

## Retrieval-Augmented Generation na Ciência da Informação: o que precisamos saber?

Patrícia Nascimento Silva<sup>1</sup>, Daiane Campos Procópio<sup>2</sup>, Danielle Teixeira de Oliveira<sup>3</sup>

Universidade Federal de Minas, Minas Gerais, Brasil

patricians@ufmg.br, campos-daiane@ufmg.br, danielle-oliveira@ufmg.br

**Resumo.** *A Ciência da Informação é uma área que, naturalmente, investiga a informação, ou seja, dados contextualizados, seus fenômenos e interações, bem como seus impactos na sociedade. Com o aumento significativo dos dados em ambientes digitais e a rápida evolução da Inteligência Artificial Generativa (IAG), essas fontes ganham um novo destaque, tendo em vista que constituem matérias-primas para modelos de IAG. No entanto, identificar como esses dados e informações serão coletados, tratados e organizados para serem utilizados e recuperados nos modelos é o questionamento central deste estudo, que investiga a geração aumentada de recuperação, ou o Retrieval-Augmented Generation (RAG). Esta pesquisa objetivou investigar como a técnica RAG tem sido implementada e utilizada na atualidade, sob a luz da organização e da representação da informação, em 2026, por meio de uma revisão de literatura em bases acadêmico-científicas. Foi observado que o RAG é uma técnica muito utilizada por diversos domínios; contudo, na Ciência da Informação, ela ainda é pouco difundida, apesar dos fundamentos da área de organização, representação e recuperação da informação estarem alinhados à técnica, o que justifica um novo olhar para a área para se aproximar de técnicas da Computação, relacionadas à Recuperação da Informação.*

---

<sup>1</sup> Professora Adjunta na Escola de Ciência da Informação (ECI) da Universidade Federal de Minas Gerais (UFMG). Professora e Pesquisadora no Programa de Pós-Graduação em Gestão & Organização do Conhecimento (PPGGOC) ECI/UFMG. Universidade Federal de Minas Gerais. patricians@ufmg.br. <https://orcid.org/0000-0002-2405-8536>

<sup>2</sup> Doutoranda e mestra (2025) em Gestão e Organização do Conhecimento pela ECI/UFMG. Graduada em Biblioteconomia pela UFMG (2013) e Análise e Desenvolvimento de Sistemas pela PUC Minas (2024). campos-daiane@ufmg.br. Universidade Federal de Minas Gerais. <https://orcid.org/0000-0002-9006-191X>

<sup>3</sup> Doutoranda e mestra (2025) em Gestão e Organização do Conhecimento pela ECI/UFMG. Coordenadora da Biblioteca da Faculdade de Educação da UFMG. Universidade Federal de Minas Gerais. danielle-oliveira@ufmg.br. <https://orcid.org/0000-0002-1958-9113>

## 1. Introdução

A Ciência da Informação, enquanto área do conhecimento, é uma ciência relativamente recente, nascendo em um período histórico de críticas ao fracasso do projeto da modernidade e ao seu modelo científico (Araújo, 2003).

No entanto, sua trajetória remonta à documentação, termo cunhado por Paul Otlet, em 1903, para designar o processo de fornecimento de documentos para a busca de informação (Alvares & Araújo Júnior, 2010). Nesse período, a área era definida como Ciência da Documentação por Otlet (Otlet, 1934), até que novas discussões em alguns eventos de informação científica mudaram essa trajetória. O destaque foi para a Conferência Internacional de Informação Científica, em 1958, na qual Robert S. Taylor apresentou um trabalho sobre o glossário de termos da documentação científica, que contribuiu de forma significativa para a definição da área como Ciência da Informação (Queiroz & Moura, 2015).

Entretanto, é na década de 1960 que são discutidos e elaborados os primeiros conceitos e definições da área e se inicia realmente o debate sobre suas origens e seus fundamentos teóricos, na tentativa de demarcar e estabelecer relações interdisciplinares com outros campos do conhecimento, bem como vislumbrar a atuação dos também novos profissionais (Pinheiro & Loureiro, 1995).

A partir da década de 1960, a informação consolidou-se como seu principal objeto de estudo da área, com a discussão de parâmetros classificatórios e teorias significativas para a área, como a Teoria da Análise Facetada (Ranganathan, 1967), a Teoria do Conceito (Dahlberg, 1978) e a Teoria Geral da Terminologia (Wüster, 1931). Tudo isso foi acompanhado das evoluções tecnológicas da época, como a criação da orientação a objetos e dos modelos relacionais, na década de 1970, os primeiros sistemas de informação, em 1980, e a evolução da catalogação, com a criação dos metadados, e a *World Wide Web*, em 1989 (World Wide Web Foundation, [20–]).

Tudo isso permitiu a implementação de ambientes digitais ampliados e cada vez mais complexos, com representações envolvendo a *web* semântica e a interoperabilidade, com os movimentos de acesso aberto nos anos 2000, iniciando pelos dados abertos de pesquisa, seguidos pelos dados abertos de governo e pelas bases generalistas e ampliadas, a exemplo do Wikidata<sup>4</sup>, base aberta com dados de diferentes domínios e processável por máquina.

É a partir dessa época, potencializada nos anos 2010, com os dispositivos móveis e sua crescente popularização, passando os anos 2020, com velocidades de *internet* mais robustas, com o 5G e fenômenos como o *Big Data*, que foram alimentadas

---

<sup>4</sup> <https://www.wikidata.org/>.

transformações digitais significativas, aceleradas pela pandemia de Covid-19 e chegando ao ápice com a Inteligência Artificial Generativa (IAG).

Nesse cenário, observa-se uma ampliação desse escopo da área, incorporando os dados como um elemento central de investigação e atuação. Considerar todas essas evoluções em torno dos dados sem pensar nas transformações da área de Ciência da Informação é inimaginável, pois seus fundamentos orientam todos esses novos processos e revoluções tecnológicas. Assim, em uma sociedade que vivencia a crescente utilização de tecnologias emergentes nas mais diversas áreas e setores, o papel da Ciência da Informação torna-se ainda mais abrangente, demandando uma constante revisão e atualização de seus fundamentos, objetos e práticas.

Todo esse preâmbulo aqui apresentado perpassou a documentação, a informação e agora os dados. Assim, quando se pensa em um contexto lógico, ao longo dos anos, a área aumenta a sua granularidade, a partir de um documento que pode estar associado ao registro de conhecimento, desde a informação até os dados.

Nessa perspectiva, a Retrieval-Augmented Generation (RAG) representa mais um avanço nesse processo de refinamento da informação. Enquanto os Large Language Models (LLMs) tradicionais respondem com base no conhecimento aprendido durante seu treinamento, o RAG permite que essas respostas sejam enriquecidas por informações recuperadas em tempo real de fontes externas, como documentos, bases de dados científicas ou conteúdos disponíveis na *internet*. Ao integrar essas fontes ao processo de geração das respostas, as informações se tornam mais atuais e contextualizadas, favorecendo a construção de conhecimento mais consistente e confiável (Bourne, 2024). A Figura 1 apresenta a lógica de construção do conhecimento contextualizado, a partir de documentos e informações que produzirão novos dados para serem utilizados na RAG.

