



Interfaces de hemerotecas digitais em português: análise crítica e desafios para o desenvolvimento de uma nova geração de acervos históricos

Eric Brasil Nepomuceno

Universidade da Integração Internacional da Lusofonia Afro-Brasileira (UNILAB),
Ceará, Brasil

ORCID: <https://orcid.org/0000-0001-5067-8475>

E-mail: profericbrasil@gmail.com

RESUMO

Este artigo analisa as interfaces gráficas de usuário (IGUs) de hemerotecas digitais em português e suas implicações para a pesquisa em história. O estudo examina a Hemeroteca Digital Brasileira (HDB), a Biblioteca Nacional Digital de Portugal (BND-PT) e a Hemeroteca Digital de Lisboa (HDL), buscando compreender como suas características influenciam o acesso, a navegação e a exploração de fontes históricas digitalizadas. Para isso, aplicamos o *Método impresso*, de Ehrmann, Bunout e Düring (2019), avaliando seis critérios principais: crítica da fonte, busca de conteúdo, filtro de conteúdo, generosidade, gerenciamento pelo usuário e conectividade. Os resultados indicam avanços em metadados, navegação e visualização, mas evidenciam lacunas significativas em enriquecimento semântico, integração com repositórios externos e suporte a metodologias computacionais avançadas. Argumenta-se que essas características limitam o potencial das hemerotecas digitais e reforçam o uso de práticas analógicas sobre os recursos digitais, restringindo abordagens inovadoras. O artigo defende a necessidade de um salto qualitativo rumo às “Hemerotecas Digitais 2.0”, promovendo maior conectividade, personalização e transparência nas decisões técnicas, consolidando essas plataformas como ferramentas essenciais para a pesquisa histórica contemporânea.

PALAVRAS-CHAVE: Interfaces Gráficas de Usuário; Hemerotecas Digitais; História Digital.



Graphical user interfaces of digital newspaper libraries in portuguese: a critical analysis and challenges for the development of next-generation historical archives

ABSTRACT

This article analyzes the Graphical User Interfaces (GUIs) of digital newspaper archives in Portuguese and their implications for historical research. The study examines the Hemeroteca Digital Brasileira (HDB), the Biblioteca Nacional Digital de Portugal (BND-PT), and the Hemeroteca Digital de Lisboa (HDL), aiming to understand how their features influence access, navigation, and exploration of digitized historical sources. To this end, we apply the *impresso Method*, developed by Ehrmann, Bunout, and Düring (2019), evaluating six key criteria: source criticism, content search, content filtering, generosity, user management, and connectivity. The results indicate progress in metadata, navigation, and visualization but reveal significant gaps in semantic enrichment, integration with external repositories, and support for advanced computational methodologies. We argue that these limitations constrain the potential of digital newspaper archives and reinforce the use of analog practices over digital resources, thereby limiting innovative approaches. The article advocates for a qualitative leap toward “Digital Newspaper Archives 2.0,” fostering greater connectivity, personalization, and transparency in technical decisions, and consolidating these platforms as essential tools for contemporary historical research.

KEYWORDS: Graphical User Interfaces; Digital Newspaper Archives; Digital History.

1. Introdução

Neste artigo,¹ busco discutir a relação entre as interfaces gráficas de usuário (doravante IGUs, do inglês *Graphical User Interfaces*) de hemerotecas digitais em português e a pesquisa em história digital. Para tanto, analiso as interfaces de três acervos digitais: a coleção de periódicos da Biblioteca Nacional Digital de Portugal, a Hemeroteca Digital de Lisboa e a Hemeroteca Digital Brasileira. A partir da análise, busco responder à seguinte questão: como suas características influenciam a pesquisa em história? Utilizo, para isso, um método de avaliação de IGUs desenvolvido por Ehrmann; Bunout e Düring (2019). Inicialmente, apresento uma breve revisão da literatura sobre o tema; em seguida, descrevo as IGUs selecionadas e o método utilizado. Por fim, apresento detalhadamente os resultados obtidos e as conclusões, encaminhando o argumento da necessidade de um salto qualitativo para o desenvolvimento de Hemerotecas Digitais 2.0, capazes de enfrentar as demandas e dilemas colocados pelo avanço das pesquisas no campo das Humanidades Digitais.²

É cada vez mais comum a utilização de acervos de fontes digitalizadas e nascidas digitais nas pesquisas históricas (Marino et al., 2021). Projetos, artigos, cursos têm recorrido a repro-

¹ Foi utilizado o ChatGPT 4o para a revisão ortográfica, gramatical e sintática, bem como para a formatação das referências. Ressalta-se que todo o conteúdo analítico, argumentativo e as ideias centrais do texto são de responsabilidade exclusiva do autor. As consultas foram realizadas entre os dias 3 e 16 de fevereiro e entre 6 e 8 de abril de 2025.

