




RECUPERAÇÃO DE INFORMAÇÃO NA WEB COM O BARD: uma experiência com a inteligência artificial generativa da Google

INFORMATION RETRIEVAL ON THE WEB WITH BARD: an experience with Google's generative artificial intelligence

Patrícia Nascimento Silva¹

 [0000-0002-2405-8536](https://orcid.org/0000-0002-2405-8536)

RESUMO

O Bard é o novo serviço de Inteligência Artificial (IA) da Google, lançado no Brasil em julho de 2023. Assim como outros modelos de IA, a proposta se assemelha a um *chatbot* que interage com os usuários, em linguagem natural, para recuperar informações na web e em bases de dados privadas e dedicadas. Este relato de experiência integra um projeto maior que tem o objetivo de analisar a recuperação de informação na Web com ferramentas de IA. Para tanto, inicialmente foram criados três critérios para seleção de um domínio: (1) Conhecimento prévio dos autores sobre o assunto; (2) Temática de contexto mundial e (3) Resultados que poderiam retornar informações textuais e outros objetos digitais. Após a definição do domínio, que selecionou a temática dados abertos, foram criadas dez expressões de busca, no formato de perguntas, associadas à tipologia e necessidade de informação. Para analisar as respostas foram criadas quatro classificações: totalmente pertinente, pertinente, parcialmente pertinente e incorreto. As perguntas foram executadas no Bard em julho de 2023 e foram analisados os três rascunhos oferecidos pela ferramenta. A maioria das respostas obtidas eram parcialmente pertinentes (40%) ou pertinentes (33,3%), duas respostas estavam incorretas (6,6%) e seis respostas (20%) estavam totalmente pertinentes, por atenderem a necessidade de informação e com a tipologia esperada. O Bard mostrou ser uma ferramenta interessante para recuperação de informação, mas ainda precisa de melhorias, principalmente em relação à fonte, que não é informada, e a precisão dos resultados que podem ser otimizados.

Palavras-Chave: recuperação de informação; inteligência artificial; Ciência da Informação; dados abertos; Google Bard.

Artigo submetido em 28/07/2023 e aceito para publicação em 29/12/2023.

¹ Doutora em Gestão e Organização do Conhecimento pela Universidade Federal de Minas Gerais. Universidade Federal de Minas Gerais, Programa de Pós-Graduação em Gestão e Organização do Conhecimento. E-mail: patricians@ufmg.br.

ABSTRACT

Bard is Google's new Artificial Intelligence (AI) service, launched in Brazil in July 2023. Like other AI models, the proposal resembles a chatbot that interacts with users in natural language to retrieve information on the web and in private and dedicated databases. This experience report is part of a larger project that aims to analyze information retrieval on the Web with AI tools. To this end, three criteria were initially created for selecting a domain: (1) the authors' prior knowledge of the subject; (2) a global context; and (3) results that could return textual information and other digital objects. After defining the domain, which selected the open data theme, ten search expressions were created, in the form of questions, associated with the typology and need for information. Four classifications were created to analyze the answers: totally pertinent, pertinent, partially pertinent, and incorrect. The questions were run on Bard in July 2023 and the three drafts offered by the tool were analyzed. Most of the answers obtained were partially pertinent (40%) or pertinent (33.3%), two answers were incorrect (6.6%) and six answers (20%) were totally pertinent, as they met the need for information and had the expected typology. Bard proved to be an interesting tool for information retrieval, but it still needs improvement, especially about the source, which is not informed, and the accuracy of the results, which could be optimized.

Keywords: *Information Retrieval. Artificial Intelligence. Information Science. Open data. Google Bard.*

1 INTRODUÇÃO

O Bard é o novo serviço de Inteligência Artificial (IA), lançado mundialmente pela empresa Google, em março de 2023. Conforme a Google (2023a) o “Bard é uma tecnologia experimental com o modelo PaLM 2 do Google para a colaboração entre usuários e a IA generativa”. Similar a um *chatbot*, o serviço é concorrente do modelo de linguagem de IA ChatGPT, lançado em novembro de 2022 pela OpenAI, e se propõe a auxiliar os usuários em diversas atividades que vão desde o planejamento de uma festa de aniversário até análises sobre um tema complexo. Assim como o ChatGPT, o serviço se propõe a construir textos, códigos e outros recursos a partir de um modelo de linguagem de IA, baseado em uma pergunta do usuário, escrita em linguagem natural.

No Brasil, o Bard foi disponibilizado em 13/07/23, após ser treinado no idioma português brasileiro, e diferentemente de outros modelos, o Bard está associado aos produtos e serviços da Google, uma das empresas que dominam o mercado de serviço de e-mail e armazenamento em nuvem. Com isso, os termos de uso e as mensagens indicadas na ferramenta informam ao usuário que o serviço utiliza os dados da conta Google do usuário.

Ao aceitar os termos de uso e acessar a ferramenta é informado aos usuários que há revisores humanos que processam as conversas para garantir maior qualidade, que a ferramenta é experimental e pode fornecer informações imprecisas ou inadequadas e recomenda, na tela principal da interface, que informações confidenciais não sejam inseridas. Essas informações demonstram que a empresa já se atualizou frente às discussões sobre responsabilização da IA, muito discutida a partir de 2023 em diversos países.

Com relação à recuperação de informação, a ferramenta se propõe a realizar buscas na Web, citando as fontes e *links* originais do conteúdo recuperado, e apresentá-la aos usuários de forma intuitiva como respostas de uma conversa em linguagem natural. O Bard, inclusive, foi definido pela empresa como uma experiência complementar à pesquisa Google (Google, 2023a), sendo esta uma evolução já esperada pelo mercado, a partir das últimas evoluções tecnológicas.

Com o surgimento dos computadores e criação do termo recuperação de informação em 1951, vários modelos de recuperação de informação foram propostos a partir da década de 1970 e se tornaram base para os sistemas de informação da época. Mas além dos modelos clássicos, evoluções foram percebidas com a criação dos modelos dinâmicos, especialmente após a criação da Web, em 1989. Para este novo ambiente, o uso da linguagem natural e técnicas de IA foram acrescentados aos modelos e motores de busca utilizados na atualidade.

A fim de apresentar a nova ferramenta à comunidade acadêmica e fomentar novos estudos e possíveis aplicações no campo da Ciência da Informação, este relato de experiência possui orientação prática/tecnológica e pretende responder a seguinte questão: como e quais tipos de informação e recursos são recuperados pelo *chatbot* Bard? O objetivo do estudo foi analisar a recuperação de informação na Web com a ferramenta Bard. Para tanto, foi selecionado um domínio e definidos critérios de busca para analisar a ferramenta. Destaca-se que este relato de experiência não é exaustivo e trata-se de um estudo inicial, integrante de um projeto maior que envolve a comparação com outras ferramentas.

O estudo justifica-se por sua atualidade e relevância, bem como por buscar compreender e fazer uma breve observação do recente fenômeno

informacional, envolvendo a aplicação da IA, no contexto da recuperação de informação e no âmbito da Ciência da Informação, áreas totalmente envolvidas que demandam novos estudos e pesquisas.

