



Aplicando o ChatGPT na condução de revisões sistemáticas da literatura: protocolo de pesquisa

Ananda Fernanda de Jesus

Doutoranda em Ciência da Informação, Universidade Estadual Paulista (UNESP), Marília, São Paulo, Brasil.

<http://lattes.cnpq.br/8553935355036462>

Email: af.jesus@unesp.br



José Eduardo Santarem Segundo

Pós-Doutor Faculdade de Engenharia da Computação da Western University, UWO, Canadá.

Docente do Programa de Pós-Graduação em Ciência da informação (PPGCI), Universidade Estadual Paulista (UNESP), Marília, São Paulo, Brasil.

Docente, Universidade de São Paulo (USP), Ribeirão Preto, São Paulo, Brasil.

<http://lattes.cnpq.br/5562746387565465>

Email: santarem@usp.br

Submetido em: 11/09/2023. **Aprovado em:** 30/04/2024. **Publicado em:** 18/07/2024.

RESUMO

O ChatGPT (Generative Pretrained Transformer), em sua terceira versão, é uma Inteligência artificial generativa, do tipo texto para texto, que tem sido testado e aplicado em diversos contextos, inclusive no contexto científico. A presente pesquisa visa como objetivo verificar as potenciais contribuições do ChatGPT para o preenchimento do protocolo de pesquisa de Revisões Sistemáticas da Literatura (RSL). Como método, foi realizada uma pesquisa empírica e comparativa, partindo da elaboração de três diferentes tipos de prompt: simples, detalhado e detalhado com exemplo. Como resultados, observou-se que foram gerados modelos diferentes de protocolos, inclusive quando uma segunda solicitação foi feita utilizando um mesmo prompt. Observou-se ainda oscilação entre os campos apresentados, principalmente os campos: objetivos, processo de seleção, estratégia de busca e formulário de extração. A presença ou ausência desses campos ocorre em diferentes respostas. Conclui-se que existem limitações ao uso do ChatGPT para o preenchimento de protocolos de pesquisa, entretanto, essas limitações podem ser minimizadas com a apresentação de um modelo de protocolo predefinido. Conclui-se ainda que a familiaridade do pesquisador tanto com o preenchimento de protocolos de pesquisa como com a temática que circunda a RSL também pode contribuir para minimizar essas limitações.

Palavras-chave: revisão sistemática da literatura; ChatGPT; inteligência artificial.

INTRODUÇÃO

Desde seu lançamento, em novembro de 2022, o *ChatGPT*, graças à sua característica de tecnologia de uso geral, passou a ser testado e aplicado em diversos contextos. “O *ChatGPT-3* (Generative Pretrained Transformer) é a terceira iteração do modelo de linguagem popular da *OpenAI*. Foi lançado em 2020 e é considerado um dos *Large Language Models* (LLM) mais avançados.” (Iftikhar; Iftikhar; Hanif, 2023, p. 1, tradução nossa)¹.

As Inteligências Artificiais (IA) Generativas são aquelas que não se limitam a realizar previsões e ações com base em conjuntos de dados, mas cujo foco é a criação de conteúdos, como textos e imagens.

Entretanto, nas gerações existentes, a criação ainda é limitada a uma recombinação dos dados utilizados para treino, com base em análise probabilística. (Gozalo-Brizuela; Garrido-Merchan, 2023).

Como indicado no próprio nome, o ChatGPT consiste em um *chatbot*, conceito que define ferramentas que se utilizam de processamento de linguagem natural (PLN) e técnicas de Machine Learning (ML) para “[...] interagir com os usuários de maneira humana” (Kooli, 2023, p. 2)².

O Processamento de Linguagem Natural (PLN) é compreendido como um campo científico que resulta em um conjunto de técnicas empregadas visando

[...] aproximar a linguagem computacional da linguagem humana, buscando oferecer formas de transformar o que uma pessoa escreve e fala em algo compreensível para as máquinas (Coneglian, 2020, p. 18).

