores), também planejamos 3 abordagens (exemplos de *prompts* preliminares no apêndice A):

- **Simplificada:** simulando como um usuário comum utilizaria os LLMs, esse *prompt* tem o intuito de averiguar se a falta de técnicas de Engenharia de *Prompt* realmente impacta na integridade dos fatos e desempenho geral dos resumos.
- **Chain-of-Thought:** Descrevendo como o processo de resumo em etapas menores e mais detalhadas, orientamos o modelo a raciocinar por partes, essa técnica permite que ele foque em uma tarefa por vez e verifique antes de produzir o resumo. Reduz omissões e melhora a organização lógica do conteúdo.
- **Extrativa-Abstrativa:** Mantendo o conceito de etapas do anterior, porém adicionando instruções requerendo a sumarização extrativa do edital seguida de um abstrativa do resumo gerado pela etapa anterior. Essa abordagem tem o intuito de formar um resumo mais contido e íntegro.

### 7.3 Geração dos resumos

Com os *prompts* e a base de editais pronta, se inicia um processo manual de criação dos resumos através dos modelos selecionados. Essa abordagem tem o intuito de replicar o processo feito por um usuário comum que utiliza essas ferramentas no dia a dia, o que seria impossível de simular se utilizarmos APIs. Todos os resumos serão efetuados pelos sites oficiais de cada uma das empresas responsáveis pelos modelos analisados, seguindo os mesmo parâmetros de versão

básica (GPT-4o, Gemini 2.0 flash, DeepSeek-V3 e o copilot não disponibiliza sua versão) e sem funções adicionais como pesquisa ou pensamento profundo.

O processo consiste na iniciação de uma conversa, importação de um edital e aplicação do *prompt*, o resumo gerado será transscrito para um documento de texto, identificado e armazenado em um *drive* (resultados que não seguirem os tópicos definidos nos *prompts* serão descartados, necessitando recomeçar o processo), em seguida adicionado a uma matriz de controle destacando o modelo e *prompt* que geraram o resumo. Esse processo se repetirá para cada edital selecionado, com cada um dos 3 *prompts* desenvolvidos, em cada um dos 4 modelos estudados, serão gerados de 60 a 120 resumos, dependendo do número de editais selecionando para a pesquisa. Com todos os resumos feitos, identificados e armazenados, inicia o processo mais demorado.

#### **7.4 Avaliação dos resumos**

Um formulário baseado no “*The data collection interface*” de (Fabbri *et al.*, 2021) será criado para cada edital no Google Forms, cada formulário irá conter instruções para os avaliadores, destacando o significado de cada um dos critérios e disponibilizando o link para o Documento original do Edital analisado. O formulário possuirá uma seção para cada resumo do edital do formulário, disponibilizando um link para o resumo analisado e novamente para o edital, logo abaixo um campo de avaliação para cada uma das dimensões de avaliação (exemplo do formulário no Apêndice B).

Dependendo da quantidade de editais selecionados e devido limitação do Google Forms de só contabilizar as respostas se todas as seções forem respondidas, pode se fazer necessário diminuir a quantidade de resumos por formulário, acarretando na criação de múltiplas versões de um formulário (Formulário Edital A - 1, Formulário Edital A - 2) que conteriam a base de resumos separada igualmente, cuja a disposição será guardada. Essa medida diminuiria a quantidade de resumos analisados por vez, evitando que o processo seja maçante para os avaliadores, visto que todos os resumos serão avaliados por cada um dos avaliadores.

Como avaliadores, serão utilizadas três voluntárias, acadêmicas de licenciatura, que irão receber o conjunto dos formulários, sem saber de qual modelo ou *prompt*

os resumos se tratam. Ao fim, espera-se ter uma base de dados contendo 3 avaliações para cada um dos resumos.