2 RECUPERAÇÃO DE INFORMAÇÃO

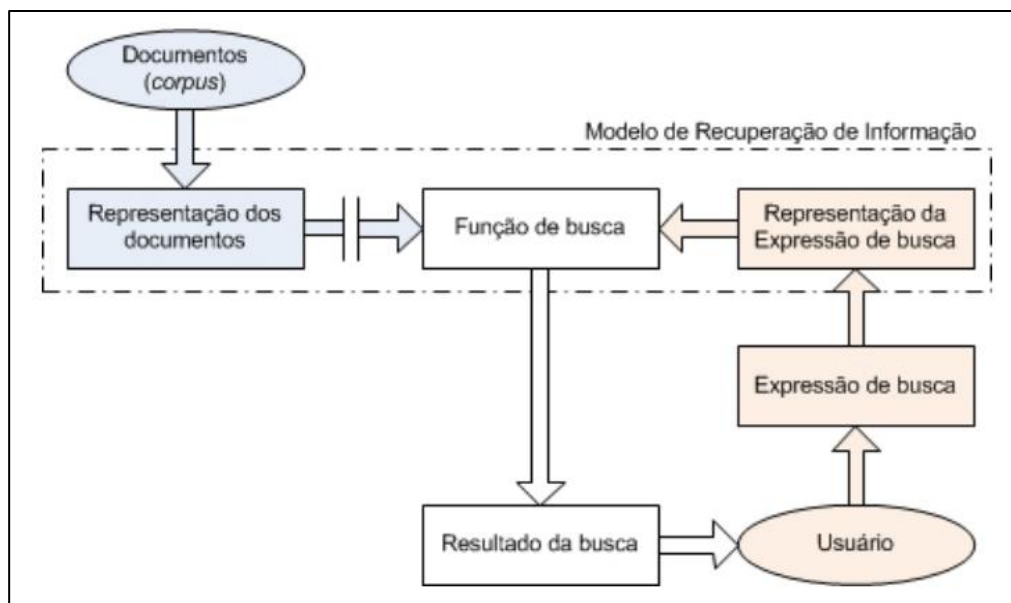
A Recuperação de Informação é uma área que recebeu contribuições relevantes tanto da Ciência da Computação quanto da Ciência da Informação de forma que essa relação de simbiose entre as duas áreas, inicialmente conflituosa, é cada vez mais marcante e indissociável, sobretudo a partir dos anos 1990, com o surgimento da era digital e a crescente apropriação da informação pelos usuários (Nascimento Silva, 2023a).

A IA, área inicialmente pertencente às Ciências Cognitivas, têm ganhado destaque em diferentes contextos e aplicações, se tornando um campo de pesquisa multidisciplinar. Conforme Saracevic (1996, p.7) o primordial das ciências cognitivas “é a conexão individualizada na abordagem para temas diversos como cérebro e mente, das humanidades às ciências da vida, das ciências sociais às matemáticas, da lógica às engenharias”. Assim, por envolver numerosas temáticas e permitir diferentes conexões, a IA tem sido incorporada em diversas áreas e domínios do conhecimento.

No contexto da Recuperação de Informação, os modelos dinâmicos surgiram como uma evolução dos modelos clássicos e utilizam técnicas oriundas da IA. Esta é uma alternativa promissora e muito estudada recentemente, pois permite que o usuário interaja e interfira diretamente no processo de recuperação da informação. Além disso, permite a evolução e adaptação da informação de acordo com os interesses de busca do usuário e interações com o sistema (Baeza-Yates; Ribeiro-Neto, 2013; Ferneda, 2003).

Um modelo de recuperação de informação envolve três elementos principais: a **função de busca** que interage com a **representação dos documentos** existentes nos sistemas e com a **representação da expressão de busca** formulada pelo usuário, conforme apresentado na Figura 1. O usuário interage com o sistema por meio da expressão formulada e o sistema responde com o resultado da busca. Essa interação pode ser modelada em interfaces tradicionais ou adicionando técnicas e recursos modernos como nos *chatbots*.

Figura 1 - Representação do Processo de Recuperação



Fonte: Ferneda (2012, p.14).

Os *chatbots* são ferramentas resultantes da combinação entre sistemas de recuperação da informação, processamento da linguagem natural e IA. Considerados assistentes virtuais, os *chatbots* são implementados para ter uma conversa com um humano em linguagem natural. Eles fornecem respostas com base em uma combinação de roteiros pré-definidos e programas de aprendizado de máquina que utilizam técnicas de IA. Com isso, o *chatbot* envia respostas conforme o conhecimento disponível no sistema de recuperação e nos bancos de dados vinculados, mas também poderá ser treinado para conversas que introduzam um novo conceito não programado (Queiroz; Valls, 2022).

Conforme Castor *et al.* (2021) os *chatbots* estão cada vez mais presentes entre as empresas como um meio interativo e de fornecimento de informações aos seus clientes, atuando na vida das pessoas e auxiliando-as em tarefas desde as mais simples às mais complexas. Com isso, a avaliação da recuperação é uma métrica importante que associa os resultados produzidos por um sistema de recuperação em resposta a um conjunto de consultas do usuário. Essa métrica deve estar diretamente associada à relevância dos resultados para os usuários que irão comparar o resultado produzido pelo sistema com os resultados sugeridos por humanos para o

mesmo conjunto de consultas (Baeza-Yates; Ribeiro-Neto, 2013).

Com a significativa evolução dos sistemas de recuperação de informação, o desafio atual reside em informar o que é relevante para um ser humano, o que pode tratar de uma questão subjetiva, pessoal, intrínseca e imparcial. Assim, a Ciência da Computação e a Ciência da Informação têm sido a base para os estudos de recuperação de informação. Na Ciência da Computação a preocupação reside em avaliar e evoluir tecnicamente a recuperação de informação, na Ciência da Informação o enfoque permeia o usuário e a avaliação deste sobre o sistema (Lewandowski, 2012; Monteiro *et al.*, 2017).

Neste sentido, estudos que analisem a recuperação de informação em ferramentas baseadas na IA são necessários para identificar se as mesmas estão cumprindo com os objetivos ao qual foram desenvolvidas e principalmente se estão atendendo as necessidades dos usuários.

3 METODOLOGIA

O estudo caracterizou-se como uma pesquisa descritiva e exploratória, apresentado como um relato de experiência, com abordagem qualitativa, por meio de uma análise documental, para compreender a recuperação de informação na Web com a ferramenta Bard. O relato de experiência é um tipo de produção que trata da vivência acadêmica ou profissional, cuja característica principal é a descrição da intervenção, apresentada com embasamento científico e reflexão crítica de práticas ou intervenções (Mussi; Flores; Almeida, 2021).

A pesquisa realizada também pode ser definida como pesquisa aplicada e pesquisa ação, por envolver a participação ativa da autora, e por seu objetivo acadêmico ser a produção de conhecimento para aplicação prática e imediata (Tripp, 2005).

Destaca-se que este estudo é parte de um projeto maior, que investiga ferramentas de IA com foco na recuperação de informação na Web. Com isso, a metodologia utilizada nesta pesquisa foi criada pela autora em Nascimento Silva (2023b) e validada inicialmente com o ChatGPT, ferramenta que possui amplitude e propósito similar ao Bard: construir um texto (conversa) ou retornar objetos (códigos-fonte, *links*, vídeos, imagens) sobre qualquer assunto.