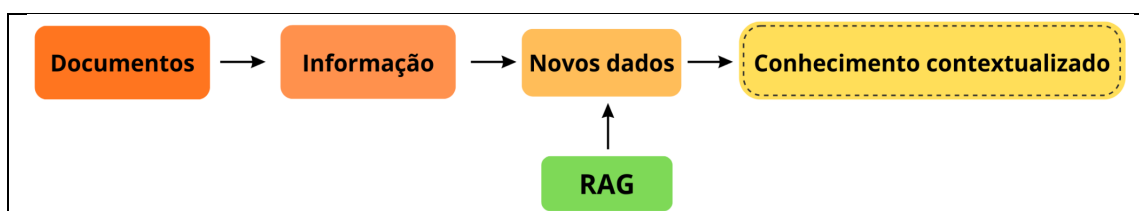


Figura 1 - Representação da RAG<sup>5</sup>

Mas, para além desse contexto, o objetivo da Ciência da Informação continua sendo o mesmo: investigar a origem, a coleção, a organização, o armazenamento, a recuperação, a interpretação, a transmissão, a transformação e a utilização da informação,

<sup>5</sup> Elaborado pelas autoras (2026).

incluindo a pesquisa sobre a representação da informação em sistemas naturais e artificiais, o uso de códigos para a transmissão eficiente da mensagem e o estudo do processamento e de técnicas aplicadas aos computadores e seus sistemas de programação (Borko, 1968, tradução nossa). Conforme afirma Ferneda (2012), identificar, em um acervo documental, os documentos que venham a atender a uma determinada necessidade de informação do usuário constitui o processo de recuperação de informação, essencial em qualquer ambiente informacional.

Assim, esse artigo fomenta a ideia de que os pesquisadores repensem oportunidades e o posicionamento da área, que, apesar de sua importância, ainda não é amplamente visualizada em interseções simbióticas com as novas tecnologias, perdendo espaço nas discussões atuais.

A aplicação da técnica de RAG é recente, datada de 2020 (Lewis *et al.*, 2020), e tem crescido exponencialmente. No entanto, há uma hipótese de que esse crescimento na produção acadêmico-científica é mais percebido em áreas da tecnologia, questionando-se: como o RAG está sendo refletido no contexto da Ciência da Informação? Desta forma, como a área tem se apropriado dessas novas técnicas é o questionamento principal deste estudo, que realiza uma pesquisa bibliográfica com o objetivo de investigar as aplicações e associações ao RAG na área da Ciência da Informação.

Como apontado por Cardoso, Ferneda e Souza (2025), existe uma lacuna na literatura da Ciência da Informação no que tange às análises sistemáticas sobre as aplicações, as potencialidades e os desafios éticos do RAG, e há a necessidade de os estudos explorarem a integração entre tecnologias de RAG e as práticas da Ciência da Informação, considerando suas dimensões sociais e éticas. Assim, essa pesquisa justifica-se por fomentar uma nova discussão, voltada a um dos fundamentos da área, a Recuperação da Informação (RI), considerando as fases da documentação, da informação e, agora, dos dados, como uma oportunidade para ampliar como a Ciência da Informação pode contribuir nos processos informacionais impactados pelas novas tecnologias.

## **2. Evolução da Recuperação da Informação na Ciência da Informação**

A Recuperação da Informação (RI) constitui um campo abrangente, originado na Ciência da Computação, cujo propósito central é possibilitar aos usuários o acesso às informações que atendam às suas necessidades e aos seus interesses (Baeza-Yates & Ribeiro-Neto, 2013). Na perspectiva de Saracevic (1996), ela também pode ser compreendida como uma dimensão tecnológica da Ciência da Informação, desenvolvida a partir da aproximação entre essa área e a Ciência da Computação.

Com o avanço da era digital e a intensificação do uso e da apropriação da informação pelos usuários, a RI passou a incorporar contribuições significativas de ambos os campos. Desse modo, a relação entre a Ciência da Computação e a Ciência da

Informação, inicialmente marcada por conflitos, tornou-se progressivamente mais próxima, simbiótica e indissociável, sobretudo a partir da década de 1990.

Com base nos estudos de Baeza-Yates e Ribeiro-Neto (2013), Luz, Coneglian e Segundo (2019), Procópio (2025), Souza e Rodas (2020), Stock e Stock (2013), em dados do IBGE (2015) e na complementação das autoras, o Quadro 1 reúne, em ordem cronológica, algumas das contribuições mais relevantes para a evolução da área de RI a partir da década de 1950.

**Quadro 1 – Contribuições para a área de RI a partir dos anos 1950<sup>6</sup>**

| Ano  | Pesquisador(es)                                                 | Contribuição                                                                                                                                                                                                                        |
|------|-----------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1951 | Calvin Mooers                                                   | Criação do termo “recuperação da informação”.                                                                                                                                                                                       |
| 1955 | Allen Kent, Madeline M. Berry, Fred U. Luehrs Jr. e J. W. Perry | Publicação do artigo <i>Machine literature searching VIII. Operational criteria for designing information retrieval systems</i> , que descreveu métricas de precisão e revocação para Sistemas de Recuperação de Informação (SRIs). |
| 1962 | Cyril Cleverdon                                                 | Estudos de Cranfield, que avaliaram e compararam SRIs.                                                                                                                                                                              |
| 1963 | Joseph Becker e Robert Hayes                                    | Publicação do primeiro livro sobre RI, intitulado <i>Information storage and retrieval: tools, elements, theories</i> .                                                                                                             |
| 1968 | Gerard Salton                                                   | Publicação do primeiro livro de RI de autoria de Gerard Salton, intitulado <i>Automatic information organization and retrieval</i> .                                                                                                |
| 1971 | N. Jardine e C. J. van Rijsbergen                               | Formulação da hipótese de agrupamento.                                                                                                                                                                                              |
| 1978 | Association for Computing Machinery (ACM)                       | Primeiro congresso da ACM sobre recuperação da informação, ACM SIGIR – <i>Special Interest Group on Information Retrieval</i> , em Rochester, NY, sendo este o principal congresso da área, reunindo pesquisadores do mundo todo.   |
| 1979 | C. J. van Rijsbergen                                            | Publicação do livro <i>Information Retrieval</i> , com abordagens sobre modelos probabilísticos.                                                                                                                                    |

<sup>6</sup> Adaptado de Baeza-Yates e Ribeiro-Neto (2013); Luz, Coneglian e Segundo (2019); Procópio (2025); Souza e Rodas (2020); Stock e Stock (2013); e IBGE (2015).