² Este artigo é um produto de pesquisa de pós-doutorado realizada entre outubro de 2022 e outubro de 2023 no Instituto de História Contemporânea (IHC), da Faculdade de Ciências Sociais e Humanas da Universidade NOVA de Lisboa, integrante do IN2PAST — Laboratório Associado para a Investigação e Inovação em Património, Artes, Sustentabilidade e Território. A pesquisa contou com a supervisão de Daniel Alves, a quem agradeço o acompanhamento atento e pelas contribuições fundamentais ao longo do processo. No Brasil, a continuidade das análises e a consolidação desta reflexão contaram com o suporte do Laboratório de Humanidades Digitais da Universidade Federal da Bahia (LABHDUFBA).

sitórios *on-line* de fontes primárias digitalizadas ou nato-digitais nos últimos anos, mesmo antes da pandemia de covid-19 (Brasil; Valverde; Veloso, 2024). Dentre esses, as hemerotecas digitais têm se destacado por fornecerem acesso a milhões de páginas digitalizadas e por possibilitarem o uso de recursos de busca, leitura e/ou *download* de páginas de variadas qualidades e formatos (Veloso, 2022). Bunout, Ehrmann e Clavert (2022) e Pfanzelter et al. (2021) convergem, ao analisar o impacto das hemerotecas digitais na pesquisa histórica, destacando tanto suas potencialidades quanto seus desafios. Ambos os estudos ressaltam que a digitalização massiva de periódicos facilita o acesso remoto e amplia as possibilidades de análise por meio de técnicas automatizadas, como mineração de dados e processamento de linguagem natural. No entanto, alertam para limitações estruturais dessas plataformas, incluindo a qualidade variável do reconhecimento óptico de caracteres (OCR), os vieses algorítmicos na busca e a fragmentação do acesso às fontes. Além disso, enfatizam que as interfaces digitais não são neutras, influenciando metodologicamente a pesquisa ao determinar como os historiadores acessam, organizam e interpretam os dados. Assim, ambos os estudos defendem a necessidade de abordagens críticas e metodologicamente rigorosas para explorar essas fontes digitais de maneira eficaz e consciente.

Pfanzelter et al. (2021) enfatizam ainda a necessidade de aprimoramento dessas plataformas, sugerindo que a integração de ferramentas, como modelagem de tópicos, análise de frequência de palavras e reconhecimento de entidades nomeadas, poderia beneficiar tanto pesquisadores acadêmicos quanto outros usuários. Ressaltam também que as interfaces atuais frequentemente não atendem às expectativas dos pesquisadores, o que revela um descompasso entre as possibilidades oferecidas e as demandas reais da pesquisa histórica. Concluem que, para que as hemerotecas digitais sejam efetivamente utilizadas como instrumentos de análise histórica aprofundada, é importante que as instituições provedoras invistam na melhoria da usabilidade, na acessibilidade a metadados estruturados e na transparência dos algoritmos de busca e classificação (Pfanzelter et al., 2021).

Ian Milligan (2013), em estudo de caso sobre a pesquisa histórica no Canadá, demonstra que, com o crescimento das bases de dados *on-line*, houve um aumento expressivo na citação de jornais digitalizados em dissertações e artigos acadêmicos, evidenciando como o acesso facilitado influencia diretamente as escolhas metodológicas dos historiadores (Milligan, 2013, p. 541-543). Além disso, o autor alerta que a ênfase na pesquisa baseada em palavras-chave pode enviesar a historiografia, ao privilegiar certos jornais em detrimento de fontes ainda não digitalizadas, criando um efeito de distorção nas abordagens historiográficas (Milligan, 2013, p. 549-553).