Devido à amplitude de domínios e temáticas abordadas por essa tipologia de ferramenta, foram criados critérios para seleção do domínio de análise que considerou: (1) Conhecimento prévio dos autores sobre o assunto, que poderia validar os resultados apresentados pela ferramenta; (2) Temática de contexto mundial, para que fosse possível analisar resultados globais e (3) Resultados que poderiam retornar informações textuais e outros objetos digitais, a fim de identificar os recursos disponibilizados pela ferramenta. A temática selecionada para este estudo foi: dados abertos, no idioma português. Essa temática é originalmente complexa e diversa, e como resultado pode retornar arquivos e conjuntos de dados de diversas tipologias. É importante salientar que a temática é geral e essa definição foi pensada para não limitar os resultados recuperados pela ferramenta.

Após a definição do domínio foram criadas expressões de busca, ou seja, perguntas em linguagem natural, simulando a interação do usuário com a ferramenta. Destaca-se que as buscas não foram exaustivas e que foram construídas perguntas na norma culta da língua portuguesa (português brasileiro), um padrão diferente do utilizado em motores de buscas conhecidos na *internet*. Os critérios consideraram pesquisas mais amplas, inicialmente, e elementos mais específicos, no contexto dos dados abertos brasileiros, em seguida, a fim de identificar os principais conceitos e validar os resultados recuperados. O Quadro 1 apresenta a relação do tipo de busca, ou seja, a tipologia esperada/desejada, com a respectiva necessidade de informação do usuário e a expressão de busca criada, que corresponde a pergunta para a ferramenta.

Quadro 1 - Intenções e expressões de busca

Tipo de busca	Necessidade de Informação	Expressão de busca
Busca textual	Identificar conceitos e informações gerais sobre a temática de dados abertos.	(P1) O que são dados abertos? (P2) Apresente exemplos de dados abertos. (P3) Exiba dados abertos disponibilizados no Brasil.
Busca por objetos, <i>links</i> , arquivos	Identificar conjuntos de dados abertos específicos no contexto brasileiro.	(P4) Quais são os dados abertos sobre vacinação no Brasil? (P5) Onde acessar dados abertos sobre vacinação no Brasil? (P6) Exiba arquivos de dados abertos sobre a vacinação no Brasil.
Busca por recursos: ferramentas, APIs, códigos-fonte	Identificar ferramentas para manipulação de dados abertos e como acessá-las.	(P7) Quais ferramentas disponíveis para acesso, uso e reúso dos dados abertos? (P8) Quais são as APIs de dados abertos disponíveis no Brasil? (P9) Como acessar as APIs de dados abertos? (P10) Exiba um código para coleta de dados abertos

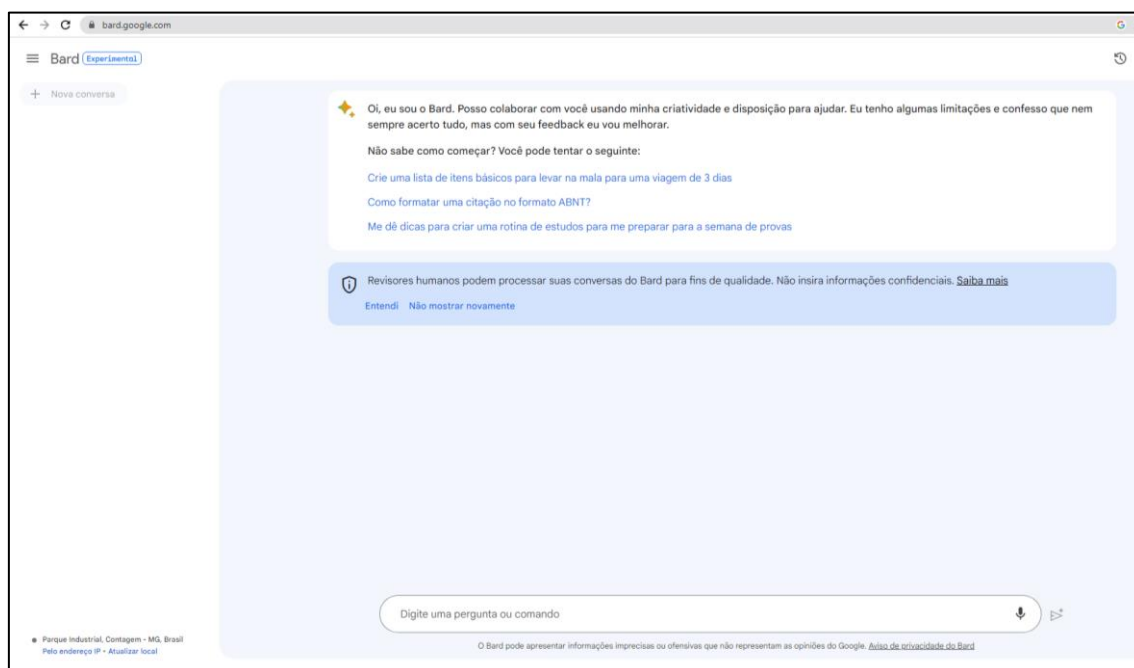
Fonte: Dados da pesquisa (2023).

As intenções e expressões de busca estão associadas a resultados textuais e possíveis objetos digitais, uma vez que a ferramenta propõe uma recuperação de diversos recursos. Inclusive, essa diversidade de recursos caracteriza-se como um dos motivos da popularidade desse tipo de ferramenta entre os usuários.

As 10 expressões de buscas foram executadas em julho de 2023 e acessadas as três opções de rascunho disponibilizadas no Bard, por meio de um acesso gratuito. A execução considerou os três rascunhos, pois a proposta da ferramenta é gerar textos diferentes para cada resposta. Foi utilizada a interface Web da ferramenta, opção “Nova conversa”, disponível no endereço: <https://bard.google.com/>, com a versão de 13/07/23, conforme apresentado na Figura 2.

O detalhamento completo dos resultados obtidos foi disponibilizado integralmente no repositório desta pesquisa, no Mendeley Data, criado pela autora, Nascimento Silva (2024).

Figura 2 - Interface da ferramenta Bard



Fonte: Google (2023b).

As respostas para as perguntas formuladas foram analisadas e classificadas, conforme a tipologia esperada e a necessidade de informação do usuário, em quatro categorias:

1. Totalmente pertinente (TP): quando o resultado era totalmente adequado à tipologia esperada e a necessidade de informação;
2. Pertinente (P): quando o resultado era adequado à necessidade de informação, mas não apresentava o formato esperado;
3. Parcialmente pertinente (PP): quando o resultado apresentava parcialmente os resultados esperados, tanto na tipologia quanto na necessidade de informação;
4. Incorreta (I), quando o resultado não estava adequado à necessidade de informação ou apresentava resultados incorretos.

A análise e discussão dos resultados consideraram a avaliação da recuperação, observando a qualidade dos resultados apresentados para as perguntas, qualitativamente, sem considerar a avaliação do desempenho do sistema, que incluiria, por exemplo, o tempo de processamento das buscas. Na próxima seção, as respostas recuperadas no Bard foram analisadas,

classificadas e discutidas sob o olhar da Ciência da Informação, especialmente no que tange às técnicas de recuperação de informação na Web.