|                |                                   |                                                                                                                                                           |
|----------------|-----------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1983           | Gerard Salton e Michael J. McGill | Publicação do livro <i>Introduction to Modern Information Retrieval</i> , com enfoque no modelo vetorial.                                                 |
| 1989           | Tim Berners-Lee                   | Criação da <i>World Wide Web</i> , que popularizou o acesso à informação em larga escala, impulsionado por mecanismos de busca como o Yahoo! e AltaVista. |
| 2001           | Tim Berners-Lee                   | Criação da <i>web</i> semântica, que contribuiu para a criação de aplicações mais inteligentes ao aprimorar a compreensão do contexto dos dados.          |
| Década de 2010 | IBGE                              | Popularização dos dispositivos móveis no Brasil, que ampliaram o acesso e a agilidade na busca por conteúdos digitais.                                    |
| Década de 2020 | OpenAI                            | Inteligência Artificial Generativa.                                                                                                                       |

No contexto científico, a pesquisa de Nascimento Silva (2023), que analisou a produção brasileira em RI na Ciência da Informação entre 2012 e 2021, identificou 27 temáticas de pesquisa. Os estudos concentraram-se principalmente em usuários e mecanismos de busca, RI e modelos de recuperação. A partir de 2020, também passaram a emergir temas relacionados à Inteligência Artificial (IA), *chatbots*, *big data*, *data science*, *linked data* e visualização da informação, demonstrando o impacto das transformações tecnológicas na área.

### 3. Metodologia

A pesquisa corresponde a uma revisão da literatura, com natureza descritiva, por buscar expor as particularidades e os atributos de um determinado grupo ou fenômeno, neste caso, a técnica RAG, e exploratória, por visar ao aprofundamento da compreensão do problema investigado, considerando o foco na Ciência da Informação, conferindo-lhe maior definição e precisão.

Dessa forma, realizou-se uma revisão bibliográfica para investigar esse cenário, que forneceu a base teórica necessária para compreender o estado atual do conhecimento sobre o tema (Gil, 2017), conduzida com base em um protocolo de revisão de literatura adaptado de Nascimento Silva (2023), detalhado no Quadro 2, que orientou as três etapas metodológicas.

**Quadro 2 - Protocolo de revisão de literatura<sup>7</sup>**

| Protocolo de revisão de literatura         |                                                                                                                                                                                        |
|--------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Critérios                                  | Descrição                                                                                                                                                                              |
| <b>Objetivo geral</b>                      | Investigar as aplicações e associações ao RAG na área da Ciência da Informação.                                                                                                        |
| <b>Questões a serem resolvidas</b>         | Como a utilização de RAG está sendo refletida no contexto da Ciência da Informação?<br>Como a Ciência da Informação tem se apropriado dessa nova técnica de recuperação da informação? |
| <b>Fontes de informação pesquisadas</b>    | Bases de dados: ACM Digital Library, BRAPCI, IEE, Scielo, Scopus, Springer Nature e Web of Science.                                                                                    |
| <b>Critérios de elegibilidade</b>          | Idioma: inglês, espanhol e português, para aumentar a abrangência das publicações.                                                                                                     |
|                                            | Disponibilidade: documentos disponíveis na íntegra em acesso aberto ou a partir do acesso institucional no Portal de Periódicos da Capes, para leitura e análise de seu conteúdo.      |
|                                            | Recorte temporal: sem delimitação temporal.                                                                                                                                            |
| <b>Critérios de inclusão e de exclusão</b> | Inclusão de documentos que abordam a temática do estudo.                                                                                                                               |
|                                            | Exclusão de documentos duplicados.                                                                                                                                                     |
|                                            | Exclusão de documentos que não foram escritos em português, inglês ou espanhol.                                                                                                        |
|                                            | Documentos cujos títulos e resumos não abordam os objetivos da pesquisa.                                                                                                               |
|                                            | Documentos que não estejam disponíveis na íntegra, via acesso aberto ou Portal de Periódicos da Capes.                                                                                 |
| <b>Campos de busca</b>                     | Título, resumo e palavras-chave.                                                                                                                                                       |
| <b>Expressões de busca</b>                 | Geração aumentada por recuperação e suas variações E ciência da informação.                                                                                                            |
|                                            | As expressões foram utilizadas em português, inglês e espanhol.                                                                                                                        |
| <b>Strings</b>                             | ((("geração aumentada por recuperação" OR "retrieval-augmented generation" OR "generación potenciada por la recuperación" OR "geração aumentada via recuperação"                       |

<sup>7</sup> Dados da pesquisa (2026).



|                                                            |                                                                                                                                                                                                            |
|------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                            | OR "generación aumentada por recuperación" OR "RAG")<br>AND ("ciência da informação" OR "information science"<br>OR "ciencia de la información"))                                                          |
| <b>Procedimentos de seleção dos documentos recuperados</b> | Após a aplicação dos critérios de elegibilidade, será feita a leitura dos títulos e resumos dos documentos recuperados com o intuito de verificar a pertinência do conteúdo ao objetivo geral da pesquisa. |
| <b>Procedimentos de análise</b>                            | Leitura completa dos documentos a fim de identificar elementos sobre a utilização de RAG no contexto da Ciência da Informação.                                                                             |
| <b>Critério de exclusão após análise dos documentos</b>    | Trabalhos que não possuem abordagem conceitual, teórica ou metodológica sobre a utilização de RAG no contexto da Ciência da Informação.                                                                    |
| <b>Tratamento</b>                                          | Google Drive para a gestão dos artefatos da pesquisa em um repositório. Ferramenta Parsif.al para apoio na seleção dos documentos.                                                                         |

A condução da revisão foi delineada a partir de um processo realizado por etapas. Na **Etapa 1**, procedeu-se à definição dos conceitos, termos e sinônimos de pesquisa, de forma estruturada, com base em pesquisas da área e leituras prévias das autoras, já que não foram localizados vocabulários controlados com o conceito central da pesquisa, resultando em uma *string* geral para a realização das buscas. Essa etapa foi importante para manter um repertório terminológico consistente e alinhado aos eixos conceituais da pesquisa.

Na **Etapa 2**, foram definidas as bases de dados e elaboradas e testadas estratégias de busca, que incluíram o recorte na Ciência da Informação. Considerando o objetivo da revisão e suas questões, a seleção das fontes de informação foi orientada por sua relevância científica, abrangência e credibilidade acadêmica. Dessa forma, foram selecionadas as bases: BRAPCI, Scopus, Springer Nature e Web of Science. As buscas também foram feitas nas bases ACM Digital Library, Scielo e IEEE; porém não houve resultados.

É importante destacar que essa pesquisa considerou somente estudos publicados em bases tradicionais e de grande reconhecimento, mesmo sabendo que a temática é frequente em bases de *preprints*, como o Arxiv. Apesar de reconhecer a importância desse tipo de fonte, ela ainda é sensível a assuntos recentes. Portanto, optou-se por utilizar somente trabalhos revisados por pares, indexados em bases com rigor científico.



