

Usando LLMs para simplificar e representar documentos médicos

Leandro Krug Wives¹, Maria José Bocorny Finatto²

¹ Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brasil

²Instituto de Letras – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.002 – 91.501-970 – Porto Alegre – RS – Brasil

Abstract. *Although the field of medicine produces a substantial number of academic papers, much of this content is inaccessible to a broad audience due to barriers in comprehension. A limited understanding can undermine the ability to make informed health decisions and adopt health-promoting practices. In this context, text simplification emerges as a valuable approach to promote knowledge dissemination in health care. In this sense, this study reports on experiments that evaluated the application of large language models (LLMs) to simplify medical documents in Portuguese, generating simpler and more accessible versions, in addition to graphical representations of their content. The results demonstrate the feasibility of this approach to provide equity in information access.*

Resumo. *Embora a medicina se destaque pela alta produtividade de artigos acadêmicos, a maior parte das pessoas enfrenta barreiras para compreender o conteúdo dessas publicações. A compreensão inadequada prejudica o acesso a informações cruciais e a adoção de práticas benéficas à saúde e ao bem-estar. Diante desse cenário, a simplificação de textos emerge como ferramenta de grande relevância, ajudando a promover a disseminação eficaz do conhecimento em saúde. Nesse contexto, este artigo descreve experimentos que avaliam o uso de LLM para a simplificação de documentos médicos em português, gerando versões mais simples e acessíveis, além de extrair representações gráficas do seu conteúdo. Os resultados foram promissores, demonstrando a viabilidade da proposta em prover equidade no acesso a essa informação.*

1. Introdução

A saúde e o bem-estar das pessoas são preocupações importantes em qualquer sociedade. A área médica é uma das mais produtivas, gerando mais de 1,5 milhão de artigos por ano [González-Márquez et al. 2024]. No entanto, grande parte dessa informação não é acessível para a maioria da população brasileira, em especial porque os documentos produzidos por pesquisadores e médicos estão em inglês e, mesmo traduzidos, são complexos. Vale destacar que boa parte da população do Brasil possui escolaridade limitada ao Ensino Fundamental e pouca experiência com a leitura de textos científicos [Teixeira et al. 2021, Barbosa et al. 2023, Lima 2016, Finatto 2025].

Nesse contexto, a modelagem conceitual rigorosa desempenha um papel central. Tal modelagem requer, entre outras atividades, a identificação e a formalização dos

conceitos-chave do domínio da saúde, a explicitação de suas inter-relações semânticas, a sistematização de suas variações terminológicas e a definição dos padrões pelos quais esses conceitos se materializam nos diferentes registros linguísticos. Esse processo fundamenta-se na criação de esquemas conceituais que assegurem a integridade, a consistência e a interoperabilidade dos dados armazenados, sendo, portanto, imprescindível para o desenvolvimento de sistemas de informação eficazes. Trata-se, assim, de uma tarefa intrinsecamente interdisciplinar, que demanda a colaboração articulada entre especialistas da área da saúde, linguistas, terminólogos, cientistas da computação e profissionais da ciência da informação, todos convergindo para a construção de modelos que sirvam de base sólida para a gestão e recuperação qualificada da informação em ambientes computacionais.

Por outro lado, com o advento de modelos de linguagem de grande porte (*Large Language Models* - LLMs), muitas dessas tarefas têm sido automatizadas. No entanto, é importante avaliar se os modelos atuais conseguem refletir com fidelidade a complexidade do português brasileiro, especialmente em contextos sensíveis e críticos, como o da saúde.

Um modelo de linguagem é uma simplificação de um fenômeno complexo, no nosso caso, o funcionamento do português do Brasil por meio de textos escritos sobre temas de saúde. Esses modelos permitem representar a organização de uma língua por ferramentas computacionais. *"Embora um modelo tente capturar as nuances do fenômeno real, justamente por ser uma simplificação, ele não tem a intenção de substituir o fenômeno real, mas representá-lo para auxiliar o nosso entendimento ou resolver algumas tarefas"*[Paes et al. 2024].

Idealmente, qualquer modelo de linguagem precisa manter alguma consistência com o fenômeno real que ele representa. Assim, um modelo de linguagem deveria respeitar princípios do vocabulário (léxico), da sintaxe e da semântica da língua em foco.

Neste artigo, partimos da hipótese de que LLMs são capazes não apenas de gerar versões mais acessíveis de textos complexos, mas também de identificar, explicitar e reorganizar os principais conceitos presentes nesses textos, incluindo seus atributos e relações conceituais, de forma a preservar a integridade do conteúdo original. Para tanto, desenvolveremos um conjunto de estratégias metodológicas que envolvem tanto a construção dos *prompts* quanto a avaliação sistemática das respostas geradas pelas LLMs. A qualidade dos *prompts* não depende apenas da sintaxe ou do vocabulário empregado, mas da articulação entre os objetivos comunicativos e os mecanismos internos de funcionamento do modelo. Assim, aliaremos princípios da linguística aplicada e da análise textual ao uso criterioso de ferramentas de IA generativa, assegurando que a reescrita proposta pelos modelos preserve os significados essenciais, respeite a terminologia técnica e amplie o acesso ao conteúdo, sem comprometer sua precisão.