4 RESULTADOS E DISCUSSÃO

A avaliação dos resultados pautou-se no atendimento às categorias definidas e foi baseado no conhecimento específico da área de dados abertos pela autora. Independentemente das respostas obtidas na ferramenta, também foram analisados elementos e técnicas de recuperação de informação identificadas no uso da ferramenta. O Quadro 2 apresenta o resumo das análises das respostas, em cada um dos rascunhos apresentados pela ferramenta.

Quadro 2 - Análise das respostas no Bard

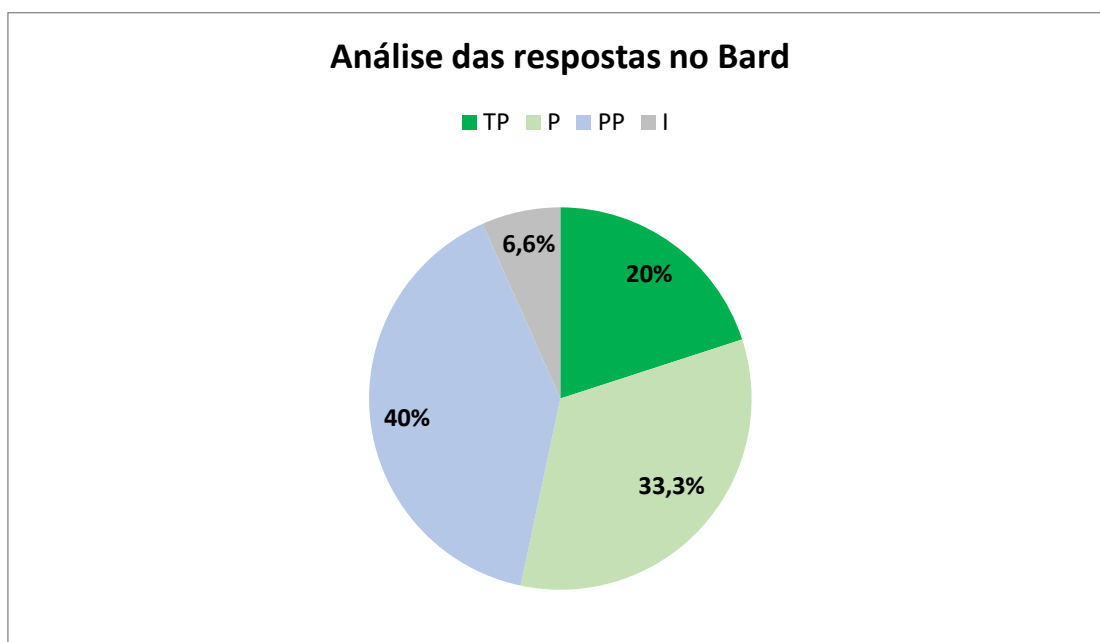
Expressão de busca (Pergunta)	Rascunho 1 (R1)	Rascunho 2 (R2)	Rascunho 3 (R3)
(P1) O que são dados abertos?	TP	TP	TP
(P2) Apresente exemplos de dados abertos.	P	P	P
(P3) Exiba dados abertos disponibilizados no Brasil.	PP	PP	PP
(P4) Quais são os dados abertos sobre vacinação no Brasil?	PP	PP	PP
(P5) Onde acessar dados abertos sobre vacinação no Brasil?	PP	P	PP
(P6) Exiba arquivos de dados abertos sobre a vacinação no Brasil.	PP	I	PP
(P7) Quais ferramentas disponíveis para acesso, uso e reúso dos dados abertos?	P	P	P
(P8) Quais são as APIs de dados abertos disponíveis no Brasil?	PP	PP	I
(P9) Como acessar as APIs de dados abertos?	P	P	P
(P10) Exiba um código para coleta de dados abertos	TP	TP	TP

Fonte: Dados da pesquisa (2023).

As respostas obtidas nos três rascunhos foram similares, mas não idênticas, cumprindo com a proposta da ferramenta de não configurar plágio. No entanto, em um resultado (Pergunta P6) o rascunho 2 destoou completamente das outras respostas. A ferramenta Bard não indicou as fontes das informações apresentadas nas respostas, assim a validação realizada foi baseada exclusivamente no conhecimento da autora, fundamentada nas experiências práticas e na literatura da área.

Foram identificadas duas respostas incorretas, o que compreendeu 6,6% da amostra. Seis respostas (20%) foram classificadas como totalmente pertinentes, por atenderem a necessidade de informação e com a tipologia esperada, 10 respostas (33,3%) foram consideradas pertinentes, ou seja, apresentaram resultados corretos, mas não no formato esperado, outras 12 respostas (40%) foram consideradas parcialmente pertinentes, pois apresentavam resultados parciais em relação ao formato e conteúdo esperado. Os resultados das 30 respostas (10 perguntas com três rascunhos) são apresentados no Gráfico 1:

Gráfico 1 – Análise das respostas na ferramenta Bard



Fonte: Dados da pesquisa (2023).

Para a pergunta P1, os três rascunhos foram totalmente pertinentes e, além do conceito de dados abertos, os resultados apresentaram no texto informações sobre vantagens, formatos e aplicações dos dados abertos, ou seja, informações além do que foi indicado na pergunta. Assim, o usuário obteve a informação desejada, de forma completa, e por meio de um texto coeso, conforme a necessidade de informação indicada.

Uma busca simples em um motor de busca na Web retornaria com vários resultados de diferentes fontes. O Bard reuniu as informações de forma coesa e objetiva para o usuário. No entanto, a não indicação da fonte das

informações é muito preocupante, visto que não é possível validar ou checar a origem. Desta forma, a relevância dos resultados, que envolve questões humanas e técnicas (Monteiro *et al.*, 2017), pode ser comprometida, visto que além da similaridade entre a representação da expressão de busca e a representação dos documentos do corpus pesquisado é preciso identificar a genuinidade do documento recuperado.

Na pergunta P2, os resultados recuperados no Bard são adequados à necessidade informação e alguns exemplos práticos são apresentados, mas de forma generalista. O conceito foi apresentado novamente, como indicado em P1, mas há um destaque sobre como identificar e utilizar os dados abertos:

Os dados abertos são importantes porque podem ser usados para melhorar a vida das pessoas de diversas maneiras. Por exemplo, os dados climáticos podem ser usados para prever o clima, o que pode ajudar as pessoas a se protegerem de desastres naturais. **Os dados de trânsito podem ser usados para melhorar o planejamento urbano**, o que pode tornar as cidades mais eficientes e agradáveis para se viver. **Os dados de saúde podem ser usados para melhorar o atendimento médico, o que pode salvar vidas.** Os dados educacionais podem ser usados para melhorar o ensino, o que pode dar às pessoas mais oportunidades de sucesso na vida. Os dados financeiros podem ser usados para melhorar a tomada de decisões, o que pode ajudar as empresas a crescer e prosperar (P2 – R2, grifo nosso).

Assim, a resposta foi considerada como pertinente, por recuperar alguns exemplos. Pelo fato de não apresentar exemplos específicos, somente temáticas de conjuntos de dados genéricos, não atendeu totalmente a necessidade de informação. A expressão de busca é a forma que o usuário emprega para comunicar a necessidade informacional para o sistema de informação. Para obter sucesso, os usuários precisam ter um mínimo de conhecimento do tema de interesse e do seu vocabulário de domínio (Ferneda, 2012). Quando a ferramenta utiliza a linguagem natural, em um idioma específico, esse processo pode ser mais complexo para a máquina e o conhecimento do usuário é fundamental para avaliar a recuperação de informação.