A pesquisa considerou, ainda, documentos nos idiomas português, inglês e espanhol, por serem as línguas de maior familiaridade das autoras, não havendo delimitação de data, uma vez que a temática ainda é recente.

A *string* geral foi, então, adaptada à sintaxe das bases, mantendo rigorosamente a mesma estratégia de busca para garantir a replicabilidade da pesquisa. Foram considerados os campos: título, resumo e palavras-chave, uma vez que, ao se ampliar para todos os campos disponíveis na base, o volume de publicações ficaria muito grande, com baixa precisão e alta revocação. Entretanto, na base BRAPCI, não há funcionalidade para selecionar os campos de pesquisa para a busca avançada. Desta forma, a pesquisa foi feita em todos os campos, gerando apenas três resultados.

A definição de critérios de elegibilidade para os documentos recuperados nas bases considerou como tipologia artigos de periódicos e trabalhos publicados em eventos, além da disponibilidade em acesso aberto ou via acesso institucional no Portal da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Portal Capes), que permite um acesso federado a bases de dados privadas incluídas neste convênio, de acordo com a vinculação de cada instituição participante no Brasil.

Na **Etapla 3**, após diversos testes e a validação das buscas em cada base selecionada, garantindo a correspondência com a *string* geral, procedeu-se à execução das buscas, exportação dos resultados e importação no Parsif.al<sup>8</sup> para apoio à seleção e à análise dos documentos. O Parsif.al é uma ferramenta utilizada como apoio a revisões sistemáticas de literatura, permitindo registrar todas as seleções realizadas, bem como os critérios utilizados e os comentários para cada documento.

Assim, na ferramenta, foi cadastrado o protocolo criado para a revisão, realizada a importação dos resultados de cada base de dados, no formato BIB, a identificação e eliminação automáticas de duplicidades e a seleção dos estudos pertinentes, aplicando-se os critérios de inclusão e exclusão, mediante leitura de títulos e resumos, de modo a verificar a aderência temática dos documentos e o alinhamento conceitual e metodológico ao objetivo geral da pesquisa.

Após essa etapa de seleção inicial, os estudos selecionados foram submetidos à leitura na íntegra, realizada por meio de uma planilha de detalhamento elaborada e preenchida manualmente pelas autoras. A planilha contemplou metadados e campos definidos a partir da estrutura dos documentos e de informações consideradas relevantes para subsidiar a análise, incluindo: identificador, título, tipo de documento,

---

<sup>8</sup> <https://parsif.al/>.

idioma, ano de publicação, citação, países ou localidades abordados, objetivo, referencial teórico (autores e temáticas em destaque), metodologia (principais aspectos), resultados, contribuições e limitações. Esse processo de leitura possibilitou a categorização dos estudos de acordo com os eixos teóricos da investigação, a saber: organização, representação e recuperação, favorecendo a identificação de padrões de convergência, lacunas de pesquisa e perspectivas teóricas predominantes, subsidiando a interpretação qualitativa e discussões dos resultados encontrados.

Destaca-se, ainda, que, tendo em vista que o objetivo do artigo possui foco na Ciência da Informação, foi realizada uma comparação nas bases de dados utilizadas, de forma a evidenciar essa diferença significativa com as áreas de tecnologia. Assim, para além do protocolo utilizado na pesquisa, foi realizada essa comparação a partir de uma busca sem o refinamento da Ciência da Informação. Os resultados são apresentados junto com a revisão, na próxima seção.

## **4. Resultados**

Nesta seção, apresentam-se os resultados da revisão de literatura, incluindo a caracterização da amostra e a discussão dos achados.

### **4.1 Amostra**

As buscas foram realizadas considerando as variações de RAG, em inglês, português e espanhol, no início de julho de 2026, abrangendo, assim, todas as publicações até o primeiro semestre de 2026, ou seja, até 30/06/2026. Tendo em vista os quantitativos pouco expressivos, tomou-se o cuidado de fazer algumas comparações a fim de confirmar a hipótese do estudo de que a técnica tem sido muito utilizada em diversas áreas, mas com pouca frequência na Ciência da Informação.

Dessa forma, foi feito um teste retirando a restrição da Ciência da Informação para aumentar a revocação, sendo identificado um número muito expressivo de resultados na Scopus (21251) e na Web of Science (12140), o que mostra que a temática é volumosa quando não se especifica a área. Também foi feita outra alteração para aumentar a revocação, ampliando a busca para todos os campos das bases, e não somente título, resumo e palavras-chave, mas mantendo a restrição da Ciência da Informação. Foi observado que não houve alterações significativas, com exceção da Web of Science (116) e da Scopus (82499); mas, ao se verificar as primeiras páginas dos

resultados, é possível observar que a precisão estava muito ruim, recuperando artigos fora do escopo da pesquisa.

Com isso, prosseguiu-se com a busca definida no protocolo, que recuperou inicialmente 205 documentos nas quatro bases selecionadas que retornaram resultados. Em seguida, foram removidos os documentos duplicados, que corresponderam a apenas uma ocorrência. Posteriormente, procedeu-se à seleção com base no título, no resumo e nas palavras-chave, resultando em 16 documentos. O grande número de exclusões está relacionado à opção de não aplicar filtros nas bases de dados, o que resultou na recuperação de documentos que continham alguns dos termos pesquisados mas não estavam relacionados ao escopo da pesquisa. O termo e a arquitetura “*Retrieval-Augmented Generation* (RAG)” foram introduzidos por Lewis et al. em 2020, na publicação *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*. No entanto, as buscas recuperaram publicações desde 1970. Optou-se por não estabelecer um recorte temporal para verificar se técnicas semelhantes à RAG já eram utilizadas antes de 2020, ainda que sob outras denominações. Com isso, o uso da sigla recuperou artigos fora do escopo da pesquisa, relacionados a outros conceitos e domínios. Ainda assim, optou-se por mantê-la na expressão de busca, a fim de contemplar todas as possibilidades de recuperação.

Antes de iniciar a leitura na íntegra, realizou-se uma leitura flutuante, com a finalidade de obter uma visão inicial dos documentos. Observou-se que quatro artigos não estavam disponíveis em acesso aberto, mesmo mediante autenticação pelo acesso federado ao Portal Capes; três estavam redigidos em outros idiomas, como mandarim ou alemão; e outros quatro não se inseriram no escopo da pesquisa no campo da Ciência da Informação. Assim, a amostra final foi composta por cinco artigos, conforme detalhado no Quadro 3.

### **Quadro 3 - Seleção de documentos<sup>9</sup>**

---

<sup>9</sup> Dados da pesquisa (2026).

| Base de dados   | Título, resumo e palavras-chave |            |              | Leitura flutuante | Leitura completa |
|-----------------|---------------------------------|------------|--------------|-------------------|------------------|
|                 | Recuperados                     | Duplicados | Selecionados |                   |                  |
| Brapi           | 3                               | 0          | 3            | 16                | 5                |
| Scopus          | 15                              | 0          | 5            |                   |                  |
| Springer Nature | 183                             | 1          | 8            |                   |                  |
| Web of Science  | 4                               | 0          | 0            |                   |                  |
| Total           | 205                             | 1          | 16           |                   |                  |

## 4.2 Estudos selecionados

Os estudos selecionados demonstraram que a pesquisa sobre RAG na Ciência da Informação ainda se encontra em estágio inicial, com publicações recentes (2024 - 2026) e participação de pesquisadores brasileiros (Quadro 4). Observou-se uma diversidade nos domínios, como saúde e patrimônio cultural, indicando o potencial do RAG para apoiar diferentes processos informacionais.