Já Machine Learning (ML), ou aprendizado de máquina, é uma área da Inteligência Artificial, baseada na

[...] construção de agentes computacionais capazes de aprender com a experiência, com base na aplicação de técnicas estatísticas, em especial, por meio de algoritmos, visando a identificação de padrões e a realização de previsões [...] (Jesus *et al.*, 2023, p. 168).

Nesse sentido, no ChatGPT, as técnicas de Processamento de Linguagem Natural são utilizadas tanto para a compreensão do questionamento do usuário como para o fornecimento da resposta a esse questionamento, de maneira a simular o processo de comunicação humana, focado, nesse caso, na comunicação escrita. Já as técnicas de Machine Learning permitem que o ChatGPT seja treinado com base em grandes quantidades de dados e ainda que aprimore o seu funcionamento apoiado nas interações e nos feedbacks fornecidos pelos usuários.

1 Original: “The ChatGPT-3 (Generative Pretrained Transformer 3) is the third iteration of OpenAI’s popular language model. It was released in 2020 and is considered one of the most advanced large language models (LLM)”. (Iftikhar; Iftikhar; Hanif, 2023, p. 1).

2 Translation: “[...] interagir com os usuários de maneira humana” (Kooli, 2023, p. 2, editorial translation).

Aplicando essas técnicas, o ChatGPT é capaz de responder a perguntas, contestar premissas incorretas e rejeita solicitações inapropriadas (Gozalo-Brizuela; Garrido-Merchan, 2023).

O mecanismo básico empregado pelo *ChatGPT* é pegar uma sequência de entrada de texto (chamada *prompt*), processá-la através do modelo e emitir o próximo *token* na sequência. Este processo é repetido várias vezes para gerar uma resposta completa ao *prompt*. (Wang *et al.*, 2023, p. 4, tradução nossa)³.

A maneira como são formulados esses *prompts* pode influenciar na granularidade e na acurácia dos resultados retornados. Wang *et al.*, (2023) listam três formas principais de elaboração: simples, detalhada e com exemplo. Enquanto a formulação simples apresenta um questionamento direto, a detalhada acrescenta o contexto relacionado à formulação. Já com exemplos, apresenta um modelo daquilo que se espera obter como resultado ao questionamento elaborado.

Compreendendo a necessidade de testes para um maior entendimento das potenciais contribuições e riscos do uso da ferramenta no contexto científico, a presente pesquisa tem por objetivo: verificar as potenciais contribuições do ChatGPT para o preenchimento do protocolo de pesquisa de Revisões Sistemáticas da Literatura. As Revisões Sistemáticas da Literatura (RSL) podem ser compreendidas como revisões de literatura conduzidas com base na elaboração e aplicação de um protocolo de pesquisa e no registro do processo decisório do pesquisador, visando garantir a sua auditabilidade e replicabilidade. (Briner; Denyer, 2012). A existência de etapas manuais e exaustivas, e o longo investimento em recursos para a condução de RSLs tornam relevantes estudos que busquem automatizar um ou mais processos relacionados à sua condução.

ROCEDIMENTOS METODOLÓGICOS

A presente pesquisa, de caráter inicial, exploratório e aplicado, tem como finalidade identificar as potenciais aplicações do *GPT* na elaboração de protocolos de pesquisa que orientam a condução de Revisões Sistemáticas da Literatura. Como método, propõe-se a realização de uma pesquisa empírica e comparativa, tomando como base a forma de elaboração de *prompts* proposta por Wang *et al.*, (2023). Enquanto Wang *et al.*, (2023) testaram o uso dos *prompts* apenas na elaboração de booleanos para a RSL, nessa pesquisa propõe-se a elaboração de *prompts* que permitam o preenchimento de todos os campos do protocolo. O quadro 1 apresenta os *prompts* utilizados.

3 Original: "The basic mechanism employed by ChatGPT is that to take an input sequence of text (called prompt), process it through the model, and output the next token in the sequence. This process is repeated several times to generate a complete response to the prompt." (Wang *et al.*, 2023, p. 4)

QUADRO 1 – Prompts utilizados no estudo empírico.