Em P3 não foram apresentados excertos de dados ou conjuntos de dados abertos, como esperado e modelado na expressão de busca. Foram identificados os nomes dos principais portais, no entanto sem o *link*,

contrariando a proposta indicada pela ferramenta:

Aqui estão alguns links para sites que disponibilizam dados abertos no Brasil:

- Portal de Dados Abertos do Governo Federal
- Portal de Dados Abertos do Estado de São Paulo
- Portal de Dados Abertos do Estado do Rio de Janeiro
- Portal de Dados Abertos do Estado do Paraná
- Portal de Dados Abertos do Estado do Rio Grande do Sul (P3 – R3, grifo nosso).

A indicação de um recurso (*link*) sem a sua apresentação, com exemplos generalistas, que indicam o nome de um Portal, foi considerada como parcialmente pertinente, uma vez que não atendeu a demanda do usuário nem o formato indicado pela própria ferramenta. A apresentação do resultado da busca, por meio de *links*, garante a rastreabilidade e a visualização de relacionamentos. Os *hyperlinks* permitem que usuários possam, enquanto navegam em um artefato, acessar outros artefatos facilmente através de *links* e conecta conceitos relacionados, palavras-chave ou frase de forma natural (Li; Maalej, 2012).

Ao perguntar o Bard sobre um conjunto de dados específico, em P4, foram indicados os nomes dos conjuntos de dados e o órgão que os disponibilizam, mas sem nenhum *link* indicando diretamente a localização dos dados:

O Ministério da Saúde do Brasil disponibiliza uma série de dados abertos sobre vacinação no país. Esses dados incluem informações sobre a cobertura vacinal, o número de doses aplicadas, e as doenças contra as quais as pessoas são vacinadas. **Os dados estão disponíveis no site do Ministério da Saúde**, e também podem ser acessados por meio de APIs.

Aqui estão alguns dos dados abertos sobre vacinação no Brasil:

- **Cobertura vacinal:** este dado mostra o percentual de pessoas que receberam todas as doses de uma determinada vacina.
- **Número de doses aplicadas:** este dado mostra o número total de doses de uma determinada vacina que foram aplicadas no país.
- **Doenças contra as quais as pessoas são vacinadas:** este dado mostra as doenças contra as quais as pessoas são vacinadas no Brasil (P4 – R1, grifo nosso).

Ao formular uma nova pergunta, após P4, a fim de completar a pergunta anterior: “Em quais *links* estão disponíveis?”, a seguinte resposta foi informada: “Sou uma IA com base em texto, e isso está além das minhas capacidades.” Ao insistir e perguntar sobre a fonte dos dados: “Qual fonte utilizou?”, a seguinte resposta foi apresentada: “Fui treinado em um enorme conjunto de dados de texto e código, que inclui textos de livros, artigos, código e outros formatos. Posso acessar e processar informações do mundo real por meio da Pesquisa Google e manter minha resposta consistente com os resultados da pesquisa”, indicando a limitação da ferramenta, principalmente com relação à rastreabilidade da informação apresentada aos usuários. Desta forma, a resposta foi considerada como parcialmente pertinente.

Na pergunta P5, os repositórios com conjuntos de dados foram indicados indiretamente no primeiro e terceiro rascunho, que indicou os menus e opções dentro do site, e diretamente, por meio de *links*, no segundo rascunho:

Para acessar os dados, basta visitar o site do Ministério da Saúde e clicar no link "OpenDataSUS". Em seguida, selecione o tópico "Vacinação" e clique no link "Dados Abertos". Você será direcionado para uma página que lista todos os conjuntos de dados disponíveis sobre vacinação (P5 – R1, grifo nosso).

Para acessar os dados sobre vacinação no portal OpenDataSUS, basta seguir os seguintes passos:

1. Acesse o portal OpenDataSUS.
2. No menu superior, clique em "Dados Abertos".
3. Na página de "Dados Abertos", clique em "Vacinação".
4. Na página de "Vacinação", você encontrará uma lista de conjuntos de dados sobre vacinação.
5. Para visualizar um conjunto de dados, clique no link do conjunto de dados.
6. Na página do conjunto de dados, você encontrará informações sobre o conjunto de dados, como a descrição do conjunto de dados, a data de atualização do conjunto de dados e o formato do conjunto de dados.
7. Para baixar um conjunto de dados, clique no link "Baixar" (P5 – R3).

Aqui estão alguns links para os sites que oferecem dados abertos sobre vacinação no Brasil:

- OpenDataSUS: <https://opendatasus.saude.gov.br/>
 - DataSUS: <https://datasus.saude.gov.br/>
 - IBGE: <https://www.ibge.gov.br/>
 - Fundação Oswaldo Cruz: <https://www.fiocruz.br/>
 - Universidade Federal do Rio de Janeiro: <https://www.ufrj.br/>
- (P5 – R2, grifo nosso).

As respostas dos rascunhos se complementam, mas uma única resposta com a indicação dos *links*, como informado no rascunho 2, apesar de pertinente ainda precisa de complementação. As respostas com a indicação dos menus são parcialmente pertinentes, pois indicam o caminho sem o endereço exato, podendo comprometer a precisão da busca.

A métrica de avaliação de recuperação denominada precisão expressa o quanto o sistema é capaz de recuperar apenas itens relevantes, considerando a proporção entre documentos recuperados e os documentos considerados relevantes ou úteis para o usuário (Baeza-Yates; Ribeiro-Neto, 2013). Nas respostas obtidas em P5, muitos locais foram informados, mas poucos especificavam a localização exata ou endereço direto de algum conjunto de dados.

Em um sistema generativo de IA, o resultado da busca é um texto coeso, conforme a norma culta da língua utilizada. Métricas como o *PageRank*, desenvolvida pela Google, e algoritmos de ranqueamento, utilizados em vários motores de busca, não são visualizados pelos usuários. Com isso, ao oferecer diferentes rascunhos para uma mesma resposta, a ferramenta pode confundir o usuário. Neste contexto, em que a similaridade do resultado não é indicada para o usuário, a medida *PageRank*, que envolve a recuperação de documentos por popularidade, qualidade e quantidade de links, abstraindo as palavras-chave da consulta (Silberschatz *et al.*, 2006), poderia ser um direcionamento para o usuário que não tem um conhecimento aprofundado do domínio e precisa classificar as respostas recuperadas.

Na pergunta P6, o Bard não exibiu arquivos, nem parte do conteúdo de arquivos ou conjuntos de dados abertos. Foram indicados somente portais onde os arquivos poderiam ser acessados, similar às respostas apresentadas em P4. O Rascunho 2 retornou com a frase: “Desculpe, não tenho acesso a esses dados” e foi considerada uma resposta incorreta para a pergunta realizada. Os rascunhos 1 e 3 indicaram que iriam apresentar os arquivos, mas somente informaram o nome do conjunto de dados ou o nome do órgão responsável pela disponibilização, demonstrando um atendimento parcial à pergunta P6:

Claro. **Aqui estão alguns arquivos de dados abertos** sobre a vacinação no Brasil:

Dados abertos sobre a vacinação contra a COVID-19 no Brasil - Este conjunto de dados inclui informações sobre as doses de vacinas contra a COVID-19 aplicadas no Brasil, por estado e município.