Para o caso brasileiro, em parceria com Leonardo Nascimento (2020), apontamos que a interface de busca da Hemeroteca Digital Brasileira (HDB) apresenta desafios estruturais que afetam a pesquisa histórica, especialmente no que diz respeito à qualidade do OCR e à falta de transparência nos processos de indexação e à limitação nas possibilidades de exportação de dados estruturados. Essa opacidade compromete a reprodutibilidade da pesquisa e impõe limitações significativas à aplicação de métodos computacionais mais avançados. Destacam, também, que a organização dos resultados de busca na HDB não permite uma visão estruturada das ocorrências no tempo, o que dificulta a construção de análises históricas mais aprofundadas.



Em outro artigo de minha autoria (Brasil, 2022), busquei avançar nessa discussão ao propor uma solução metodológica para lidar com as limitações da HDB por meio da ferramenta *pyHDB*. Esse *software*, baseado em *web scraping*, permite a coleta sistemática dos dados disponibilizados na plataforma, garantindo o registro detalhado dos parâmetros de busca e a organização estruturada das informações recuperadas. Enquanto Pfanzelter et al. (2021) enfatizam a necessidade de aprimoramento das interfaces digitais para viabilizar análises mais sofisticadas, propus uma abordagem prática para superar algumas dessas barreiras, permitindo que historiadores extraíam, organizem e analisem dados da HDB de forma mais eficiente. Ao integrar métodos computacionais ao processo de pesquisa histórica, o autor demonstra como o desenvolvimento de ferramentas auxiliares pode ampliar as possibilidades heurísticas e garantir maior rigor metodológico na análise de fontes digitalizadas (Brasil, 2022).

A partir dos estudos aqui apresentados, fica evidente que a relação entre as IGUs e as estratégias de acesso, seleção e interpretação de dados digitais constitui um aspecto crítico da prática histórica, que demanda análises mais aprofundadas e metodologicamente rigorosas. Entre os principais desafios identificados estão a falta de transparência nos processos de busca, a organização dos resultados e a ausência de ferramentas avançadas para análise e exploração dos dados. Diante dessas questões, este artigo busca examinar IGUs de hemerotecas digitais em português de forma sistemática, utilizando a metodologia proposta por Ehrmann, Bunout e Düring, doravante referido como *Método impresso* (2019). Essa abordagem permite uma análise detalhada de como as características técnicas e visuais das interfaces influenciam o processo de pesquisa, afetando diretamente os métodos heurísticos e hermenêuticos adotados pelos pesquisadores.

2. Acervos estudados: características e descrição geral

Analizamos aqui as IGUs de três importantes acervos em língua portuguesa: a coleção de periódicos da Biblioteca Nacional Digital de Portugal (doravante BND-PT), a Hemeroteca Digital de Lisboa (HDL) e a Hemeroteca Digital Brasileira (HDB) da Biblioteca Nacional do Brasil. Em 2002, a Biblioteca Nacional de Portugal lançou a Biblioteca Nacional Digital como parte de um esforço estratégico para modernizar seus serviços e ampliar o acesso ao vasto acervo da instituição. Inicialmente tratada como um projeto especial, a BND-PT rapidamente se consolidou como um serviço regular da biblioteca, refletindo uma transformação estrutural e estratégica que visava integrar coleções físicas e digitais em uma única infraestrutura operacional (Patrício, 2011). Atualmente, os itens digitalizados abrangem livros, manuscritos, jornais, mapas, partituras e materiais iconográficos, com publicações desde o século XII até o século XXI (Campos; Cabral; Borbinha, 2002; Patrício, 2011). Compõe o acervo digital da BND-PT uma vasta coleção de periódicos, reunindo jornais, revistas e outras publicações com variadas periodicidades, com acesso direto pela interface da BND-PT.³

³ Disponível em: <<https://bndigital.bnportugal.gov.pt/>>. Acesso em: 07 de fevereiro de 2025

A Hemeroteca Digital Brasileira (HDB)⁴ foi lançada em 2012 com o objetivo de digitalizar e disponibilizar periódicos históricos, possibilitando o acesso remoto a jornais e revistas que integram o acervo da Biblioteca Nacional do Brasil. Inicialmente, o projeto contou com o apoio da Financiadora de Estudos e Projetos (FINEP), que permitiu a digitalização de 10 milhões de páginas de periódicos (Bettencourt; Pinto, 2013, p. 1). Essa iniciativa incluiu o uso de OCR, que facilita buscas textuais e melhora a recuperação de informações (Giordano, 2016, p. 19) e sua IGU adota um padrão de buscas que permite a combinação de palavras-chave com parâmetros de filtro por datas, localidades e coleções específicas (Giordano, 2016, p. 21). A HDB está inserida em um projeto mais amplo chamado *BNDigital*, lançado em 2006, que consolidou “um programa sistemático de digitalização do acervo documental mantido pela Biblioteca Nacional” (Meirelles, 2020, p. 57).