2. Trabalhos Relacionados

Uma ferramenta específica para auxiliar pessoas que se ocupem da simplificação textual na área da saúde é a MedSimples¹, que consiste em um sistema *on-line* gratuito. A ferramenta não realiza simplificação automática de um texto, mas oferece uma série de sugestões de reescrita, sobre terminologias, com definições simplificadas, e sobre o

¹<http://www.ufrgs.br/textecc/acessibilidade/page/cartilha/>

vocabulário em geral que pode ser complexo para os leitores em foco. Seu principal objetivo é promover a acessibilidade textual e terminológica, auxiliando a apresentação de informações mais fáceis de entender para um público amplo e diverso [Zilio et al. 2020]. Com técnicas de Processamento de Linguagem Natural e de um *corpus* de referência de português simples, o CorPop, a MedSimples ajuda a reformular conteúdos complexos, adaptando a linguagem para níveis mais acessíveis, sem comprometer a precisão e a qualidade da informação. Atualmente, todos os elementos fornecidos por ela ao usuário provêm de um *corpus* textual fechado, um *corpus* de domínio, segmentado por especificidades, que fornece padrões de vocabulário complexo-simples, associados a listas de palavras e a glossários terminológicos internos, manualmente construídos e validados por profissionais de saúde [Villar and Finatto 2024]. Assim, a ferramenta tende a desempenhar um papel fundamental na promoção da inclusão e na disseminação de conhecimento sobre saúde de forma mais equitativa e eficaz. O público-alvo principal da ferramenta inclui profissionais de saúde e redatores técnicos que precisam produzir conteúdos claros e diretos para trabalhadores brasileiros com escolaridade limitada e pouca experiência de leitura. No entanto, MedSimples não utiliza recursos de LLM ou de IA.

Atualmente, diferentes experimentos têm sido feitos com o uso de ferramentas de IA Generativa, nos quais se buscam simplificações automáticas de conteúdo, customizadas para diferentes finalidades e para o uso de diferentes tipos de leitores, conforme seus graus de letramento linguístico. [Leal et al. 2023], ao explorarem as plataformas Bard da Google, ao Copilot da Microsoft e ao ChatGPT 3.5 da OpenAi, solicitaram, via *prompt*, a seguinte ação: *Simplifique o seguinte texto para que um aluno do quarto ano do Ensino Fundamental consiga compreender, após a simplificação, forneça passo a passo os detalhes das mudanças e motivos para fazer as adaptações no texto.* Conforme os autores concluíram, foram surpreendidos positivamente pelos resultados desses sistemas na tarefa de simplificação textual automática. Todavia, como verificaram problemas, como a alucinação do sistema BARD, que “inventou” uma informação e a inseriu no texto simplificado, apontam ser necessária uma revisão humana da qualidade dessas simplificações, pelo menos por enquanto.

Em um contexto mais específico, em [Makhmutova et al. 2024] os autores discutem o uso de LLMs para simplificar e sumarizar textos médicos de maneira confiável. Eles propõem a criação de um *dataset* multi-modal para auxiliar a simplificação de textos médicos e permitir a comparação e a avaliação de modelos. Eles concluem que o estado atual das LLMs de propósito geral ainda não é confiável o suficiente para o contexto médico, embora sejam capazes de simplificar textos muito bem. Segundo os autores, que utilizaram o GPT como base, o modelo é capaz de resolver abreviações, reescrever textos de maneira mais compreensível e simplificar termos médicos. Já os maiores problemas estão relacionados com a criação de textos com viés, contendo eventual linguagem ofensiva ou fatos criados ao acaso (alucinação).

No âmbito da extração de entidades, em [Wang et al. 2025b] os autores utilizaram a arquitetura *transformer* para melhorar a extração de entidades na literatura da área médica. Eles avaliaram vários modelos pré-treinados, incluindo BERT, BioBERT, PubMedBERT e ClinicalBERT. Os experimentos por eles conduzidos mostraram que é possível realizar extrações com alta performance quando os modelos são contextualizados na área (F1-score de 88.6% para o PubMedBERT). Em cenários de poucos recursos,

abordagens de *few-shot learning* são efetivas, mesmo em condições limitadas.

De maneira similar, [García-Barragán et al. 2024] avaliam a capacidade de LLMs em extrair entidades a partir de registros médicos, escritos em espanhol, relacionados com câncer de mama. Para tanto, avaliam técnicas de *prompting* e estratégias de *fine-tuning*, comparam técnicas de extração de entidades tradicionais combinadas com BERT e GPT com *few-shot learning* e integração de conhecimento externo. O artigo identifica que a abordagem baseada em LLM com few-shot e conhecimento externo alcança valores similares com os do método tradicional, com BERT, eventualmente com F1-Score maior. Concluem que, pelo fato de não necessitar de anotação de dados, o potencial da LLM é grande para o processamento de dados.