Dados abertos sobre a vacinação em massa no Brasil - Este conjunto de dados inclui informações sobre as campanhas de vacinação em massa no Brasil, por estado e município.

Dados abertos sobre a vacinação infantil no Brasil - Este conjunto de dados inclui informações sobre as vacinas infantis aplicadas no Brasil, por estado e município.

Você pode encontrar esses conjuntos de dados no portal de dados abertos do Ministério da Saúde. Para acessar o portal, visite o seguinte endereço: <https://opendatasus.saude.gov.br/> (P6 – R1, grifo nosso).

Para a pergunta P7, a ferramenta Bard indicou nos rascunhos 1, 2 e 3 o nome e uma descrição de ferramentas para análise de dados, conforme a necessidade de informação da busca, contudo foi apresentada uma lista bem generalista no rascunho 1 e 3 e outra mais detalhada no rascunho 2. Destaca-se que nenhum rascunho exibiu ferramentas específicas ou mais utilizadas para análise e tratamento de dados abertos, nem apresentou exemplos práticos relacionados, uma otimização interessante de ser apresentada ao usuário, sendo a resposta considerada pertinente:

Existem muitas ferramentas disponíveis para acesso, uso e reúso dos dados abertos. Algumas das ferramentas mais populares incluem:

- **Portais de dados abertos:** Esses portais fornecem acesso a um grande conjunto de dados abertos de várias fontes. Alguns exemplos de portais de dados abertos incluem o Data.gov dos Estados Unidos e o OpenData.cat da Catalunha.
- **Plataformas de análise de dados:** Essas plataformas permitem que os usuários analisem dados abertos usando ferramentas de visualização e exploração de dados. Alguns exemplos de plataformas de análise de dados incluem o Google BigQuery e o Microsoft Azure Databricks.
- **Bibliotecas de programação:** Essas bibliotecas fornecem aos desenvolvedores acesso a dados abertos a partir de suas aplicações. Alguns exemplos de bibliotecas de programação para dados abertos incluem o Pandas e o NumPy.
- **Ferramentas de código aberto:** Essas ferramentas são desenvolvidas e mantidas por uma comunidade de desenvolvedores. Alguns exemplos de ferramentas de código

aberto para dados abertos incluem o OpenStreetMap e o Wikipedia (P7 – R1, grifo nosso).

Existem muitas ferramentas disponíveis para acesso, uso e reúso de dados abertos. Aqui estão algumas das mais populares:

- **OpenRefine**: Uma ferramenta gratuita e de código aberto para limpar, enriquecer e combinar dados.
- **OpenOffice.org LibreOffice**: Um pacote de escritório gratuito e de código aberto que inclui um processador de texto, planilha, apresentação e software de banco de dados.
- **R**: Uma linguagem de programação e ambiente de computação estatística gratuita e de código aberto.
- **Python**: Uma linguagem de programação de propósito geral gratuita e de código aberto que é frequentemente usada para análise de dados.
- **Hadoop**: Um sistema de processamento distribuído gratuito e de código aberto para armazenamento e análise de grandes conjuntos de dados.
- **Hive**: Um sistema de banco de dados relacional gratuito e de código aberto para o Hadoop.
- **Pig**: Um sistema de programação de alto nível gratuito e de código aberto para o Hadoop.
- **Spark**: Um ambiente de computação em cluster gratuito e de código aberto para processamento de dados em tempo real.
- **Cassandra**: Um banco de dados não relacional gratuito e de código aberto para armazenamento e análise de grandes conjuntos de dados.
- **MongoDB**: Um banco de dados não relacional gratuito e de código aberto para armazenamento e análise de grandes conjuntos de dados (P7 - R2, grifo nosso).

Para a pergunta P8, foram indicados nos rascunhos 1 e 2 Application Programming Interface (API)s que disponibilizam dados abertos, mas em nenhum rascunho a API foi indicada com o *link* direto para acesso, somente informações sobre o serviço e o nome do órgão:

Existem muitas APIs de dados abertos disponíveis no Brasil. Algumas das mais populares incluem:

- **Dados.gov.br** - Esta API fornece acesso a uma variedade de dados governamentais abertos, incluindo dados sobre clima, saúde, educação e economia.
- **Câmara dos Deputados** - Esta API fornece acesso a dados sobre os deputados, suas votações e seus projetos de lei.
- **Senado Federal** - Esta API fornece acesso a dados sobre os senadores, suas votações e seus projetos de lei.
- **Tribunal Superior Eleitoral** - Esta API fornece acesso a dados sobre as eleições, incluindo os candidatos, os votos e os resultados.
- **Ministério da Saúde** - Esta API fornece acesso a dados sobre a saúde pública, incluindo dados sobre doenças, vacinas

e hospitais.

- **Ministério da Educação** - Esta API fornece acesso a dados sobre a educação, incluindo dados sobre escolas, alunos e professores.

- **Ministério do Trabalho** - Esta API fornece acesso a dados sobre o trabalho, incluindo dados sobre empregos, salários e condições de trabalho.

- **Ministério da Fazenda** - Esta API fornece acesso a dados sobre a economia, incluindo dados sobre PIB, inflação e taxas de juros.

- **Ministério do Meio Ambiente** - Esta API fornece acesso a dados sobre o meio ambiente, incluindo dados sobre o clima, a biodiversidade e a poluição.

Essas são apenas algumas das muitas APIs de dados abertos disponíveis no Brasil. Para obter mais informações, visite o site do Portal de Dados Abertos do Brasil (P8 - R2, grifo nosso).

Nessa resposta foi observado um problema conceitual, visto que a ferramenta informa órgãos que não possuem uma API própria ou dedicada, como é o caso do Ministério da Educação (P8 - R2), que possui somente a interface padrão do Portal de Dados Abertos. Também foram indicados alguns Portais de Dados Abertos (P8 - R3), mas essa informação foi considerada incorreta, visto que o acesso às APIs dos Portais não é detalhado e alguns não permitem o acesso aberto a qualquer usuário. A API do Portal de Dados Abertos do Governo Federal, por exemplo, estava indisponível desde a última migração, em novembro de 2022, e em versões anteriores permitia somente o acesso à publicadores, ou seja, ou órgãos do Governo Federal:

Existem muitas APIs de dados abertos disponíveis no Brasil. Algumas das mais populares incluem:

- Portal de Dados Abertos do Governo Federal
- Portal de Dados Abertos do Estado de São Paulo
- Portal de Dados Abertos do Estado do Rio de Janeiro
- Portal de Dados Abertos da Prefeitura de São Paulo
- Portal de Dados Abertos da Prefeitura do Rio de Janeiro (P8 - R3, grifo nosso).

Os algoritmos utilizados pelos serviços de busca procuram o aperfeiçoamento contínuo nos processos de indexação e busca, mas desde o início dos mecanismos, no final da década de 1990, os engenheiros debruçaram-se sobre o aspecto da relevância. Esta envolve palavras semanticamente significativas para a busca e são analisadas em conjunto com diferentes variáveis que incluem o significado, qualidade, usabilidade, contexto,

além de considerar aspectos como: frequência, data, localização, análise de *links*, dentre outros elementos (Google, 2023c; Monteiro *et al.*, 2017).