Embora os estudos possuam caráter aplicado, voltado ao desenvolvimento de sistemas e à análise de casos de uso, há discussões que reforçam a necessidade de fundamentação nos princípios da organização e representação da informação para aumentar a confiabilidade, a contextualização e a qualidade das respostas produzidas por sistemas baseados em RAG.

**Quadro 4 - Estudos selecionados<sup>10</sup>**

| ID | Título                                                                                                              | Autores                                  | Ano  | País(es) | Objetivo                                                                                                                                            | Resultados                                                                                                                                                                                                                                                                                                            |
|----|---------------------------------------------------------------------------------------------------------------------|------------------------------------------|------|----------|-----------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | A Brapci como ferramenta do fortalecimento da ciência aberta: inovações na organização e disseminação da informação | Rene Faustino Gabriel Junior             | 2025 | Brasil   | Descrever as inovações desenvolvidas na Brapci para promover a interoperabilidade e a organização eficiente da informação em repositórios digitais  | Evidencia a importância das bases de dados especializadas em Ciência da Informação na promoção da Ciência Aberta e na democratização do acesso à informação científica                                                                                                                                                |
| 2  | A new “sense of position”: from the logic of subject indexing to the geometry of embeddings                         | João Alberto de Oliveira Lima            | 2025 | Brasil   | Propor uma estrutura conceitual demonstrando como os <i>embeddings</i> de texto oferecem uma solução para o problema da "indeterminação de assunto" | Demonstra que a tecnologia de <i>embeddings</i> viabiliza a mudança da indexação baseada em regras humanas rígidas para um modelo geométrico impulsionado por similaridade de dados                                                                                                                                   |
| 3  | FolkRAG: a retrieval-augmented generation system for cultural heritage materials                                    | Paul Kelly, Jonathan Schild, Amir Jafari | 2025 | EUA      | Introduzir um sistema de prova de conceito que utiliza RAG para integrar e facilitar a pesquisa em coleções de patrimônio cultural                  | Demonstra como a integração de metadados de diversas fontes de dados é usada para construir um banco de dados vetorial que alimenta sistemas RAG avançados, melhorando o acesso a materiais do patrimônio cultural tanto para usuários iniciantes quanto para especialistas, sem sacrificar as práticas arquivísticas |

<sup>10</sup> Dados da pesquisa (2026).

|   |                                                                                                                                                 |                                     |      |        |                                                                                                                                                                                                                             |                                                                                                                                                                         |
|---|-------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|------|--------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 4 | Modelos de linguagem (LLM) com geração aumentada por recuperação (RAG) na saúde: a Ciência da Informação no centro dos desafios e oportunidades | José Eduardo Santarem Segundo       | 2025 | Brasil | Examinar a convergência tecnológica sob uma perspectiva multidisciplinar, de forma a apresentar alguns elementos de abordagem, o estado da arte, oportunidades e desafios do uso e aplicação de LLMs e RAG na área da saúde | Foram identificados oito eixos que evidenciam que há muitas oportunidades para desenvolvimento de pesquisas na aplicação de LLMs com RAG na área de informação em saúde |
| 5 | Recuperação da informação e inteligência artificial generativa com large language model e retrieval-augmented generation                        | Henrique Leal Tavares <i>et al.</i> | 2024 | Brasil | Apresentar um modelo de RI utilizando como interface um <i>chatbot</i> e aplicando as técnicas de IAG, LLM e RAG                                                                                                            | Criação de um protótipo com um ambiente de recuperação da informação construído a partir da aplicação das técnicas de IAG e RAG                                         |

### 4.3 Discussão

Inicialmente, observou-se que há divergências na definição e nas características dos LLMs entre pesquisadores, principalmente de domínios diferentes. No estudo de Tavares *et al.* (2024, seção "Recuperação da informação...", para. 6), é utilizada a definição de Zhao *et al.* (2023) para classificar LLMs como "modelos de linguagem Transformer que contêm centenas de bilhões (ou mais) de parâmetros [...]". Porém, a arquitetura de Transformer foi introduzida por Vaswani *et al.* em 2017, com a publicação seminal *Attention Is All You Need*. Antes disso, as arquiteturas mais utilizadas eram baseadas em Recurrent Neural Networks (RNN) e, posteriormente, em Long Short-Term Memory (LSTM) (Santarem Segundo, 2025). Desta forma, os LLMs atuais são predominantemente baseados na arquitetura Transformer, mas seu conceito não está restrito a essa arquitetura.

Há uma divergência de opiniões quanto ao acesso na literatura. Enquanto Santarem Segundo (2025, seção "Modelos de linguagem...", para. 4) afirma, citando Erickson *et al.* (2025), que "o acesso fácil e econômico a LLMs [...] permitiu que uma geração de desenvolvedores e pesquisadores integrasse a IA moderna baseada em LLMs em aplicações cotidianas". Em outra perspectiva, Barnett *et al.* (2024) relatam que o uso de LLMs é computacionalmente custoso e dependente de infraestrutura de alto desempenho, o que limita sua adoção em contextos com restrições técnicas ou orçamentárias, mesmo com a utilização de abordagens como RAG. Nesse sentido, Carta *et al.* (2024), Siciliani *et al.* (2023) e Raja *et al.* (2024) também discutem que limitações de *hardware* e recursos financeiros afetaram a escalabilidade e o desempenho dos sistemas desenvolvidos em suas pesquisas.

Mas, para além da questão conceitual e das perspectivas, esta pesquisa objetiva discutir a organização, a representação e a recuperação da informação e do conhecimento sob a perspectiva da Ciência da Informação para implementação do RAG.

Apesar das "facilidades" proporcionadas pelo expressivo número de ferramentas atualmente disponíveis para a criação de sistemas RAG integrados a modelos de linguagem, a organização da informação e do conhecimento permanece como elemento essencial para assegurar a qualidade dessas aplicações.

O conceito de base de conhecimento apresenta estreita relação com o RAG e constitui um instrumento basilar para a construção de aplicações de qualidade. Na Ciência da Informação, as bases de conhecimento estão apoiadas nos processos de organização e representação da informação e do conhecimento, envolvendo sistemas de classificação, taxonomias e metadados de diferentes tipologias, empregados para descrever, contextualizar e relacionar dados, informações e conhecimentos.