Já em [Jiang et al. 2023], os autores desenvolvem um sistema interativo que converte as respostas textuais de um LLM em diagramas gráficos com o objetivo de facilitar tarefas de busca de informação e *Question-Answering*. O sistema foi denominado “Graphologue”, e aplica estratégias de *prompting* design de interface que realizam a extração de entidades e seus relacionamentos a partir de respostas de uma LLM, construindo diagramas em forma de grafos em tempo real. A ferramenta permite que os usuários interajam com o diagrama para explorar a informação e ajustar o *prompt* de maneira contextualizada para obter mais informações. Com a ferramenta, torna-se possível realizar diálogos não lineares com a LLM, facilitando a compreensão e a exploração da informação.

Finalmente, em [Wang et al. 2025a] os autores desenvolveram um *framework* que faz uso de LLMs para automatizar a extração e visualização de conhecimento estruturado a partir de documentos de especificação de requisitos. O *framework* é baseado em *prompts* específicos que utilizam o modelo GPT-4 para extrair informações-chave a partir de textos não estruturados, permitindo a identificação de requisitos funcionais e não funcionais. O resultado é utilizado para criar grafos de conhecimento. Os autores encontram uma série de limitações, em especial, desafios do GPT-4 em lidar com dependências hierárquicas e desambiguação de entidades, complexidades da linguagem (no caso, textos em chinês), falta de conhecimento específico de domínio e limitações relacionadas a alucinações e inconsistências.

Todos esses estudos demonstram que muitos avanços já foram feitos tanto na área de simplificação quanto na de extração de entidades. No entanto, ainda há uma série de desafios, em especial para a língua portuguesa, em contextos com poucos recursos (*low-resources*) e focados em áreas especializadas. Nossa contribuição reside em ilustrar o estado atual de ferramentas disponíveis ao público em geral, sem treinamento e conhecimento específico de contexto, em auxiliar pessoas leigas a compreender e extraír dados provenientes de bancos de dados abertos e não estruturados, como textos disponíveis na web.

3. Avaliando a Simplificação de Textos Médicos

Diante da possibilidade de uso de LLMs para realizar várias tarefas de simplificação, extração e representação de conteúdo, conforme exemplificado na seção de trabalhos relacionados, optamos por realizar uma avaliação, um experimento focado, permitindo-nos avaliar a real possibilidade de uso de tais ferramentas no domínio da saúde, eventualmente integrando LLM com ferramentas baseadas em PLN e conhecimento de domínio, como a

MedSimples.

Para este experimento, selecionamos os seguintes modelos: Claude3.7-sonnet* (Claude), GPT-4o* (GPT), DeepSeek V3-FW* (DeepSeek) e Gemini Advanced 2.0 flash (Gemini)².

O *prompt* apresentado a seguir foi elaborado e utilizado como base em todos os modelos: “*Olá. Você pode simular um profissional da área de saúde? Mais especificamente um médico que trabalhe com pacientes com câncer de próstata. Esse profissional deve ter o perfil de linguista, sabendo realizar tarefas de simplificação textual. Você receberá um texto complexo, da área médica, e sua tarefa será simplificar esse texto, para que uma pessoa adulta, leiga e com educação formal de 8 anos possa compreender o resultado. Não alucine e não inclua informações que não estejam no texto original*”.

Após o *prompt*, para cada LLM foi passado o conteúdo do texto sobre Câncer de Próstata disponível online no site do Instituto Nacional de Câncer (INCA), versão para profissionais da saúde³. Cabe salientar que esse conteúdo já possui duas versões, uma complexa, para profissionais da saúde, e uma simplificada, para pessoas leigas. Ambas foram desenvolvidas por humanos e serão utilizadas por nós como *ground-truth*, e esses dois documentos são nosso *dataset*.

Após analisar o conteúdo do texto complexo (denominado de Original Complexo) sobre Câncer de Próstata para profissionais da saúde, cada um dos modelos elaborou uma versão simplificada desse texto. Essas versões foram avaliadas por diferentes métricas desenvolvidas pelo NILC e disponibilizadas via sistema NILC-Metrix⁴ [Leal et al. 2023]. Como a quantidade de métricas é muito vasta (200), optou-se por um subconjunto, conforme sugerido por [Zilio et al. 2023].