A Hemeroteca Digital de Lisboa (HDL) foi criada em 2005 como uma iniciativa da Hemeroteca Municipal de Lisboa para a digitalização e disponibilização de periódicos em domínio público (Oliveira; Matos, 2015). O projeto surgiu como resposta à deterioração das coleções físicas e à necessidade de ampliar o acesso remoto ao acervo (Matos, 2001, p. 2). A HDL passou por uma atualização significativa, com a criação de uma nova plataforma digital, cuja versão atualizada foi lançada apenas após a realização desta pesquisa,⁵ razão pela qual a análise aqui apresentada foi baseada na versão anterior da plataforma⁶. Assim, os resultados desta investigação refletem as funcionalidades e limitações da antiga interface, sem levar em conta possíveis melhorias ou mudanças implementadas na nova versão.

3. O Método impresso

Como parte das pesquisas desenvolvidas pela equipe do projeto “*impresso – Media Monitoring of the Past*”, Ehrmann, Bunout e Düring (2019) desenvolveram uma metodologia para a análise crítica de IGUs de hemerotecas digitais. O projeto *impresso* é uma iniciativa interdisciplinar que integra linguistas computacionais, designers e historiadores para a criação de um *corpus* semântico de jornais históricos digitalizados, em múltiplos idiomas, como francês e alemão. Desenvolvido com financiamento do *Swiss National Science Foundation* (SNSF), o projeto busca transformar o conteúdo textual não estruturado desses jornais em dados estruturados, vinculados e semanticamente indexados (Ehrmann; Bunout e Düring, 2020, p. 959)⁷.

O *Método impresso* possibilita uma abordagem sistemática que combina critérios de avaliação qualitativos e quantitativos, para compreender como essas IGUs impactam a pesquisa histórica e a experiência do usuário. Baseia-se na identificação e análise de funcionalidades técnicas e práticas acessíveis diretamente na página *web* dessas interfaces. Para isso, foram definidos seis *CrITÉrios de Avaliação de Alto Nível*: crítica de fontes; busca de conteúdo; filtragem de conteúdo; generosidade; gestão e exploração de conteúdo pelo usuário e conectividade. Os autores elabo-

⁴ Disponível em: <<https://memoria.bn.gov.br/hdb>>. Acesso em: 07 de fevereiro de 2025

⁵ Disponível em: <<https://hemerotecadigital.lisboa.pt/>>. Acesso em: 07 de fevereiro de 2025.

⁶ Disponível em: <<https://hemerotecadigital.cm-lisboa.pt/>>. Acesso em: 07 de fevereiro de 2025

⁷ Disponível em: <<https://impresso-project.ch/>>. Acesso em: 07 de fevereiro de 2025



raram um formulário contendo 125 funcionalidades agrupadas em 12 famílias (Tabela 1). Essas funcionalidades foram mapeadas para os critérios de avaliação acima mencionados. A coleta de dados é feita diretamente nas IGUs por meio de observação sistemática e registradas em planilhas para posterior análise em termos de frequência e cobertura das funcionalidades (Ehrmann; Bunout; Düring, 2019, p. 6-7).

TABELA 1. Critérios de Avaliação de Alto Nível

Critério de Avaliação	Descrição
Crítica de Fontes	Examina a transparência das informações sobre os documentos digitalizados, incluindo metadados, proveniência e processos de digitalização.
Busca de Conteúdo	Avalia a amplitude e sofisticação das ferramentas de busca, como pesquisa por palavras-chave, operadores booleanos, busca por frases exatas e filtragem baseada em metadados.
Filtragem de Conteúdo	Analisa a existência de mecanismos que permitem refinar os resultados da busca, como filtros por período de tempo, títulos de jornais e local de publicação.
Generosidade	Avalia se a interface proporciona caminhos alternativos para a descoberta de conteúdos, incluindo sugestões de documentos relacionados e modos diferenciados de exibição de resultados.
Gestão e Exploração de Conteúdo pelo Usuário	Examina a capacidade da interface de permitir que os usuários organizem, marquem e comparem conteúdos em espaços personalizados.
Conectividade	Mede a integração do acervo com outras fontes e repositórios, avaliando a presença de conexões diretas, identificadores padronizados e exportação de dados para ferramentas externas.