### **7.5 Classificação das avaliações**

Após todo o processo de avaliação, através da função nativa do Google Forms, iremos transferir as respostas para uma planilha no Google Planilhas, em seguida, ajustar a estrutura. Caso os formulários possuam mais de uma versão, temos que mesclar as tabelas. Em seguida será calculada a média das 3 avaliações de cada resumo.

Adiante, teremos que organizar os resumos, identificando cada um através da disposição deles em cada Formulário e separando eles por modelo e *prompt* através da sua identificação na matriz de controle. Ao fim desse processo, espera-se uma tabela para cada edital, que destaque as notas de cada resumo e os separe por *prompt* e por modelo.

### **7.6 Comparação e análise das melhores avaliações**

Nesta etapa, realizaremos o ranqueamento dos modelos, dos *prompts* e de suas combinações. Para uma análise mais representativa do desempenho geral das objetos avaliados, calcularemos a média das notas atribuídas a cada conjunto modelo/*prompt*, a cada modelo isoladamente e a cada *prompt*. Com esses dados, poderemos identificar qual combinação apresentou o melhor desempenho, qual modelo obteve melhores resultados independentemente do *prompt* utilizado e, da mesma forma, qual *prompt* se destacou independentemente do modelo. Essa metodologia visa evitar conclusões precipitadas baseadas em casos isolados. Por exemplo, uma determinada combinação pode ter obtido a melhor avaliação em um edital específico, mas apresentado desempenho inferior nos demais.

## **8 RESULTADOS ESPERADOS**

Após o processo realizado pela metodologia apresentada, definimos três resultados que podem servir de marco da pesquisa realizada:

- Espera-se construir *prompts* eficientes tanto para um usuário leigo quanto para o profissional conhecedor de engenharia de *prompt* que atendam a demanda exigida.

- Espera-se obter uma base de resumos avaliados que cumpra seu papel nessa pesquisa e que sirva para futuros trabalhos.
- Com comparação dos resultados obtidos, espera-se ser extraído um ranqueamento dos modelos, *prompts* e combinações e comentários sobre os casos observados.

## 9 IMPACTO OU APLICAÇÃO

A contribuição deste documento apoia-se principalmente na apresentação de uma utilização dos LLMs, de uma forma frequentemente utilizada nos dias de hoje. Apresentando um ponto de interesse entre a possível experiência do usuário e padrões estabelecidos pelas respostas dos modelos disponíveis no mercado.

## 10 CRONOGRAMA DE EXECUÇÃO

<b>Etapas</b>	<b>MESES</b>					
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
Coleta dos editais	X					
Construção dos <i>prompts</i>	X	X◊				
Geração dos resumos		X	X			
Avaliação dos resumos			X	X	X◊	
Classificação dos editais avaliados					X	
Comparação e análise						X
Classificação dos LLMs avaliados						X◊

## 11 ORÇAMENTO

Não haverá necessidade de investimentos nesta pesquisa, visto que trata-se de um experimento que utiliza versões gratuitas e acessíveis ao usuário geral.

## 12 USO DE INTELIGÊNCIA ARTIFICIAL GENERATIVA

Este trabalho segue as orientações da Resolução n. 2/2024-COBES, fazendo uso de Inteligência Artificial Generativa (IAG) da seguinte forma:

- Ferramenta Chat GPT para levantamento de referências e correção textual.
- Ferramenta SciSpace para levantamento de referências.

## REFERÊNCIAS

Fabbri, A. R. et al. SummEval: Re-evaluating Summarization Evaluation. **Transactions of the Association for Computational Linguistics**, v. 9, p. 391–409, 26 abr. 2021.

Disponível em:

[https://direct.mit.edu/tacl/article/doi/10.1162/tacl\\_a\\_00373/100686/SummEval-Re-evaluating-Summarization-Evaluation](https://direct.mit.edu/tacl/article/doi/10.1162/tacl_a_00373/100686/SummEval-Re-evaluating-Summarization-Evaluation) Acesso em: 19 de maio de 2025.