ID	Tipo	Prompt	Objetivo da formulação
P1	Simple	Elabore um protocolo de pesquisa para uma Revisão Sistemática da Literatura Intitulada “Avaliação de qualidade de dados publicados como Linked Data”.	Verificar o resultado obtido com base em uma solicitação direta, apenas fornecendo o título da RSL.
P2	Detalhado	Considerando que você é um pesquisador com vasta experiência na condução de Revisões Sistemáticas da Literatura, elabore um protocolo de pesquisa para uma revisão intitulada “Avaliação de qualidade de dados publicados como Linked Data”, cujo objetivo é compreender de que forma tem sido discutida a qualidade de dados no contexto do Linked Data. Procure elaborar um protocolo que seja o mais eficiente possível, ou seja, recupere os documentos pertinentes, mas que priorize recuperar o mínimo possível de documentos irrelevantes.	Verificar o resultado quando mais detalhes são fornecidos, e é solicitado que a ferramenta preze pela eficiência, reduzindo ao máximo os ruídos.
P3	Detalhado	Considerando que você é um pesquisador com vasta experiência na condução de Revisões Sistemáticas da Literatura, elabore um protocolo de pesquisa para uma revisão intitulada “Avaliação de qualidade de dados publicados como Linked Data”, cujo objetivo é compreender de que forma tem sido discutida a qualidade de dados no contexto do Linked Data. Procure elaborar um protocolo que seja o mais abrangente possível, ou seja recupere o máximo de documentos pertinentes possível.	Verificar o resultado quando mais detalhes são fornecidos, e é solicitado que a ferramenta seja abrangente, recuperando o máximo de documento possível.
P4	Detalhado com exemplos	Considerando que você é um pesquisador com vasta experiência na condução de Revisões Sistemáticas da Literatura, e está pesquisando sobre avaliação de qualidade de dados publicados como Linked Data, para elaboração da sua monografia de conclusão de curso. Você está tentando obter uma visão geral sobre como tem sido discutida a qualidade de dados no contexto do Linked Data. Preencha um protocolo de pesquisa da revisão sistemática da literatura, contendo os seguintes campos: título; descrição; objetivo; questão principal; questões secundárias; palavras-chave e sinônimos; critérios para seleção das bases de dados; linguagem do estudo; bases de dados consultadas; critérios de inclusão; critérios de exclusão, critérios de avaliação de qualidade dos estudos; formulário de extração de dados; sumarização dos resultados.	Verificar o resultado quando um modelo de protocolo é fornecido.

Fonte: Elaborados pelos autores (2023).

Em relação a pesquisas correlatas, com base em uma análise exploratória, não foram identificados em periódicos científicos revisados por pares, outros artigos, além do trabalho de Wang *et al.*, (2023), que buscassem testar o potencial de aplicação do *ChatGPT* no preenchimento de protocolos de busca da RSL.

RESULTADOS

A presente seção será dividida em 4 subseções que irão refletir cada um dos *prompts* elaborados e apresentados no quadro 1⁴

P1–Prompt Simples

Nessa rodada, foram feitas duas solicitações utilizando o mesmo *prompt*, que retornaram dois modelos bastante distintos de protocolo de pesquisa. Ambos os modelos apresentam os campos fundamentais para embasar a condução da pesquisa, com exceção do campo formulário de coleta, considerado um campo importante por refletir aquilo que será efetivamente coletado nos documentos.

Em relação ao protocolo retornado na primeira solicitação, esse tem início diretamente na pergunta de pesquisa, não sendo apresentados os objetivos. O *ChatGPT* retorna três perguntas de pesquisa, não fazendo distinção entre pergunta principal e pergunta secundária. Das três perguntas apresentadas, apenas uma delas não é pertinente—Qual é a qualidade dos dados publicados como *Linked Data*? Já que não é uma prática viável identificar qual é a qualidade de um conjunto de dados. Esse modelo não apresenta uma estratégia de busca, apenas uma orientação sobre como elaborar essa estratégia. Os procedimentos de busca são divididos em bases de dados, repositórios de *preprint* e busca manual. Na busca manual, é sugerida a consulta às referências indicadas nos artigos aceitos, técnica conhecida com *snowballing*, muito realizada em conjunto com a RSL. Esses procedimentos, que incluem busca manual e outros tipos documentais, são geralmente associados a RSLs cujo objetivo é ser abrangente, recuperando o máximo possível de documentos publicados a respeito da temática.