Fonte: Produzido pelo/a autor/a

No estudo de Ehrmann, Bunout e Düring (2019), a seleção das interfaces gráficas seguiu critérios pragmáticos que buscavam equilibrar diversidade e viabilidade dentro dos limites da pesquisa. Foram incluídas interfaces que refletissem diferentes níveis de sofisticação tecnológica, abrangendo desde plataformas mais simples até as mais avançadas. A prioridade foi dada a interfaces mantidas por bibliotecas nacionais, regionais ou institucionais amplamente reconhecidas por suas coleções de jornais históricos. Além disso, as interfaces selecionadas operavam em idiomas compreendidos pelos pesquisadores, permitindo uma análise detalhada e precisa (Ehrmann; Bunout; Düring, 2019, p. 6-7).

O estudo abrangeu um total de 24 instituições, incluindo bibliotecas nacionais, regionais, portais comerciais e consórcios semipúblicos. Observou-se, contudo, uma predominância de plataformas baseadas em países europeus e nos Estados Unidos. Embora reconhecendo essa limitação, e explicitando que não tenham buscado dar conta da totalidade das hemerotecas digitais disponíveis, os autores afirmam que a amostra foi suficiente para mapear tendências e identificar boas práticas no *design* de interfaces de acesso a periódicos históricos (Ehrmann; Bunout; Düring, 2019, p. 6).

Os resultados revelaram que a cobertura das funcionalidades analisadas nas IGUs de hemerotecas digitais ainda é limitada. Cerca de três quartos das funcionalidades avaliadas estavam presentes em menos da metade das interfaces estudadas. Embora a maioria das plataformas ofereça ferra-

mentas básicas, como busca por palavras-chave e visualização de fac-símile, funcionalidades mais sofisticadas, como enriquecimento semântico, conectividade com outros repositórios e personalização da experiência do usuário, ainda são pouco comuns (Ehrmann; Bunout; Düring, 2019, p. 9).

Argumento que a ampliação do escopo da pesquisa, para incluir IGUs de hemerotecas digitais em língua portuguesa, contribui para um entendimento mais abrangente das dinâmicas envolvidas no acesso e na organização de acervos históricos digitalizados. O português, falado por mais de 260 milhões de pessoas no mundo, está presente em países com acervos documentais extensos que registram períodos históricos globalmente significativos, como a colonização e os processos de independência, Revoluções e Movimentos Sociais variados. A análise e comparação dessas IGUs de hemerotecas contribuem para uma visão mais diversa das práticas e soluções tecnológicas empregadas no acesso a coleções digitalizadas.

4. Aplicação do *Método impresso*

A aplicação do *Método impresso* foi conduzida a partir do preenchimento do formulário de avaliação para as três IGUs em português – BDP-PT, HDB e HDL –, considerando exclusivamente a página *web* onde a consulta é realizada e onde os resultados são acessados. Funcionalidades adicionais presentes em outras seções das plataformas não foram incluídas na análise. O preenchimento dos formulários ocorreu em junho de 2023, no âmbito da pesquisa de pós-doutorado, conforme registrado no log⁸. Os dados e os códigos utilizados na escrita deste artigo, assim como todos os gráficos gerados, estão disponíveis no repositório <https://github.com/ericbrasiln/artigo_interfaces>.⁹

Após a conclusão do preenchimento do formulário (acessíveis na pasta */data* do referido repositório), executei o *notebook*¹⁰ de autoria de Ehrmann, Bunout e Düring, com pequenas adaptações. O arquivo *.ipynb* está disponível na pasta */notebooks*, e os gráficos gerados podem ser acessados na pasta */charts*.

O primeiro passo do *notebook* consiste na preparação dos dados: o formulário de avaliação é carregado e estruturado para facilitar a análise; os dados são organizados em uma estrutura de índice múltiplo, garantindo uma categorização clara das funcionalidades das interfaces; antes de prosseguir com a análise, são realizadas limpezas e ajustes nos dados, como a remoção de espaços em branco desnecessários e a conversão de respostas categóricas para valores numéricos.