Kryscinski, W. et al. **Neural Text Summarization**: A Critical Evaluation. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). Anais...Stroudsburg, PA, USA: Association for Computational Linguistics, 2019.

Disponível em: <https://www.aclweb.org/anthology/D19-1051> Acesso em: 19 de maio de 2025.

Martins, C. B. et al. **Introdução à sumarização automática**. Relatório Técnico RT-DC, v. 2, 2001. <https://sites.icmc.usp.br/taspardo/rtdc00201-cmartinetal.pdf> Acesso em: 29 de abril de 2025.

Moon, I. H. **Benchmarking Large Language Models for Calculus Problem-Solving**: A Comparative Analysis. 30 mar. 2025. Disponível em:

<<https://arxiv.org/abs/2504.13187>> Acesso em: 19 de maio de 2025.

Nze, S. U. AI-Powered Chatbots. **Global Journal of Human Resource Management**, v. 12, n. 1, p. 10-25, 2024. Disponível em:

<https://ejournals.org/gjhrm/vol12-issue-6-2024/ai-powered-chatbots/> Acesso em: 13 de maio de 2025.

Shen, C. et al. **Large Language Models are Not Yet Human-Level Evaluators for Abstractive Summarization**. Findings of the Association for Computational Linguistics: EMNLP 2023. Anais...Stroudsburg, PA, USA: Association for Computational Linguistics, 2023. Disponível em:

<https://aclanthology.org/2023.findings-emnlp.278> Acesso em: 19 de maio de 2025.

Son, M.; Won, Y. J.; Lee, S. **Optimizing Large Language Models**: A Deep Dive into Effective Prompt Engineering Techniques. *Applied Sciences* (Switzerland), v. 15, n. 3, 1

fev. 2025. Disponível em: <https://www.mdpi.com/2076-3417/15/3/1430> Acesso em: 19 maio 2025.

Tabosa, H. R. et al. **Avaliação do desempenho de um software de sumarização automática de textos**. Informação & Informação, v. 25, n. 1, p. 189, 1 abr. 2020. <http://www.uel.br/revistas/uel/index.php/informacao/article/view/35928> Acesso em: 19 maio 2025.

Yampolskiy, Roman V. Taxonomy of Pathways to Dangerous Artificial Intelligence. In: **AAAI Workshop: AI, Ethics, and Society**. 2016. p. 143-148. Disponível em: <https://cdn.aaai.org/ocs/ws/ws0156/12566-57418-1-PB.pdf> Acesso em: 13 maio 2025.

Zhang, H.; Liu, X.; Zhang, J. **Extractive Summarization via ChatGPT for Faithful Summary Generation**. Findings of the Association for Computational Linguistics: EMNLP 2023. Anais... Stroudsburg, PA, USA: Association for Computational Linguistics, 2023. Disponível em: <https://aclanthology.org/2023.findings-emnlp.214> Acesso em: 19 maio 2025.

## APÊNDICE A - EXEMPLO DE PROMPTS.

Esses são exemplos preliminares que sofrerão alterações ao decorrer da melhor definição dos mesmos como destacado no item 2 da metodologia.

<b>Abordagem</b>	<b>Prompt</b>
<b>Simplificada</b>	<p>Envie um edital em PDF, Faça para mim um resumo separado nos seguintes tópicos:</p> <ul style="list-style-type: none"> <li>- Título do edital</li> <li>- Resumo geral</li> <li>- Cargos e vagas</li> <li>- Requisitos</li> <li>- Critérios de seleção</li> <li>- Datas importantes</li> <li>- Taxas valores envolvidos</li> </ul> <p>Me diga as informações com linguagem clara e objetiva.</p>
<b>Chain-of-Thought</b>	<p>Você receberá um edital de concurso público. Para garantir que todas as informações importantes sejam extraídas corretamente, siga o processo abaixo:</p> <ol style="list-style-type: none"> <li>1. Leia o edital atentamente e verifique se ele contém cada um dos tópicos a seguir.</li> <li>2. Para cada tópico, liste as informações relevantes separadamente.</li> <li>3. Ao final, organize o resumo final com linguagem clara e objetiva,</li> </ol>