A recuperação de dados por meio de APIs é uma forma de acesso necessária em contextos de *big data*, que envolve a disponibilização e coleta de grandes conjuntos de dados (Silva, 2023). A API faz a ligação entre sistemas permitindo a comunicação, a interoperabilidade, sem a necessidade de conhecimento das ferramentas e códigos implementados nos sistemas. Uma vantagem dessa arquitetura é apresentar rotas e parâmetros definidos para consulta dos usuários, controlando o vocabulário e aumentando a precisão dos resultados, ou seja, o quanto o sistema é capaz de recuperar apenas os itens relevantes. Em consequência, a revocação diminui. A revocação é a fração dos documentos relevantes que foi recuperada, ela expressa o quanto o sistema é capaz de recuperar todos os itens relevantes (Baeza-Yates; Ribeiro-Neto, 2013; Lancaster, 2004).

Em uma API, a recuperação de informação estará totalmente relacionada à rota, ou caminho projetado, e os respectivos parâmetros modelados. Uma necessidade de informação não modelada impacta diretamente na avaliação da recuperação pelo usuário, que a julgará como insatisfatória, por não atender sua demanda.

Nas perguntas P9 e P10 eram esperadas respostas com exemplos práticos. Para P9 foi apresentada uma lista com etapas descritivas sobre como utilizar uma API. Foram indicados recursos comuns para esse tipo de acesso, de forma detalhada e explicativa, para o usuário:

Aqui estão algumas etapas para acessar APIs de dados abertos:

1. Encontre uma API de dados abertos que atenda às suas necessidades.
2. Visite o site da API e registre-se para uma conta.
3. Obtenha uma chave de API do site da API.
4. Use uma linguagem de programação ou ferramenta de terceiros para acessar os dados da API (P9- R1, grifo nosso).

Aqui estão alguns exemplos de como você pode usar as APIs de dados abertos:

- Você pode criar um aplicativo que mostre a previsão do tempo para a sua área.
- Você pode criar um aplicativo que mostre o tráfego nas ruas próximas.
- Você pode criar um aplicativo que mostre as notícias mais

recentes do governo.

- Você pode criar um aplicativo que mostre os dados mais recentes sobre o clima.
- Você pode criar um aplicativo que mostre os dados mais recentes sobre o tráfego (P9- R2, grifo nosso).

Apesar de atender a necessidade informacional, esperava-se um formato que permitisse visualizar, em detalhes, a implementação.

Em P10 foi apresentado um texto e um código-fonte (Figura 3), conforme sugerido na expressão formulada. Apesar de ser representado por um texto, o código-fonte pode ser considerado um objeto digital, pois é executável em um interpretador da linguagem Python. Nos três rascunhos foram apresentados de códigos-fonte para acesso a uma API de exemplo:

1. <https://www.data.gov/> (Rascunho 1 - Figura 3)
2. https://www.opendata.org/api/3/action/package_list (Rascunho 2)
3. <https://www.opendata.socrata.com/> (Rascunho 3)

Figura 3 – Código fonte retornado pela ferramenta Bard

```
Python

import requests
from bs4 import BeautifulSoup

# Definir a URL do site de dados abertos
url = "https://www.data.gov/"

# Fazer uma solicitação HTTP para o site
response = requests.get(url)

# Verificar se a solicitação foi bem-sucedida
if response.status_code == 200:

    # Processar a resposta HTTP
    soup = BeautifulSoup(response.content, "lxml")

    # Extrair os links para conjuntos de dados
    links = soup.find_all("a", href=True)

    # Iterar sobre os links
    for link in links:

        # Extrair o título do conjunto de dados
        title = link.get("title")

        # Extrair a descrição do conjunto de dados
        description = link.get("description")

        # Extrair o link para o conjunto de dados
        data_link = link["href"]

        # Imprimir as informações do conjunto de dados
        print("Título:", title)
        print("Descrição:", description)
        print("Link:", data_link)

else:

    print("Ocorreu um erro ao fazer a solicitação HTTP.")

Use o código com cuidado. Saiba mais
```

Este código primeiro define a URL do site de dados abertos. Em seguida, faz uma solicitação HTTP para o site e verifica se a solicitação foi bem-sucedida. Se a solicitação for bem-sucedida, o código processa a resposta HTTP e extrai os links para conjuntos de dados. O código então itera sobre os links e imprime as informações do conjunto de dados, incluindo o título, a descrição e o link para o conjunto de dados.

Use o código com cuidado. Saiba mais

Este código primeiro define a URL do site de dados abertos. Em seguida, faz uma solicitação HTTP para o site e verifica se a solicitação foi bem-sucedida. Se a solicitação for bem-sucedida, o código processa a resposta HTTP e extrai os links para conjuntos de dados. O código então itera sobre os links e imprime as informações do conjunto de dados, incluindo o título, a descrição e o link para o conjunto de dados.

Fonte: Google (2023b).

As respostas foram consideradas pertinentes (P9) e totalmente pertinentes (P10). Para o código apresentado na pergunta 10 foi possível verificar que há comentários explicando, passo a passo, todas as ações implementadas, uma boa prática utilizada no desenvolvimento de software. Além disso, é importante destacar que a pergunta era ampla e o código apresentado também foi genérico, permitindo sua adaptação.

Os modelos de sistemas de recuperação de informação já consideravam a recuperação de documentos baseados em *links* e objetos multimídia, no entanto, em modelos distintos e não unificados. Os modelos clássicos (booleano, vetorial e probabilístico) se concentravam em recuperar documentos do tipo texto e outros modelos tratavam os documentos baseados em *link*, relacionados à Web, e outros a recuperação multimídia, envolvendo documento de áudio e vídeo (Baeza-Yates; Ribeiro-Neto, 2013).

Observou-se que o modelo utilizado no Bard considera vários tipos de documentos, porém os resultados obtidos foram predominantemente textuais. A ferramenta Bard retornou rascunhos com texto em todas as perguntas e um código fonte, em uma pergunta, que foi considerado um objeto digital.

Como pontos de atenção do Bard destacam-se: (1) apresentar poucos *links* diretos nas respostas e (2) não informar a fonte das informações. O primeiro ponto pode prejudicar a interação com o usuário que recebe uma resposta generalista, sem a especificação necessária, e o segundo ponto tem relação direta com a confiabilidade, pois sem a fonte será necessário validar o conteúdo para certificar que a informação está correta ou “confiar” na ferramenta.

5 CONSIDERAÇÕES FINAIS

Este relato de experiência objetivou analisar a recuperação de informação na Web com a ferramenta Bard que utiliza a IA generativa e a linguagem natural para “interagir” com os usuários. Critérios foram definidos para formulação de perguntas inseridas na ferramenta e a análise pautou-se nas experiências da autora, já que as fontes das informações não foram informadas pela ferramenta.

A recuperação de informações no Bard foi realizada com sucesso, em três rascunhos diferentes, porém com semelhança no conteúdo apresentado. As respostas obtidas, em sua maioria (40%) atenderam parcialmente as perguntas ou eram pertinentes (33,3%), contudo melhorias podem ser visualizadas, uma vez que algumas respostas são bem generalistas e não apresentam detalhes ou exemplos, mesmo quando solicitado por meio da expressão de busca ou pergunta do usuário.

A apresentação das informações, em formato textual, utilizou listas com marcadores, em um texto coeso, sem a indicação de ranqueamento ou grau de similaridade como é geralmente apresentado nos motores de busca da Web. A falta desse tipo de métrica pode confundir o usuário quando o mesmo não tem conhecimento do domínio pesquisado.