Em seguida, o *notebook* gera estatísticas descritivas, calculando a quantidade de interfaces que apresentam cada funcionalidade listada no formulário. As funcionalidades analisadas foram agru-

⁸ O repositório no GitHub relativo à pesquisa de pós-doutorado mantém um histórico detalhado das alterações realizadas, permitindo a rastreabilidade das etapas do estudo e a documentação das atividades associadas ao preenchimento dos formulários. Para acessar o repositório da pesquisa completo de pós-doutorado, <<https://github.com/ericbrasiln/hemdig-framework>>, e o arquivo específico com os logs: <https://github.com/ericbrasiln/hemdig-framework/blob/main/log_main.csv>.

⁹ O repositório está armazenado no Zenodo, com o DOI 10.5281/zenodo.15177362.

¹⁰ Um Jupyter Notebook é um ambiente interativo de computação que permite a execução de código, a inclusão de texto explicativo e a visualização de gráficos em um único documento. Ele é amplamente utilizado em pesquisa científica e análise de dados para documentar e reproduzir experimentos computacionais.

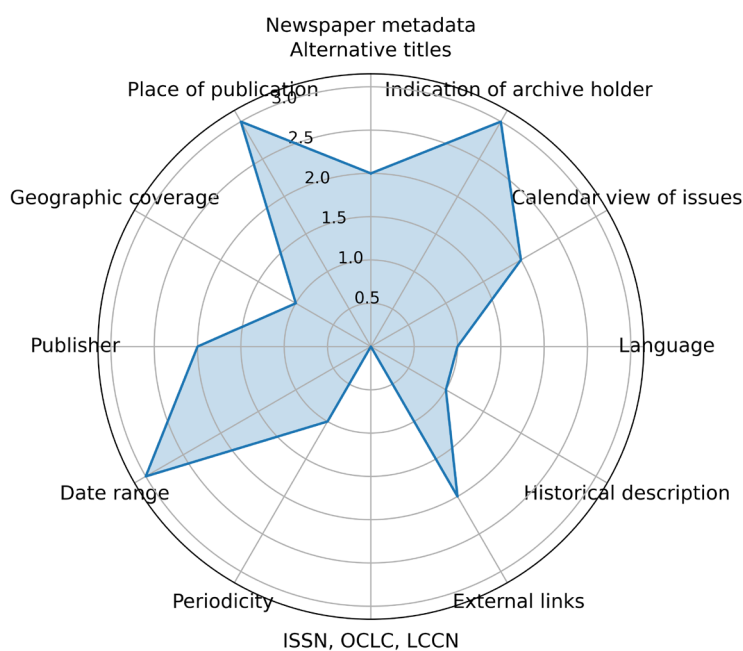


padas em 12 famílias: *Metadados dos periódicos*; *Visualização*; *Navegação*; *Busca*; *Exibição dos resultados*; *Filtros de Resultados*; *Classificação de resultado*; *Informações sobre a digitalização*; *Interação do Usuário*; *Conectividade*; *Enriquecimento* e *APIs*. E para cada uma é gerado um gráfico radial. Como o número de hemerotecas analisadas é reduzido, optou-se por trabalhar com valores absolutos ao invés de porcentagens nessas visualizações. Veremos agora o detalhamento desses resultados.