Como critérios de exclusão, nessa versão, o *ChatGPT* apresenta campos como idioma, período e tipos de documentos, tendo delimitado a busca para artigos publicados a partir de 2010. Ao ser questionado sobre a exclusão de artigos publicados antes de 2010, ele lista três motivos: avanços tecnológicos da temática; maturidade do campo (documentos mais antigos podem não ter considerado alguns aspectos importantes) e relevância (visando eliminar potenciais ruídos). Essa decisão de excluir os primeiros artigos, como os publicados em 2006, ano do lançamento do *Linked Data*, contradiz a abordagem abrangente da RSL.

Os critérios de inclusão foram divididos de maneira a refletir os resultados esperados e facilitar a síntese dos resultados posteriormente, não realizar essa distinção é uma prática comum de pesquisadores iniciantes em RSL, que tendem a criar um critério único, do tipo: “(l) documentos que abordam a temática de interesse”. Em relação ao processo de seleção,

4 Por uma questão de limitação de espaço não foi possível apresentar os protocolos recuperados e as respostas obtidas na íntegra. Entretanto, esse material pode ser consultado em: <https://docs.google.com/document/d/1UtY6jgqHQ8wwJB3z6SnJYD0Pjzkv4PEE/edit?usp=sharing&oid=117525836395805599466&rtpof=true&sd=true>

o *ChatGPT* indica que esse será realizado em duas etapas: leitura dos títulos e resumos, e leitura completa. A análise dos resultados foi preenchida de maneira genérica, apontando que os resultados serão analisados quantitativa e qualitativamente.

O protocolo retornado por meio da segunda solicitação possui um caráter mais instrutivo, indicando o que deveria conter em cada campo, seguido de exemplos. O modelo contém uma introdução, enquanto o anterior iniciou na pergunta de pesquisa. Nessa introdução, o *ChatGPT* orienta que sejam apresentados os objetivos, justificativas, definições pertinentes e potenciais limitações.

Esse protocolo contém mais perguntas de pesquisa, entretanto, essas não são condizentes com o objetivo indicado no campo anterior. Nesse modelo de protocolo, são incluídas bases de dados no campo de estratégia de busca, diferente da primeira versão. Como exemplos de estratégia de busca, nessa versão de protocolo, o *ChatGPT* apresenta palavras-chave em inglês, mas novamente não elabora uma estratégia de busca. Os critérios de inclusão dessa segunda versão foram menos precisos, sendo adicionados critérios que seriam mais adequados à exclusão do que à inclusão.

P2–Prompt detalhado direcionado para eficiência dos resultados

Nessa versão, o *ChatGPT* não apresentou uma estrutura campo-preenchimento e sim uma sequência de etapas necessárias para preencher um protocolo. Essa resposta também possui um caráter mais explicativo, indicando a importância da presença de cada campo. Nas etapas indicadas, o primeiro passo seria o preenchimento das palavras-chave, não a identificação das perguntas de pesquisa, como normalmente ocorre. São apresentados potenciais termos em português, seguidos das bases a serem consultadas.

Nesse modelo, não são sugeridos bases de *preprint* e busca manual, o que é condizente com uma abordagem mais eficiente de RSL, que foi solicitada à ferramenta. São apresentadas as estratégias de busca para cada uma das bases de dados listadas, considerando as particularidades de pesquisa de cada base.