Além da quantidade de palavras (Pal.), sentenças (Sent.) e parágrafos (Par.), as métricas adotadas foram as seguintes:

- **Diversidade de Palavras de Conteúdo** (DPC): Proporção de *types* de palavras de conteúdo em relação à quantidade de *tokens* de palavras de conteúdo no texto. Quanto maior o valor, maior a complexidade.
- **Proporção de palavras simples** (PPC): Proporção de palavras de conteúdo simples, sobre o total de palavras de conteúdo do texto. Quanto maior a proporção, menor a complexidade textual.
- **Índice Brunet** (IB): Avalia a riqueza e a variedade do vocabulário de um texto, considerando o tamanho do texto para identificar padrões de repetição de palavras. Valores típicos variam entre 10 e 20, sendo que um texto mais rico (e complexo) produz valores menores.
- **Fórmula Dale Chall adaptada** (DCA): A fórmula de leitabilidade de Dale-chall adaptada combina a quantidade de palavras não familiares com a quantidade média de palavras por sentença. Quanto maior o valor da métrica, maior a complexidade textual.
- **Índice Flesch** (IF): O Índice de Leitabilidade de Flesch busca uma correlação

²Claude, GPT e DeepSeek e Grookhttp://poe.com/, enquanto que Gemini foi acessado diretamente, via http://gemini.google.com

³<https://www.gov.br/inca/pt-br/assuntos/cancer/tipos/prostata>

⁴<http://fw.nilc.icmc.usp.br:23380/nilcmetrix>)

entre tamanhos médios de palavras e sentenças. Quanto maior o resultado da métrica, menor a complexidade textual.

- **Estatística de Honoré (EH):** A estatística de Honoré é um tipo de *type/token* ratio que leva em consideração, além da quantidade de *types* e *tokens*, a quantidade de *hapax legomena*. Quanto mais alto o valor, mais rico o texto é lexicalmente, o que está associado a maior complexidade.

Os resultados são apresentados na tabela seguinte, onde “Original Complexo” refere-se ao texto original criado para profissionais da saúde e “Original simplificado” é o texto original criado para leigos, conforme mencionado anteriormente nesta seção. Os demais são textos gerados pelos respectivos modelos.

Recurso	Pal.	Sent.	Par.	DPC	PPC	IB	DCA	IF	EH
Original complexo	2492	195	93	0,76306	0,51196	12,53301	12,28575	30,9838	1017,87925
Original simplificado	903	51	35	0,82437	0,82437	12,33088	10,51246	37,21104	962,31694
Claude	720	82	71	0,83331	0,75165	12,14392	9,62044	45,6953	1035,05412
GPT	518	60	44	0,65668	0,65909	12,94864	9,94786	72,01541	992,28776
DeepSeek	353	48	38	0,90918	0,76327	10,99445	10,99445	37,82292	1198,99717
Gemini	298	30	30	0,72435	0,76630	12,02331	9,16292	49,39629	895,43065

Tabela 1. Exemplo de tabela com nomes de textos e valores em ponto flutuante.

Pode-se perceber que o texto original, criado para profissionais da área da saúde, é realmente complexo, tendo uma grande quantidade de palavras, sentenças e parágrafos com expressões e elementos direcionados a esse perfil de leitor. Todos os indicadores de complexidade textual utilizados confirmam essa percepção. O Índice Flesch (IF), por exemplo, está próximo de 30, valor que indica texto com baixo grau de leitabilidade (quanto mais alto o valor, mais simples é o texto). Fato curioso é relacionado ao texto original simplificado, criado para pessoas leigas, pois ele também não apresenta valores adequados em termos de legibilidade para o público em geral.

Em relação aos textos gerados pelos modelos, todos obtiveram IF melhores do que os originais, tendo o GPT efetivamente obtido índice próximo ao que é esperado para pessoas com escolaridade mais limitada. Em termos de EH, os valores foram variados, inclusive com Claude atingindo valor mais alto do que o do texto original complexo. Mas outras LLMs obtiveram os melhores índices (os mais baixos). Isso ilustra que as LLMs, no geral, são capazes de obter textos mais simples, com uma quantidade menor de palavras.

4. Avaliando a Representação de Textos Médicos

Após verificar a capacidade de simplificação, nosso segundo objetivo consistiu em verificar se os principais elementos dos documentos originais estavam sendo mantidos em suas versões simplificadas. Para tanto, nossa proposta consistiu em gerar representações gráficas dos documentos – Diagramas E-R – tanto dos originais quanto dos gerados pelas LLMs. Pressupondo que os LLMs seriam capazes de identificar os principais conceitos (as entidades) e seus relacionamentos a partir de um documento, um diagrama E-R foi gerado para cada documento, i.e., um para o Original Complexo, um para o Original Simplificado e um para a versão simplificada por LLM (pela complexidade, optamos por utilizar somente o GPT, visto que foi o que obteve melhor *score* de IF). Considerando que esses diagramas representariam as principais entidades (conceitos) e seus relacionamentos, comparando-os entre si poderíamos verificar se os mesmos elementos estavam presentes em todas as versões.

Desta vez, definimos o seguinte *prompt*: “*Assuma o papel de um engenheiro de requisitos. Você receberá um texto e deverá realizar uma modelagem conceitual. Identifique as principais entidades e conceitos presentes nesse texto, seus atributos e relações*”.