5. Resultados

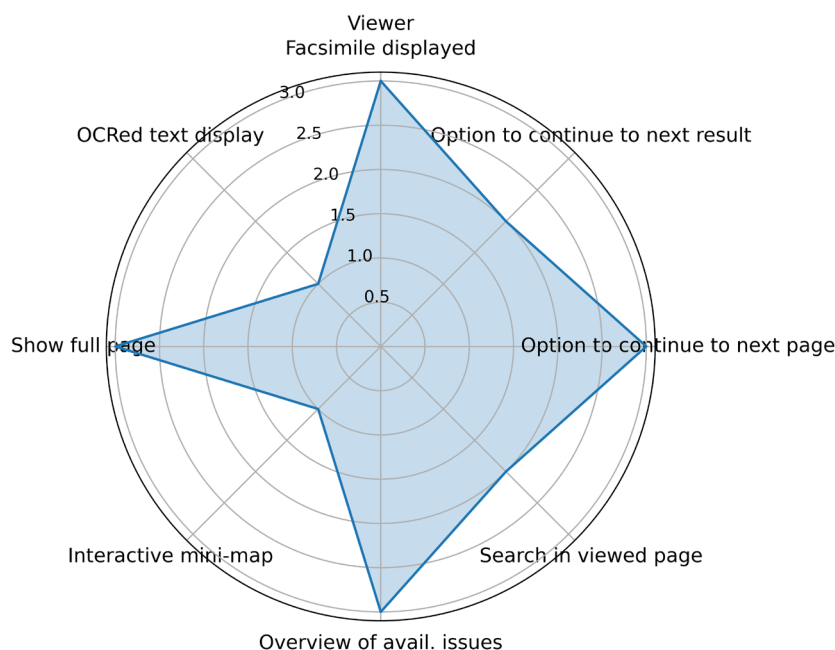
As funcionalidades agrupadas na família *Metadados dos periódicos* está bem representada, com destaque para *Local de publicação*, *período de cobertura* e *indicação da instituição responsável*. Outras funcionalidades, como *Títulos alternativos*, *Editores*, *Links externos* e *Visualização em calendário*, estiveram presentes em duas das três interfaces. Por outro lado, algumas características são pouco exploradas ou inexistentes: *ISSN*, *OCLC* e *LCCN* não aparecem em nenhuma interface, enquanto *Cobertura geográfica*, *Periodicidade*, *Descrição histórica* e *Idioma* registraram apenas uma em três (ver Gráfico 1).

GRÁFICO 1. Dados globais da família *Metadados de Periódicos* para todas as IGUs



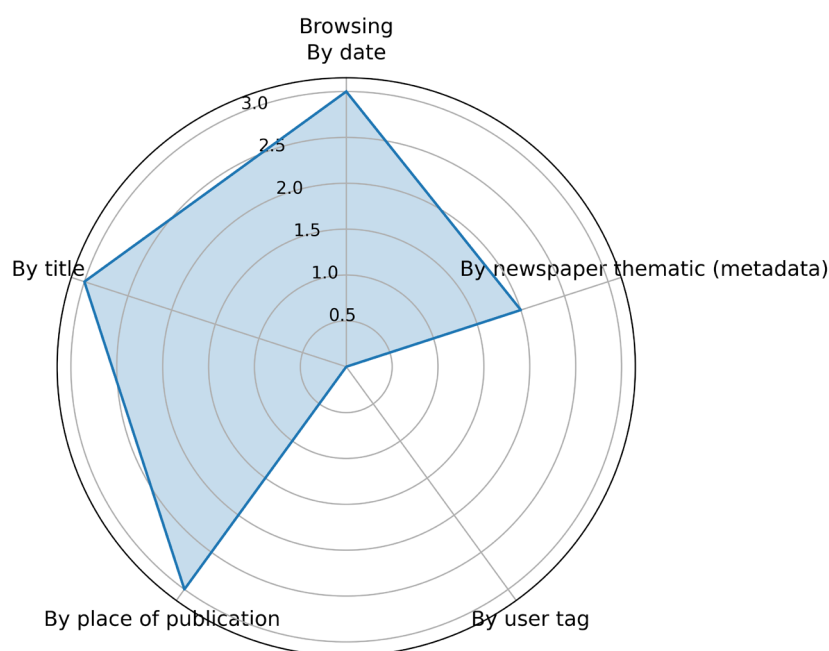
Fonte: Elaborado pelo autor

A família *Visualização* apresenta uma boa representação de funcionalidades, com pelo menos uma ocorrência em todas as oito funcionalidades listadas. Dentre elas, *Exibição de fac-símile*, *Mostrar página completa*, *Visão geral de edições disponíveis* e *Opção de continuar para a próxima página* são comuns às três IGUs. A *Opção de continuar para próximo resultado* e *Busca em páginas visualizadas* estão presentes em duas IGUs e *Exibição do texto de OCR* e *Minimapa interativo* constam em uma das três (ver Gráfico 2).