Paralelamente, a curadoria perpassa todo o processo de construção e manutenção da base de conhecimento. Trata-se de uma atividade fundamental tanto para a seleção



inicial dos documentos que irão compor a base quanto para a avaliação e a inclusão contínua de novos conteúdos ao longo de seu ciclo de vida. Esse processo também contribui para assegurar a relevância, a confiabilidade, a atualização e a adequação temática dos documentos recuperados. Santarem Segundo (2025) destaca a importância da curadoria ao indicar que, no RAG, o *prompt* é tão relevante quanto a curadoria dos documentos que integram a base externa.

Nesse contexto, destaca-se a contribuição da **organização da informação e do conhecimento** para o RAG, uma vez que elas possibilitam a criação de bases estruturadas e organizadas a partir de critérios previamente definidos para a seleção, a disposição e o relacionamento dos documentos, estruturadas por taxonomias hierárquicas ou facetadas. As taxonomias facetadas, muito utilizadas nos ambientes atuais, permitem representar a complexidade envolvida no tratamento de uma ampla variedade de informações, documentos, temas e contextos.

Em resposta a essas demandas, têm sido desenvolvidos serviços destinados ao gerenciamento de bases de conhecimento, como o Amazon Bedrock Knowledge, que, em seu modelo Titan, permite um processamento mais rápido em escala, apesar dos potenciais desafios com chamadas de Application Programming Interface (API) de alto volume (Kelly, Schild & Jafari, 2025). Além disso, esse serviço incorpora recursos relacionados à representação semântica, aos *embeddings* e aos grafos de conhecimento, possibilitando estabelecer vínculos entre conteúdos provenientes de diferentes fontes de dados. Tais recursos também se relacionam aos conceitos e às aplicações das ontologias.

Quando baseada em ontologias, a técnica RAG é chamada de Ontology-Grounded Retrieval-Augmented Generation (OG-RAG). Ela fundamenta os documentos na ontologia do domínio, convertendo-os num hipergrafo ancorado na ontologia e, em seguida, recuperando um conjunto mínimo de hiperarestas para reconstruir o contexto. A OG-RAG pode melhorar a recuperação em comparação com o RAG padrão. Já a Ontology-Guided Retrieval-Augmented Generation (ORAG) orienta a recuperação para a classificação de entidades temáticas específicas com base numa ontologia de rótulos, mitigando as alucinações e melhorando a precisão em modelos de base LLM de alto desempenho (Lima, 2026; Sharma, Kumar e Li, 2024; Xiao *et al.*, 2024). Cardoso, Ferneda e Souza (2025) apontam que as ontologias e a recuperação semântica podem ser utilizadas para estabelecer conexões mais profundas entre conceitos e termos utilizados nas buscas.

Esse cenário evidencia que o cientista da informação pode desempenhar um papel relevante no desenvolvimento de aplicações dessa natureza, uma vez que sua formação abrange fundamentos essenciais ao RAG, embora esses elementos também sejam investigados por outras áreas do conhecimento e em diferentes domínios de aplicação.

No âmbito da **representação da informação e do conhecimento**, os conceitos e processos de indexação constituem fundamentos para que o conteúdo dos documentos

selecionados e submetidos à curadoria seja representado em formatos processáveis por máquina. A definição de campos, atributos e esquemas de metadados nas bases de conhecimento permite que os documentos sejam indexados de maneira contextualizada, favorecendo sua identificação, seu relacionamento e sua posterior recuperação.

No estudo de Kelly, Schild e Jafari (2025), a técnica RAG é utilizada para enfrentar as limitações decorrentes de metadados arquivísticos fragmentados e de ferramentas de busca desconectadas em extensas coleções arquivísticas, tendo como objeto as coleções do American Folklife Center, da Library of Congress. A partir da coleta de dados por meio de *web scraping* e APIs, foi projetado um banco de dados vetorial para lidar com a ausência de padronização e com a natureza especializada dos metadados presentes em múltiplos conjuntos de dados. A proposta buscou articular, de forma inovadora, princípios fundamentais da Biblioteconomia, como serviço, responsabilidade e autoridade, com o uso de ferramentas computacionais. Para isso, foi desenvolvido um esquema de metadados baseado em uma abordagem modular, com o objetivo de integrar, de maneira unificada, diferentes fontes de descrição arquivística e viabilizar sua utilização em um sistema RAG. Com a solução, foi possível integrar metodologias avançadas de RI a recursos de geração de linguagem, contribuindo para ampliar a interatividade e a eficiência da busca e da navegação nas coleções folclóricas analisadas.

Ainda nesse cenário, com um volume expressivo de dados e considerando os ambientes digitais contemporâneos com padrões diferentes e não integrados, a criação de grandes bases de conhecimento, como os chamados *data spaces* comerciais, configura-se como uma oportunidade de atuação para os profissionais da Ciência da Informação. Estes ambientes constituem serviços robustos que reúnem grandes conjuntos de dados provenientes de diferentes fontes, em sua maioria abertas, para utilização em diversos domínios e, mais recentemente, para alimentar modelos de IA. Essa possibilidade de atuação, contudo, apresenta um requisito essencial: a competência em dados.

Conforme destacam Guedes e Nascimento Silva (2025), novas competências relacionadas aos dados são necessárias ao profissional da informação. Esse profissional precisa saber acessar diferentes fontes, seja por meio da coleta de dados em uma API, seja pelo manuseio ou pela análise desses dados. Ele também deve ser capaz de criar e manipular bancos de dados, reconhecer tipos e formatos, compreender estruturas de dados e aplicar os resultados obtidos. Essas atividades frequentemente envolvem a criação de algoritmos e o domínio de linguagens de programação, mas não exigem que sejam realizadas por um programador.

Por fim, a **recuperação da informação** constitui o processo responsável por localizar, selecionar e disponibilizar os documentos ou os trechos mais relevantes para responder às solicitações realizadas pelos usuários. No RAG, esse processo pode combinar mecanismos de busca semântica, indexação, filtros e metadados, de modo a considerar não apenas a correspondência entre termos, mas também o contexto e o

significado das informações. Os conteúdos recuperados são, então, fornecidos ao modelo de linguagem como elementos de fundamentação para a geração das respostas.

Conforme Cardoso, Ferneda e Souza (2025), a recuperação não se restringe à resposta pontual a uma consulta, mas envolve a construção de significados, o suporte à aprendizagem e a ampliação da competência informacional dos sujeitos. Assim, a técnica RAG, ao combinar a recuperação de documentos com a geração textual contextualizada, amplia as possibilidades de atendimento às necessidades informacionais dos usuários.

A técnica RAG representa um avanço significativo no processamento de linguagem natural ao combinar LLMs com capacidades de recuperação de informação direcionada com fontes de dados externas, armazenadas em bancos de dados vetoriais, permitindo-lhes superar limitações como alucinações e dados de treinamento estáticos. A arquitetura central compreende um recuperador que localiza informações relevantes do banco de dados vetorial e um gerador que incorpora essas informações para produzir respostas precisas e ricas em contexto Lewis *et al.* (2020); Veturi *et al.* (2024); Guu *et al.* (2020); Karpukhin *et al.* (2020); Izacard e Grave (2020).