O GPT gerou descrições textuais de um diagrama E-R, utilizando uma DSL (*Domain Specific Language*) compatível com ferramentas de diagramação. Com isso, usando uma ferramenta compatível com a DSL em questão, podemos gerar automaticamente diagramas. Para tal tarefa, utilizamos a ferramenta Mermaid⁵.

O resultado do texto complexo original pode ser observado na Figura 1. O do texto simples, também original, na Figura 2. O resultado da simplificação por LLM encontra-se na Figura 3.

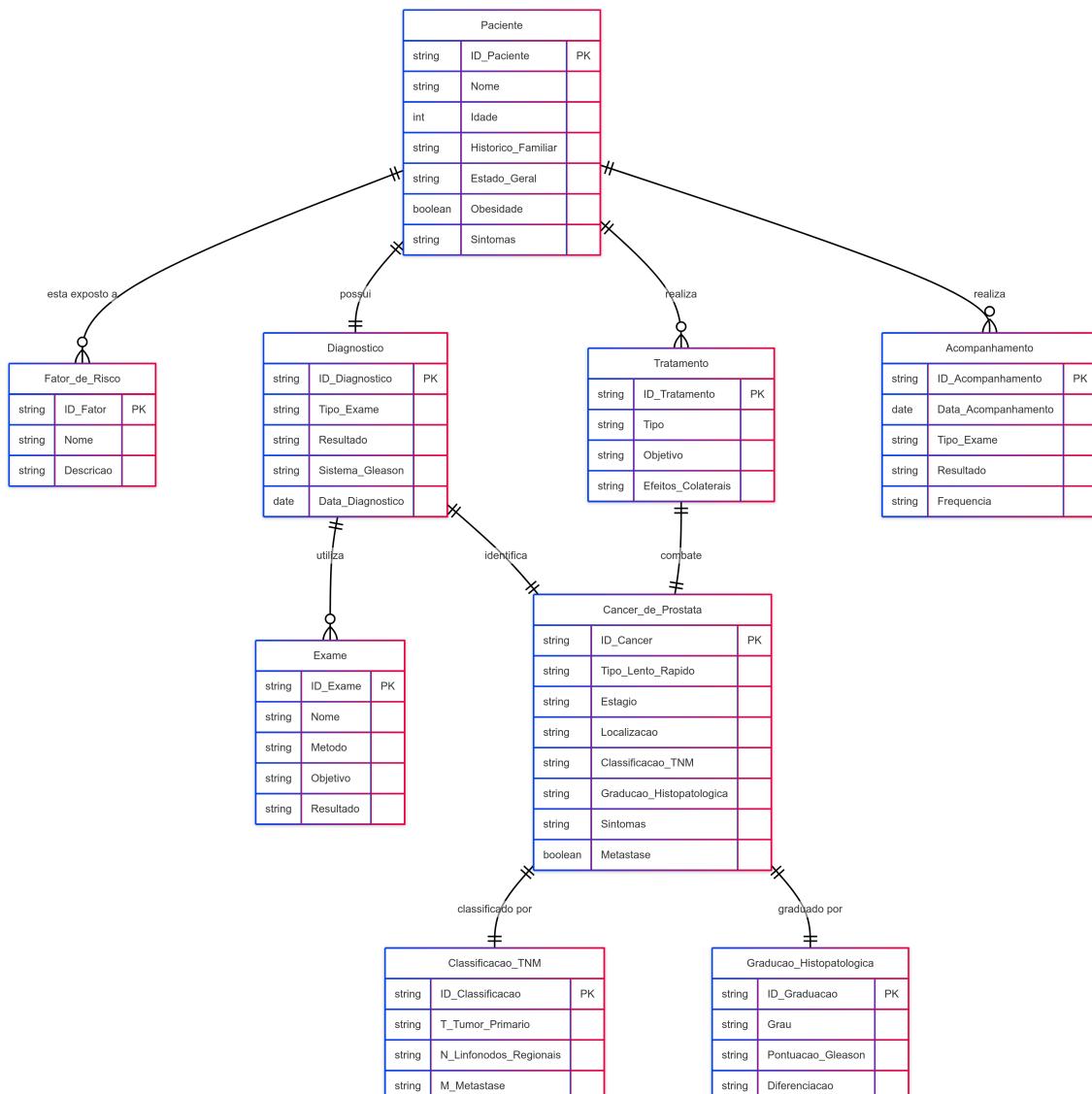


Figura 1. Diagrama do texto original complexo

É importante ressaltar que o modelo mais completo, i.e., com mais entidades e

⁵<http://www.mermaidchart.com/>

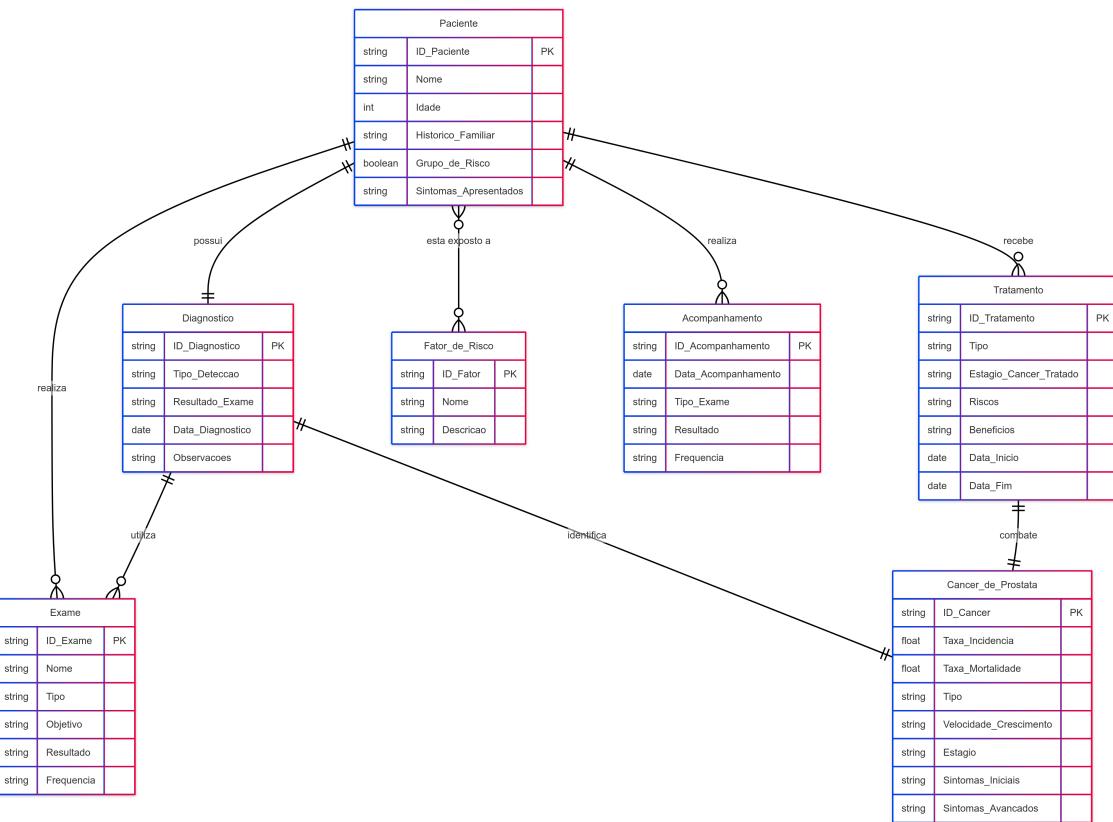


Figura 2. Diagrama do texto original simples

relações, foi gerado pelo texto original complexo. Isso era esperado. Apesar de sua complexidade, percebe-se que a LLM é capaz de identificar as principais entidades, seus atributos e relações. No diagrama, podemos identificar que *Pacientes* estão expostos a *Fatores de Risco* e através de *Diagnóstico* que utiliza *Exame* pode identificar *Cancer de Próstata* (com diferentes classificações e graus), o qual pode ser combatido através de um *Tratamento*. Ou seja, podemos supor que um modelo gerado por IA efetivamente serve para representar e até mesmo sumarizar os principais elementos de um documento.

O texto original simplificado (Figura 2) também contém os elementos principais do contexto (Homens podem desenvolver câncer de próstata, dependendo de fatores de risco; há sintomas específicos que podem ser detectados por diagnóstico precoce, dependente de exames).

Finalmente, da mesma forma, o texto simplificado por LLM (Figura 2) apresenta conceitos e relações similares, o que sugere que a IA consegue simplificar tão bem quanto o ser humano, extraíndo os conceitos e relações principais de determinado documento.

Cabe aqui salientar que o texto original simplificado consiste em parágrafos soltos, organizados em tópicos que não parecem estar encadeados ou sequenciados (ao menos textualmente). Já o texto simplificado possui encadeamento e organização lógica, fazendo com que o leitor siga uma linha de raciocínio e compreenda melhor o seu conteúdo.