A próxima etapa mencionada foi o estabelecimento dos critérios de inclusão e exclusão. Os critérios apresentados refletem o idioma e o tipo de documento, sugerida a exclusão de artigos que não abordem diretamente a questão da RSL e artigos indisponíveis *on-line*. Um critério pouco usual indicado foi o de excluir documentos que não são revisões sistemáticas ou estudos primários. Os critérios de inclusão foram ainda mais elaborados do que os obtidos aplicando o p1, eles descrevem bem os resultados esperados, separando-os por categorias que já fornecem pistas sobre o que será coletado, permitindo uma pré-categorização dos documentos aceitos. O processo de seleção teve uma descrição bastante genérica e limitada, indicando que seria feita a aplicação dos critérios de inclusão e exclusão. Esse é o primeiro modelo a apresentar o campo formulário de extração, embora ainda seja

um formulário genérico e pouco relacionado com os objetivos e perguntas de pesquisa. Uma das principais limitações identificadas nesse modelo de protocolo foi a ausência de uma ou mais perguntas de pesquisa e a não explicitação do objetivo.

P3–Prompt detalhado visando abrangência nos resultados obtidos

Esse protocolo começa pela pergunta de pesquisa, sendo apresentada mais de uma pergunta, sem fazer distinção entre pergunta principal e secundária. Ambas as questões diferem das já apresentadas, provavelmente pelo contexto e pela apresentação do objetivo no *prompt* utilizado. Alguns dos critérios de inclusão apresentados seriam mais relevantes se apresentados como critérios de exclusão (limitações de idioma e de recorte temporal). Dois critérios de inclusão refletem os principais tipos de documentos que devem ser recuperados. Um dos critérios de inclusão é pouco pertinente, “Estudos que discutem a publicação de dados estruturados e conectados no ambiente *Web*,” tendo em vista que o critério “Estudos que abordam a qualidade de dados publicados como *Linked Data*,” é mais adequado ao tema, dispensando a inclusão do mesmo.

Os critérios de exclusão repetem os critérios apresentados anteriormente como critérios de inclusão, o que poderia causar confusão no processo de análise dos resultados. O protocolo apresenta termos mais abrangentes, como “Semantic Web quality”; “RDF quality”; “OWL quality” o que é condizente com a abordagem solicitada. Nesse protocolo, também não foi apresentada uma estratégia de busca.

O processo de seleção foi descrito de maneira detalhada, em etapas condizentes com um processo de seleção de RSL, especialmente RSLs com muitos resultados. O campo de análise de resultados indica que os artigos serão classificados por tema, sendo adotada uma análise qualitativa. Apresenta ainda um tópico genérico indicando que os resultados irão ser sintetizados, mas não descreve o processo de síntese. Nesse modelo aparece o campo de informações éticas, preenchido como não pertinente, com a justificativa de que o estudo não envolve participação de animais e seres humanos. Esse campo não é comum em RSL, por se tratar de um estudo secundário, sendo mais comum em modelos de formulários e questionários de coleta de campo.

P4 Prompt detalhado com apresentação de exemplo

Seguindo um protocolo pré-fornecido, a ferramenta apresentou respostas relacionadas diretamente com o contexto da pesquisa e com o que se espera encontrar de informações nos campos indicados, com exceção do campo onde deveria constar a justificativa para as bases de dados selecionadas. Esse modelo tem início na descrição da RSL. Essa descrição cobre o tema, o objetivo e os resultados esperados, embora não apresente o contexto de elaboração (monografia de conclusão de curso) que foi fornecido na descrição do *prompt*.

A questão principal está vinculada com o contexto, as questões secundárias representam bem os principais resultados esperados, divididos em: identificação de conceitos e abordagens, identificação de metodologias e identificação dos principais desafios. As palavras-chave foram apresentadas em português, com poucos sinônimos, algumas palavras-chave não são pertinentes e levariam à recuperação de ruídos, como Dados Abertos e Revisão Sistemática da Literatura.

Os critérios de inclusão são representativos, embora tenham incluído o recorte temporal que ficaria melhor como critério de exclusão. Os critérios refletem bem as questões secundárias, embora alguns critérios pudessem ser condensados em um único critério. Os critérios de exclusão são pertinentes.