	<p>separando em tópicos conforme abaixo:</p> <ul style="list-style-type: none"> <li>- Título do edital</li> <li>- Resumo geral</li> <li>- Cargos e vagas disponíveis</li> <li>- Requisitos para candidatura</li> <li>- Critérios de seleção</li> <li>- Datas importantes</li> <li>- Taxas de inscrição e valores envolvidos</li> </ul> <p>Importante: só resuma após identificar claramente as informações para cada tópico. Se algum tópico estiver ausente, sinalize com "informação não encontrada".</p>
<b>Extrativa-Abstrativa</b>	<p>Você irá processar um edital público enviado em PDF. Para gerar um resumo estruturado e confiável, siga estas duas etapas:</p> <p><b>**Etapa 1 – Extração (Sumarização Extrativa):**</b></p> <p>Leia o edital e localize, para cada um dos tópicos abaixo, os trechos exatos onde a informação aparece. Copie esses trechos diretamente do edital (mantendo a fidelidade).</p> <p>Tópicos:</p> <ul style="list-style-type: none"> <li>- Título do edital</li> </ul>

	<ul style="list-style-type: none"><li>- Resumo geral</li><li>- Cargos e vagas disponíveis</li><li>- Requisitos para candidatura</li><li>- Critérios de seleção</li><li>- Datas importantes</li><li>- Taxas de inscrição e valores envolvidos</li></ul> <p><b>**Etapa 2 – Reescrita (Sumarização Abstrativa):**</b></p> <p>Com base nas informações extraídas na Etapa 1, reescreva com suas próprias palavras um resumo mais claro e objetivo. O resumo final deve conter os mesmos tópicos, mas com linguagem resumida e compreensível para o público geral.</p> <p>Se não encontrar alguma informação, indique com “não localizado no edital”.</p>
--	--

## APÊNDICE B - EXEMPLO DO FORMULÁRIO DE AVALIAÇÃO.

Esse é um exemplo preliminar do formulário de avaliação que será desenvolvido, como destacado no item 4 da metodologia.

### Tela inicial:

**Avaliação dos resumos do Edital ##**

**Instruções**

- Leia cuidadosamente o Edital, esteja atento às informações que ele contém.
- Leia os resumos propostos.
- Avalie cada resumo em uma escala de 1 (pior) a 5 (melhor) com base em sua relevância, consistência, fluência e coerência.
- #Explicar caso de invalidez#

**Definições**

**Relevância:** A avaliação mede o quanto bem o resumo captura os pontos chave do artigo. Considere se todos os aspectos importantes (e somente eles) estão contidos no resumo.

**Consistência:** A avaliação mede se os fatos no resumo são consistentes com os fatos do artigo original. Considere se o resumo reproduz todos os fatos com precisão e não inventa informações falsas.

**Fluência:** Esta avaliação mede a qualidade das sentenças individualmente, se estão bem escritas e gramaticalmente corretas. Considere a qualidade das sentenças individualmente.

**Coerência:** Esta avaliação mede a qualidade de todas as sentenças coletivamente, se se encaixam e soam naturalmente. Considere a qualidade do resumo como um todo.

**Edital analisado:** ##Título do Edital## [@Link do Edital](#)

[Próxima](#) [Limpar formulário](#)

### Exemplo de seção:

**Avaliação dos resumos do Edital ##**

**Resumo ##**

**Edital analisado:** ##Título do Edital## [@Link do Edital](#)  
**Resumo avaliado:** ##Identificação do Resumo## [@Link do Resumo](#)

	1	2	3	4	5
Relevância	<input type="radio"/>				
Consistência	<input type="radio"/>				
Fluência	<input type="radio"/>				
Coerência	<input type="radio"/>				

[Voltar](#) [Próxima](#) [Limpar formulário](#)