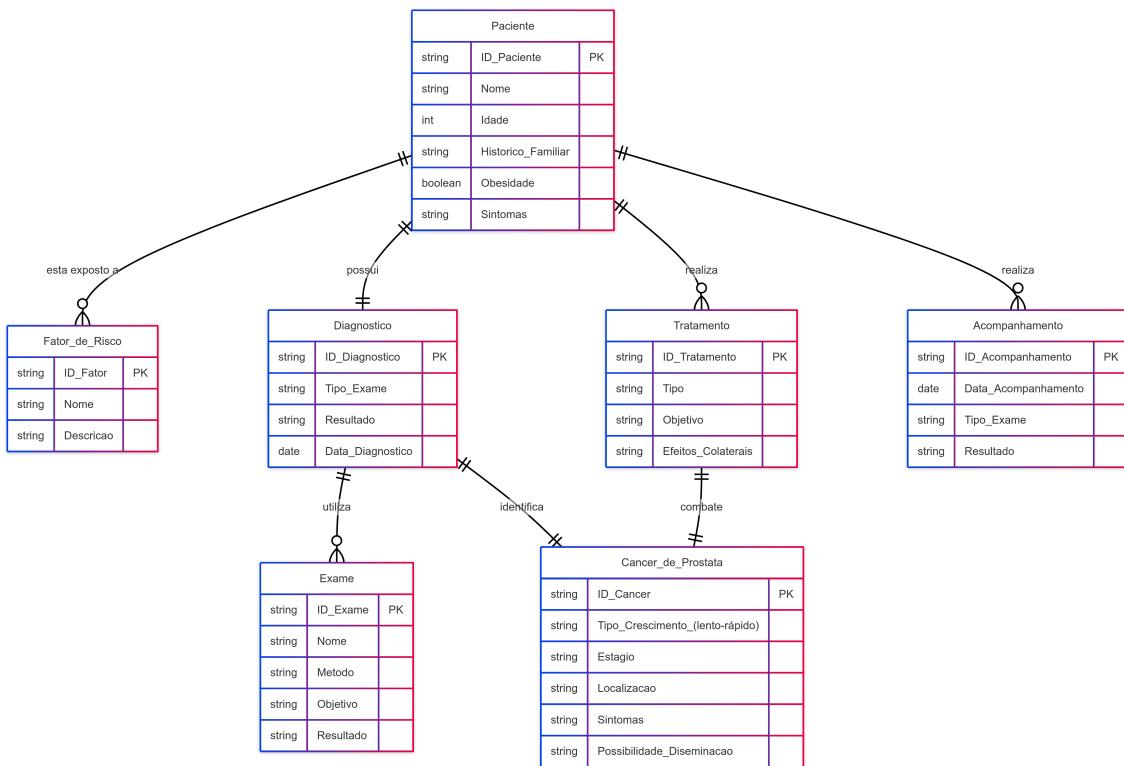


Figura 3. Diagrama do texto simplificado por LLM (GPT4o)

5. Considerações Finais

Este trabalho apresentou uma estratégia de simplificação e compreensão textual que faz uso de modelos de linguagem de larga escala (LLMs) com o objetivo de qualificar processos de simplificação textual no domínio da saúde. Partimos da premissa de que a capacidade de tais modelos em produzir reescritas acessíveis e semanticamente preservadas depende fortemente da qualidade dos *prompts* fornecidos, cuja construção deve se apoiar em conhecimento linguístico especializado, contemplando aspectos semânticos, sintáticos e pragmáticos da linguagem natural.

A importância desse tipo de análise é muito relevante, visto que grande parte da informação está disponível na web, que pode ser considerada um grande banco de dados textual, com informações mistas (estruturadas e desestruturadas), conflitantes e de diferentes complexidades, sem considerar desinformação e informação incorreta. Com isso, torna-se fundamental o desenvolvimento de abordagens que permitam aos usuários localizar, resumir e simplificar informações nesse âmbito. O estudo apresentado vai nessa linha, ilustrando que é possível simplificar e identificar conceitos-chave em documentos com o apoio de LLMs.

As principais limitações envolvem o domínio limitado do experimento, baseado em um único documento com duas versões (uma complexa e outra simplificada, mas ambas por humanos), e não termos feito uma análise mais profunda e comparativa do conteúdo que foi gerado, em especial passando por especialistas na área médica. Cabe salientar que nosso intuito foi o de verificar a capacidade de simplificação dos modelos, e os resultados ilustram que esse tipo de tarefa é viável. Há de se avaliar possível viés de trei-

namento, incluindo especificidades culturais e locais, além da veracidade das informações e a ocorrência de alucinações.

Como continuidade desta pesquisa, portanto, pretendemos agregar técnicas de *Retrieval Augmented Generation* (RAG) a fim de agregar conhecimento de domínio, além de considerarmos a geração ou o *fine-tuning* de *Small Language Models* (SLMs) específicos para o contexto médico. Também propomos o aprofundamento na formalização de esquemas ontolinguísticos para domínios especializados, a construção de *corpora* anotados com marcação semântico-pragmática e a definição de métricas específicas para avaliar a qualidade da simplificação textual com base em critérios conceituais e discursivos. Tais avanços permitirão a construção de *pipelines* mais controláveis e auditáveis, com potencial aplicação em ambientes educacionais, clínicos e administrativos. Este estudo reforça a necessidade de abordagens interdisciplinares e tecnicamente fundamentadas no uso de LLMs, especialmente quando aplicadas a áreas sensíveis como a comunicação científica e a disseminação de informações em saúde. Afinal, conforme pesquisa INAF de 2024, 58% da população do Brasil tem alfabetismo limitado aos níveis rudimentar e elementar [INAF 2024], sendo significativa a demanda por materiais acessíveis.