Com a solicitação de preenchimento do critério de avaliação de qualidade, que também não foi apresentado nos modelos escolhidos pela ferramenta, foi apresentada uma metodologia de avaliação baseada em um referencial teórico relevante para a área de RSL. O formulário de coleta foi um pouco mais completo e menos genérico, indicando a coleta das abordagens utilizadas e os métodos de avaliação de qualidade de dados citados, mas ainda não reflete todas as problemáticas indicadas, como a questão dos desafios do processo. A sumarização é condizente, mas também bastante genérica.

CONSIDERAÇÕES FINAIS

A presente pesquisa é caracterizada como experimental e inicial, sem a pretensão de explorar todas as possibilidades e problemáticas relacionadas ao uso do *ChatGPT* para o preenchimento de protocolos de pesquisa de RSLs. A proposta da pesquisa foi analisar o potencial de contribuição da ferramenta, a partir de três diferentes formas de solicitação, demonstrando assim que tipos de resultados podem ser esperados, indicando algumas limitações dos resultados obtidos, com caráter qualitativo, com base nas experiências prévias dos pesquisadores com a condução de RSLs.

Ao analisar os resultados, observou-se que a ferramenta não possui preferência por um modelo base de protocolo de pesquisa, apresentando diferentes modelos, inclusive quando uma mesma pergunta é repetida. Existe uma oscilação entre os campos apresentados, principalmente os campos objetivos, processo de seleção, estratégia de busca e formulário de extração, cuja presença ou ausência ocorre em diferentes respostas. Devido à variedade de modelos de protocolo apresentados, ressalta-se a importância do conhecimento e seleção de um modelo de protocolo de pesquisa adequado, que terá influência direta nos resultados obtidos, minimizando algumas das limitações observadas. A escolha permite ainda que mais solicitações sejam feitas à ferramenta e que os resultados sejam diretamente comparados, já que os campos apresentados permaneceriam os mesmos. Outra limitação que se destaca é relacionada ao tipo de abordagem da RSL. Nem sempre as decisões tomadas foram condizentes com o tipo de abordagem solicitada, e quando essa abordagem não foi explicitada, houve escolhas contraditórias em um mesmo protocolo.

Dadas as limitações apresentadas, entende-se que o uso da ferramenta pode ser mais eficiente quando realizada por um pesquisador com familiaridade tanto com a condução de RSLs como com a temática pesquisada, o que permite a identificação mais precisa de falhas no protocolo apresentado. Entende-se que o uso do *ChatGPT* para essa finalidade por pesquisadores menos familiarizados com RSL e com a temática deve-se limitar a um caráter de aprendizagem e de exploração, já que esses pesquisadores podem ter dificuldades em identificar falhas no preenchimento.

Com base nos resultados apresentados, pode-se concluir que existe um potencial de aplicação do ChatGPT como uma ferramenta semiautomática para preenchimento de protocolos de pesquisa, entretanto, algumas problemáticas da aplicação do ChatGPT nesse contexto, que fugiram aos objetivos da presente pesquisa, precisam ser discutidas em profundidade em estudos futuros, tais como:

- Não identificação das fontes consultadas – o ChatGPT não apresenta referências para as informações fornecidas ou a fonte dos dados consultados na elaboração das respostas;
- Respostas diferentes para uma mesma pergunta – a cada novo prompt, o ChatGPT retorna diferentes resultados, o que prejudica a auditabilidade desses resultados, aspecto importante da condução de uma RSL;
- Alucinações e erros – que culminam na elaboração de resultados imprecisos ou incorretos, gerados quando o ChatGPT não possui uma resposta adequada à pergunta. Justamente para evitar que esses erros ocorram, a liberdade de criação da ferramenta precisa ser limitada, o que pode levar à próxima questão, a possibilidade de plágio.
- Possibilidade de plágio – como o modelo generativo baseia seus resultados na recombinação de informações provenientes da alimentação previa e das interações com outros usuários, é possível que as informações geradas impliquem em plágio. Essa interação também poderia comprometer o ineditismo de pesquisas em andamento.