Desse modo, no RAG, a qualidade da recuperação depende também da organização, da representação e da curadoria envolvidas na construção desse banco vetorial, reafirmando a contribuição dos fundamentos da Ciência da Informação para o desenvolvimento de aplicações RAG mais contextualizadas, precisas e confiáveis.

## 5. Considerações finais

Tendo em vista a evolução tecnológica e fenomenológica dos últimos anos, em paralelo com a área da Ciência da Informação, é possível observar que a área mantém seu propósito, mas é inegável que precisa acompanhar as mudanças, especialmente as tecnológicas, que alteram a forma de produzir e utilizar a informação.

Com a IAG, popularizada a partir de 2022 e inicialmente mais presente nas áreas de tecnologia, as diversas áreas do conhecimento passaram a dispor de ferramentas e respectivas técnicas para apoiar diferentes atividades e processos. Entre elas, destaca-se a técnica RAG, que utiliza fontes externas para otimizar, potencializar e contextualizar cenários. Nesse contexto, surgiu o questionamento desta pesquisa, relacionado à forma como a Ciência da Informação está inserida nesses processos. Assim, a pesquisa objetivou investigar as aplicações e associações do RAG na área da Ciência da Informação, a partir de uma revisão da literatura, realizada em 2026. Destaca-se que, neste estudo, não foram utilizadas bases de *preprints*, por não apresentarem publicações validadas por pares, apesar do expressivo volume de trabalhos sobre a temática.

Como resultados, foram inicialmente selecionados 205 documentos e, após a aplicação dos critérios de inclusão e exclusão e a leitura completa, somente cinco documentos correspondem à amostra final. Trata-se de um número muito pequeno, mas

que foi validado nas bases e cuja restrição está diretamente relacionada ao recorte da Ciência da Informação.

Ao analisar a tecnologia RAG sob a perspectiva da organização, da representação e da recuperação, observa-se que a Ciência da Informação está presente, seja na criação de taxonomias e classificações para a organização, seja na indexação e na atribuição de metadados para a representação, seja na linguagem natural e nos aspectos semânticos para a recuperação. Além disso, a questão que poderá ampliar a inserção do profissional da informação nessa temática é a competência em dados, que envolve a habilidade de compreender e analisar estruturas de dados, bem como utilizar recursos tecnológicos, como bancos de dados e linguagens de programação, para sua manipulação.

Diante desse contexto, o profissional da informação precisa se reinventar e se adaptar às mudanças para continuar aplicando os fundamentos da área em diferentes contextos, assim como ocorre em outras áreas do conhecimento. Além disso, são importantes formações complementares aos cursos de graduação e pós-graduação que envolvam a motivação dos profissionais para buscar novos conhecimentos e aceitar desafios, saindo da zona de conforto para alçar novos voos.

Como trabalhos futuros, sugere-se aplicar a técnica, seja por meio de ferramentas já consolidadas ou customização de modelos, com casos específicos da Ciência da Informação, a fim de fomentar novas utilizações na área.

## **Agradecimento**

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio à pesquisa, processo 303721/2025-1.

À Agência Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro concedido.

## **6. Referências**

Alvares, L. & Araújo Júnior, R. H. de. (2010). Marcos históricos da ciência da informação: breve cronologia dos pioneiros, das obras clássicas e dos eventos fundamentais. *Transinformação*, 22(3), 195–205.  
<https://www.redalyc.org/pdf/3843/384334885001.pdf>

Araújo, C. A. Á. (2003). A ciência da informação como ciência social. *Ciência da Informação*, 32(3), 21–27. <https://doi.org/10.1590/S0100-19652003000300003>

- Baeza-Yates, R. & Ribeiro-Neto, B. (2013). *Recuperação de informação: conceitos e tecnologia das máquinas de busca* (2nd ed.). Bookman.
- Barnett, S., Kurniawan, S., Thudumu, S., Brannelly, Z. & Abdelrazek, M. (2024). Seven failure points when engineering a retrieval augmented generation system. In *Proceedings of the 3rd International Conference on AI Engineering – Software Engineering for AI* (pp. 194–199). ACM.  
<https://doi.org/10.1145/3644815.3644945>
- Borko, H. (1968). Information science: What is it? *American Documentation*, 19(1), 3–5. <https://doi.org/10.1002/asi.5090190103>
- Bourne, K. (2024). *Unlocking Data with Generative AI and RAG*. Birmingham: Packt.
- Cardoso, F., Ferneda, E. & Souza, C. B. dos S. de. (2025). A contribuição da Retrieval-Augmented Generation para a Ciência da Informação: inovações, aplicações e implicações éticas. In XI EDICIC Ibérico: Formação e investigação em Ciência da Informação: Oportunidades e desafios.  
<https://doi.org/10.34630/xiedicic.vi.6699>
- Carta, S., Giuliani, A., Manca, M. M., Piano, L., Podda, A. S., Pompianu, L., & Tiddia, S. G. (2024). A zero-shot strategy for knowledge graph engineering using GPT-3.5. *Procedia Computer Science*, 246, 2235–2243.  
<https://doi.org/10.1016/j.procs.2024.09.573>
- Dahlberg, I. Teoria do Conceito. (1978). *Ciência da Informação*, 7(2), 101-107.  
[https://upload.wikimedia.org/wikipedia/commons/6/66/Teoria\\_do\\_conceito\\_%28Cionline\\_115%29.pdf](https://upload.wikimedia.org/wikipedia/commons/6/66/Teoria_do_conceito_%28Cionline_115%29.pdf)
- Erickson, J. S., Santos, H., Pinheiro, V., McCusker, J. P. & McGuinness, D. L. (2025). LLM experimentation through knowledge graphs: Towards improved management, repeatability, and verification. *Journal of Web Semantics*, 85, Article 100853. <https://doi.org/10.1016/j.websem.2024.100853>
- Ferneda, E. (2012). *Introdução aos modelos computacionais de recuperação de informação*. Ciência Moderna.