6. Agradecimentos

Este trabalho é parcialmente suportado e financiado pelo Projeto CNPq/MCTI No 10/2023 - Universal (402086/2023-6). Na parte linguística, tem apoio de bolsa PQ-CNPq (306273/2023-3).

Referências

- [Barbosa et al. 2023] Barbosa, J. P., Conrado, A. L., and Belusci, H. T. (2023). Pressupostos para a atualização do indicador de alfabetismo funcional. *Estudos em Avaliação Educacional*, 34:1–23.
- [Finatto 2025] Finatto, M. J. B. (2025). Coleta de dados em compreensão de leitura de textos médicos: Letramento em saúde e acessibilidade textual e terminológica. *Domínios De Lingu@gem*, 19:e019015.
- [García-Barragán et al. 2024] García-Barragán, Á., González Calatayud, A., Solarte-Pabón, O., Provencio, M., Menasalvas, E., and Robles, V. (2024). Gpt for medical entity recognition in spanish. *Multimedia Tools and Applications*.
- [González-Márquez et al. 2024] González-Márquez, R., Schmidt, L., Schmidt, B. M., Berrens, P., and Kobak, D. (2024). The landscape of biomedical research. *Patterns (N Y)*, 5(6):100968.
- [INAF 2024] INAF 2024. Alfabetismo no brasil. <https://alfabetismofuncional.org.br/alfabetismo-no-brasil/>. Accessed: 2025-05-05.
- [Jiang et al. 2023] Jiang, P., Rayan, J., Dow, S. P., and Xia, H. (2023). Graphologue: Exploiting large language model responses with interactive diagrams. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, UIST '23, New York, NY, USA. Association for Computing Machinery.

- [Leal et al. 2023] Leal, S. E., Duran, M. S., Scarton, C. E., Hartmann, N. S., and Aluísio, S. M. (2023). Nilc-metrix: assessing the complexity of written and spoken language in brazilian portuguese. *Lang Resources & Evaluation*.
- [Lima 2016] Lima, A. L. D. (2016). *Estudo especial sobre alfabetismo e competências socioemocionais na população adulta brasileira*, chapter Tópico 1: Estudo especial sobre alfabetismo e competências socioemocionais na população adulta brasileira. Instituto Paulo Montenegro; Instituto Ayrton Senna; eduLab21.
- [Makhmutova et al. 2024] Makhmutova, L., Salton, G., Perez-Tellez, F., and Ross, R. (2024). Automated medical text simplification for enhanced patient access. In *Proceedings of the 17th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 2: HEALTHINF*, pages 208–218. INSTICC, SciTePress.
- [Paes et al. 2024] Paes, A., Vianna, D., and Rodrigues, J. (2024). Modelos de linguagem. In Caseli, H. M. and Nunes, M. G. V., editors, *Processamento de Linguagem Natural: Conceitos, Técnicas e Aplicações em Português*, book chapter 17. BPLN, 3 edition.
- [Teixeira et al. 2021] Teixeira, E. M., Guaraciaba, A. P. B., dos Santos, A. C. A., Guedes, B. L. d. S., Terra, C. M. D., Campos, I. B., de Souza, L. M., Amaral, P., Dias, A. M. N., Mendes, N. B. d. E. S., and Corrêa, C. R. (2021). O impacto das informações médicas obtidas através da internet na relação médico-paciente. *Brazilian Journal of Health Review*, 4(6):25225–25239.
- [Villar and Finatto 2024] Villar, G. S. and Finatto, M. J. B. (2024). Acessibilidade textual e terminológica: novos glossários sobre oncologia para a ferramenta medsimple. *Mandinga: Revista de Estudos Linguísticos*, 7:23–42.
- [Wang et al. 2025a] Wang, L., Wang, M.-C., Zhang, Y.-R., Ma, J., Shao, H.-Y., and Chang, Z.-X. (2025a). Automated identification and representation of system requirements based on large language models and knowledge graphs. *Applied Sciences*, 15(7).
- [Wang et al. 2025b] Wang, X., Liu, G., Zhu, B., He, J., Zheng, H., and Zhang, H. (2025b). Pre-trained language models and few-shot learning for medical entity extraction.
- [Zilio et al. 2020] Zilio, L., Braga Paraguassu, L., Leiva Hercules, L. A., Ponomarenko, G., Berwanger, L., and Bocorny Finatto, M. J. (2020). A lexical simplification tool for promoting health literacy. In Gala, N. and Wilkens, R., editors, *Proceedings of the 1st Workshop on Tools and Resources to Empower People with REading DIfficulties (READI)*, pages 70–76, Marseille, France. European Language Resources Association.
- [Zilio et al. 2023] Zilio, L., Finatto, M. J. B., Vieira, R., and Quaresma, P. (2023). Uma abordagem de processamento de linguagem natural para avaliação de complexidade em literatura médica do século xviii. *Domínios de Lingu@gem*, 17:e1753.