Como estudos futuros, destaca-se ainda a necessidade de ampliação dos testes. Incluir temas diversos de pesquisa pode permitir a identificação de outras limitações, além das apresentadas. Também se torna pertinente ampliar a quantidade de solicitações feitas utilizando o mesmo *prompt*, visando mapear a frequência com que aparecem os mesmos resultados e as diferentes respostas retornadas. Conclui-se que existem muitas limitações ao uso do *ChatGPT* para o preenchimento de protocolos de pesquisa de RSLs, e que o uso do *chat* para essa finalidade exige cautela. Conclui-se ainda que as limitações apresentadas podem ser minimizadas com a elaboração de um *prompt* mais completo, contendo um protocolo predefinido. Outro fator que pode minimizar essas limitações é o nível de familiaridade do pesquisador tanto com o preenchimento de protocolos de pesquisa como com a temática que circunda a RSL.

REFERÊNCIAS

BRINER, R.; DENYER, D. Systematic Review and Evidence Synthesis as a Practice and Scholarship Tool. In: ROUSSEAU, D. M. (ed.). **The Oxford Handbook Of Evidence-Based Management: companies, classroom and research**. [S. l.]: Oxford University Press, 2012. cap. 7, p. 112-129. DOI 10.1093/oxfordhb/9780199763986.013.0007.

CONEGLIAN, C. S. **Recuperação da Informação com abordagem semântica utilizando Linguagem Natural: a Inteligência Artificial na Ciência da Informação**. Orientador: Dr. José Eduardo Santarem Segundo. 2020. 194 f. Tese (Doutorado) – Curso de Programa de Pós-Graduação em Ciência da Informação, Universidade Estadual Paulista, Marília, 2020. Disponível em: https://repositorio.unesp.br/bitstream/handle/11449/193051/coneglian_cs_dr_mar.pdf?sequence=3&isAllowed=y. Acesso em: 8 set. 2022.

GOZALO-BRIZUELA, R.; GARRIDO-MERCHAN, E. C. ChatGPT is not all you need. A State of the Art Review of large Generative AI models. **Arxiv**, [s. l.], v. 1, n. 1, p. 1-22, 2023. Disponível em: <http://dx.doi.org/10.48550/ARXIV.2301.04655>. Acesso em: 18 maio 2023.

IFTIKHAR, L.; IFTIKHAR, M. F.; HANIF, M. I. DocGPT: impact of ChatGPT-3 on health services as a virtual doctor. **Ec Paediatrics**, [s. l.], v. 12, n. 3, p. 45-55, Mar. 2023. Disponível em: https://www.researchgate.net/publication/369013064_DocGPT_Impact_of_ChatGPT-3_on_Health_Services_as_a_Virtual_Doctor. Acesso em: 18 maio 2023.

KOOLI, C. Chatbots in Education and Research: a critical examination of ethical implications and solutions. **Sustainability**, [s. l.], v. 15, n. 7, p. 5614, mar. 2023. Disponível em: <http://dx.doi.org/10.3390/su15075614>. Acesso em: 29 jun. 2023.

JESUS, A. F.; TRIQUES M. L.; SEGUNDO, J. E. S.; ALBUQUERQUE, A. C. Processamento de linguagem natural e machine learning na categorização de artigos científicos: um estudo em torno do “patrimônio cultural”. **Revista Ibero-Americana de Ciência da Informação**, [s. l.], v. 16, n. 1, p. 167–184, jan./abril. 2023. DOI: 10.26512/rici.v16.n1.2023.47537. Disponível em: <https://periodicos.unb.br/index.php/RICI/article/view/47537>. Acesso em: 29 jun. 2023

WANG, S.; SCELLS, H.; ZUCCON, G.; KOOPMAN, B. Can ChatGPT Write a Good Boolean Query for Systematic Review Literature Search? **Arxiv**, [s. l.], v. 1, n. 1, p. 1-19, Feb. 2023. Disponível em: <http://dx.doi.org/10.48550/ARXIV.2302.03495>. Acesso em: 18 maio 2023.

AGRADECIMENTOS

Agradecemos à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo financiamento recebido para o desenvolvimento dessa pesquisa. PROCESSO FAPESP nº 2021/03349-0.