- Gabriel Junior, R. F. (2025). A Brapci como ferramenta do fortalecimento da Ciência Aberta: inovações na organização e disseminação da informação. *Revista Ciência Aberta Lusófona*, 1, 289-296. <https://doi.org/10.82524/recal.2025.5>
- Gil, A. C. (2017). Como elaborar projetos de pesquisa (6th ed.). Atlas.
- Guedes, S. F. M. de S. & Nascimento Silva, P. (2025). Competência em dados para o profissional da informação: proposta de aplicativo para prevenção e gestão de desastres ambientais em Belo Horizonte. *Revista ACB*, 30(1), 1–28. <https://revista.acbsc.org.br/racb/article/view/2216>
- Guu, K., Lee, K., Tung, Z., Pasupat, P. & Chang, M-W. (2020). Retrieval augmented language model pre-training. In: International conference on machine learning [Preprint]. *Arxiv*. <https://doi.org/10.48550/arXiv.2002.08909>
- Instituto Brasileiro de Geografia e Estatística (IBGE). (2015). *Pnad 2013: internet pelo celular é utilizada em mais da metade dos domicílios que acessam a Rede*. Agência de Notícias IBGE. <https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/9840-pnad-2013-internet-pelo-celular-e-utilizada-em-mais-da-metade-dos-domicilios-que-acessam-a-rede>
- Izacard, G. & Grave, E. (2020) Leveraging Passage Retrieval with Generative Models for Open Domain Question Answering [Preprint]. *Arxiv*. <https://doi.org/10.48550/arXiv.2007.01282>
- Karpukhin, V., Oğuz, B., Min, S., Lewis, P., Wu, L., Edunov, S., Chen, D. & Yin, W-T. (2020). Dense passage retrieval for open-domain question answering [Preprint]. *Arxiv*. <https://doi.org/10.48550/arXiv.2004.04906>
- Kelly, P., Schild, J. & Jafari, A. (2025). FolkRAG: A retrieval-augmented generation system for cultural heritage materials. *Neural Computing and Applications*, 37, 20281–20297. <https://doi.org/10.1007/s00521-025-11455-4>
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-T., Rocktäschel, T., Riedel, S. & Kiela, D. (2020). Retrieval-



- augmented generation for knowledge-intensive NLP tasks [Preprint]. *Arxiv*.  
<https://doi.org/10.48550/arXiv.2005.11401>
- Lima, J. A. de. O. (2025). A new “sense of position”: From the logic of subject indexing to the geometry of embeddings. *Journal of Documentation*, 82(1), 136–153. <https://doi.org/10.1108/JD-07-2025-0204>
- Luz, L. P., Coneglian, C. S. & Segundo, J. E. S. (2019). Tecnologias da web semântica para a recuperação da informação no Wikidata. *Revista Digital de Biblioteconomia e Ciência da Informação*, 17, e019003.  
<https://doi.org/10.20396/rdbci.v17i0.8651791>
- Nascimento Silva, P. (2023). Recuperação de informação na Ciência da Informação: produção acadêmico-científica brasileira (2012-2021). *Transinformação*, 35, e237336, 1-17.  
<https://doi.org/10.1590/2318-0889202335e237336>
- Otlet, P. (1934). *Traité de documentation: Le livre sur le livre, théorie et pratique*. Editions Mundaneum.
- Pinheiro, L. V. R. & Loureiro, J. M. M. (1995). Traçados e limites da ciência da informação. *Ciência da Informação*, 24(1).  
<https://revista.ibict.br/ciinf/article/view/609>
- Procópio, D. C. (2025). *Large language models para recuperação da informação em acervos históricos digitalizados* [Dissertação de mestrado, Universidade Federal de Minas Gerais]. Repositório Institucional da Universidade Federal de Minas Gerais. <https://hdl.handle.net/1843/3339>
- Queiroz, D. G. de C. & Moura, A. M. M. de. (2015). Ciência da Informação: história, conceitos e características. *Em Questão*, 21(3), 26–42.  
<https://seer.ufrgs.br/index.php/EmQuestao/article/view/57516>
- Raja, H., Munawar, A., Mylonas, N., Delsoz, M., Madadi, Y., Elahi, M., Hassan, A., Abu Serhan, H., Inam, O., Hernandez, L., Chen, H., Tran, S., Munir, W., Abd-Alrazaq, A. & Yousefi, S. (2024). Automated category and trend analysis of scientific articles on ophthalmology using large language models: Development

and usability study. *JMIR Formative Research*, 8, e52462.

<https://doi.org/10.2196/52462>

Ranganathan, S. R. (1967). *Prolegomena to library classification* (3rd ed.). London: Asia Publishing House.

Santarem Segundo, J. E. (2025). Modelos de linguagem (LLMs) com geração aumentada por recuperação (RAG) na saúde: A Ciência da Informação no centro dos desafios e oportunidades. In *Anais do XXV Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ENANCIB)*, 5(1).

[https://enancib.ancib.org/enancib/pt\\_BR/article/view/1195](https://enancib.ancib.org/enancib/pt_BR/article/view/1195)

Saracevic, T. (1996). Ciência da Informação: origens, evolução e relações.

*Perspectivas em Ciência da Informação*, 1(1), 41-62.

<https://periodicos.ufmg.br/index.php/pci/article/view/22308>

Sharma, K., Kumar, P. & Li, Y. (2024). OG-RAG: Ontology-grounded retrieval-augmented generation for large language models [Preprint]. *Arxiv*.

<https://arxiv.org/abs/2412.15235>

Siciliani, L., Taccardi, V., Basile, P., Di Ciano, M. & Lops, P. (2023). AI-based decision support system for public procurement. *Information Systems*, 119,

Article 102284. <https://doi.org/10.1016/j.is.2023.102284>

Souza, R. A. S. & Rodas, C. M. (2020). Recuperação da informação em dispositivos móveis. *Biblos: Revista do Instituto de Ciências Humanas e da Informação*, 34(2),

147–166. <https://doi.org/10.14295/biblos.v34i2.11840>

Stock, W. G. & Stock, M. (2013). *Handbook of information science*. De Gruyter.

Tavares, H. L., Coneglian, C. S., Torino, E., Vidotti, S. A. B. G. & Santarem Segundo, J. E. (2024). Recuperação da informação e inteligência artificial generativa com

large language model e retrieval-augmented generation. In *Anais do XXIV Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da*



*Informação (ENANCIB)*, 4(1).

[https://enancib.ancib.org/enancib/pt\\_BR/article/view/1801](https://enancib.ancib.org/enancib/pt_BR/article/view/1801)

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. & Polosukhin, I. (2017). Attention is All You Need. In I. Guyon *et al.* (Eds.), *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates, Inc.

[https://proceedings.neurips.cc/paper\\_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf)

Veturi, S., Vaichal, S., Jagadheesh, R. L., Tripto, N. I. & Yan, N. (2024). RAG based Question-Answering for Contextual Response Prediction System [Preprint]. *Arxiv*. <https://doi.org/10.48550/arXiv.2409.03708>

Xiao, J., Ding, L., Barry, J., Elkaref, M., De Mel, G. & Han, J. (2024). Ontology-guided retrieval-augmented generation for theme-specific entity typing. In *Proceedings of the First Conference on Language Modeling (COLM 2024)*. <https://openreview.net/pdf?id=cKBmZ2PZ6c>

World Wide Web Foundation. (s.d.). *History of the Web*.

<https://webfoundation.org/about/vision/history-of-the-web/>

Wüster, E. (1931). *Internationale Sprachnormung in der Technik, besonders in der Elektrotechnik: Die nationale Sprachnormung und ihre Verallgemeinerung*. VDI-Verlag.

Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., ... Wen, J.-R. (2023). A survey of large language models [Preprint]. *Arxiv*. <https://doi.org/10.48550/arXiv.2303.18223>