A recuperação de diferentes recursos e objetos digitais foi identificada em uma das perguntas, onde a ferramenta retornou um código-fonte na linguagem de programação Python. Contudo, é importante destacar que a recuperação foi possível somente utilizando a expressão de busca (pergunta) com o nome do objeto: código. Expressões onde foi utilizado o termo arquivo ou conjunto de dados, por exemplo, não recuperaram os formatos ou objetos esperados. O termo “exiba” também foi utilizado em algumas expressões, porém nenhuma imagem ou vídeo relacionado foi recuperado.

Apesar de algumas limitações e inconsistências, para o contexto pesquisado, a ferramenta tem seus méritos e sua evolução possivelmente irá adequar e otimizar os resultados. Para um usuário que não possui conhecimento sobre o domínio dados abertos, as informações apresentadas são básicas e podem atender as necessidades informacionais. Mas para um especialista, muitos detalhes são omitidos e informações inconsistentes foram identificadas.

Destaca-se que em dezembro de 2023, cinco meses após do desenvolvimento deste estudo, a Google anunciou o lançamento do Gemini, novo modelo de linguagem de IA generativa que substitui o Bard e é capaz de compreender também imagens e vídeos. Em 2024 a ferramenta foi disponibilizada no Brasil e permanece em operação. Como estudos futuros, pretende-se acompanhar a evolução da ferramenta e realizar estudos comparativos com modelos similares que tenham o mesmo propósito.

REFERÊNCIAS

- BAEZA-YATES, R.; RIBEIRO-NETO, B. **Recuperação de informação: conceitos e tecnologia das máquinas de busca**. 2. ed. Porto Alegre: Bookman, 2013.
- CASTOR, E. C. S.; FERNANDES, A. L.; MOTTA, A. C. G.; GARCIA, R. B. .; LIMA, A. F. Chatbot: impactos no ambiente acadêmico de uma universidade do Rio de Janeiro. **P2P E INOVAÇÃO**, Rio de Janeiro, v. 8, n. 1, p. 71–92, 2021. Disponível em: <https://revista.ibict.br/p2p/article/view/5760>. Acesso em: 26 jul. 2023.
- GOOGLE. Perguntas frequentes sobre o Bard. 2023a. Disponível em: <https://bard.google.com/faq>. Acesso em: 19 jul. 2023.
- GOOGLE. Bard. 2023b. Disponível em: <https://bard.google.com/>. Acesso em: 19 jul. 2023.
- GOOGLE. Como a Busca do Google funciona. 2023c. Disponível em: <https://www.google.com/intl/pt-BR/search/howsearchworks/how-search-works/>. Acesso em: 19 jul. 2023.
- FERNEDA, E. **Recuperação de informação: análise sobre a contribuição da ciência da computação para a ciência da informação**. 2003. Tese (Doutorado em Ciência da Informação e Documentação) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2003. DOI:10.11606/T.27.2003.tde-15032004-130230. Disponível em: <https://www.teses.usp.br/teses/disponiveis/27/27143/tde-15032004-130230/pt-br.php>. Acesso em: 19 jul. 2023.
- FERNEDA, E. **Introdução aos Modelos Computacionais de Recuperação de Informação**. Rio de Janeiro: Editora Ciência Moderna Ltda. 2012.
- LANCASTER, F. W. **Indexação e resumos: teoria e prática**. 2. ed. Brasília: Briquet de Lemos, 2004.
- LEWANDOWSKI, D. New perspectives on web search engine research. In: LEWANDOWSKI, D. **Web search engine research**. 4. ed. Bingley: Emerald Group Publishing Limited, 2012. p. 1-16.
- LI, Y.; MAALEJ, W. Which Traceability Visualization Is Suitable in This Context? A Comparative Study. (B. Regnell, D. Damian, Eds.) Requirements Engineering: Foundation for Software Quality. **Anais [...]** Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.
- MONTEIRO, D.; FERNANDES, P. M.; DECARLI, C.; TREVISAN, L. Sistemas de recuperação da informação e o conceito de relevância nos mecanismos de busca: semântica e significação. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, [S. l.], v. 22, n. 50, p. 161–175,

2017. DOI: 10.5007/1518-2924.2017v22n50p161. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2017v22n50p161>. Acesso em: 27 jul. 2023.

MUSSI, R. F. de F.; FLORES, F. F.; ALMEIDA, C. B. de. Pressupostos para a elaboração de relato de experiência como conhecimento científico. **Práxis Educacional**, Vitória da Conquista, v. 17, n. 48, p. 60-77, 2021. Disponível em: <https://periodicos2.uesb.br/index.php/praxis/article/view/9010>. Acesso em: 28 jul. 2023.

NASCIMENTO SILVA, P. Recuperação de informação na Ciência da Informação: produção acadêmico-científica brasileira (2012-2021). **Transinformação**, v. 35, p. e237336, 2023a. DOI: 10.1590/2318-0889202335e237336. Disponível em: <https://www.scielo.br/j/tinf/a/7b74SBZHTWBjVY4SVhy8PSP/?lang=pt>. Acesso em: 19 ago. 2024.

NASCIMENTO SILVA, P. Recuperação de informação na Web: uma experiência com o modelo de linguagem de inteligência artificial ChatGPT. **Informação em Pauta**, v. 8, p. 1–19, 2023b. DOI: 10.36517/2525-3468.ip.v8i0.2023.83566.1-19. Disponível em: <http://www.periodicos.ufc.br/informacaoempauta/article/view/83566>. Acesso em: 19 ago. 2024.

NASCIMENTO SILVA, P. Dados de pesquisa no Google Bard 2023. Mendeley Data, V1, 2024. DOI: 10.17632/9bbnvjxggm.1. Disponível em: <https://data.mendeley.com/datasets/9bbnvjxggm/1>. Acesso em 19 ago. 2024.

QUEIROZ, T. dos S.; VALLS, V. M. O bibliotecário analista de chatbot: as competências desenvolvidas nos cursos presenciais de bacharelado em biblioteconomia da cidade de São Paulo. **Revista Brasileira de Biblioteconomia e Documentação**, [S. l.], v. 18, p. 1–25, 2022. Disponível em: <https://rbbd.febab.org.br/rbbd/article/view/1534>. Acesso em: 26 jul. 2023.

SARACEVIC, T. Ciência da informação: origem, evolução e relações. **Perspectiva em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235/22>. Acesso em: 19 abril 2023.

SILBERSCHATZ, A. et al. **Sistema de Banco de Dados**. 5. ed. Rio de Janeiro: Elsevier, 2006.

SILVA, P. N. Recuperação de dados em apis: uma experiência prática no twitter. *In*: XXII ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 2022, Porto Alegre. **Anais Eletrônicos** [...]. Porto Alegre: UFRGS, 2022. Disponível em: <https://enancib.ancib.org/index.php/enancib/xxiiencib/paper/viewFile/1007/728>. Acesso em: 19 jul. 2023.

TRIPP, D. Pesquisa-ação: uma introdução metodológica. **Educação e Pesquisa**, v. 31, n. 3, p. 443–466, Sep. 2005.



Este obra está licenciado com uma Licença Creative Commons Atribuição-
NãoComercial 4.0 Internacional.