GRÁFICO 2. Dados globais da família *Visualização* para todas as IGUs

Fonte: Elaborado pelo autor

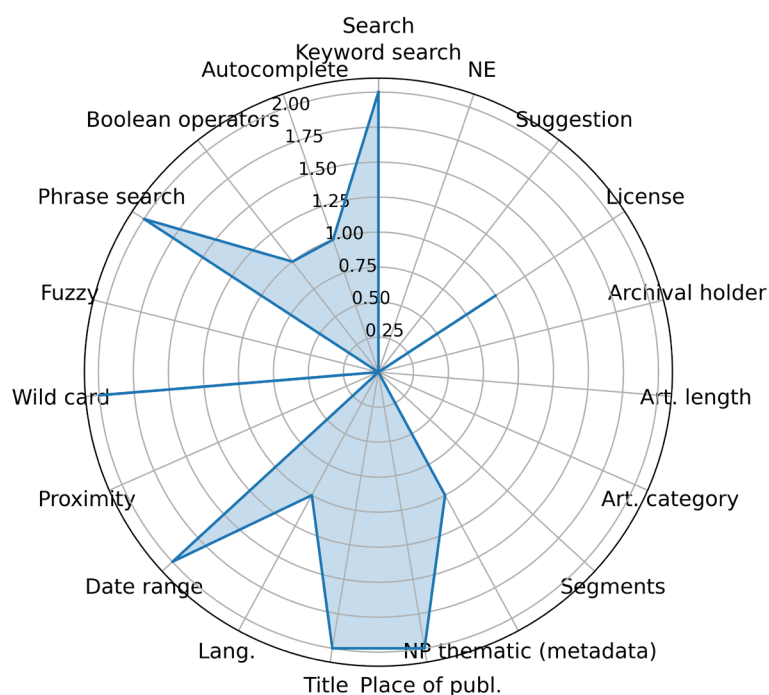
A família *Navegação* apresenta um padrão de organização técnica comum aos três casos estudados: a navegação se dá primordialmente por *data*, *título* e *local de publicação*. Outras opções de navegação, como *temática* (apenas uma ocorrência) e *tags* de usuários (nenhuma ocorrência), que favoreceriam um tipo de navegação mais voltado para aspectos qualitativos e de conteúdo, não estão presentes nas IGUs (ver Gráfico 3).

GRÁFICO 3. Dados globais da família *Navegação* para todas as IGUs

Fonte: Elaborado pelo autor

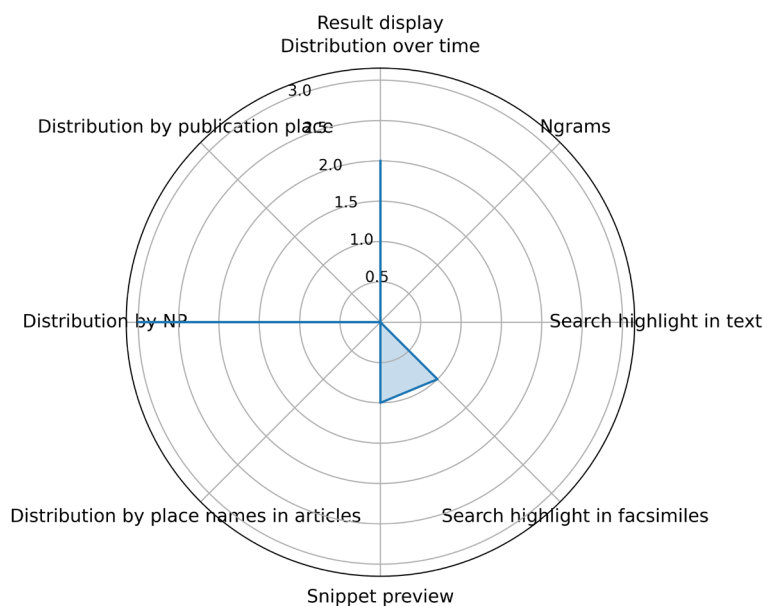
A análise da família *Busca* nos mostra uma significativa disparidade de ocorrências no conjunto de funcionalidades. *Busca por palavra, por expressão exata, limite por recorte temporal, por local de publicação e por título e uso de wildcard* estão presente com mais frequência (em pelo menos duas das três IGUs). No entanto, um número considerável de funcionalidades apresenta ausência total, como *Busca por Entidade Nomeada, Fuzzy e por proximidade, limite por extensão e categoria de artigos, limite por segmento e instituição mantenedora e sugestões de busca*. Algumas funcionalidades, como *autocomplete de buscas, operadores Boolean, limite por licença, linguagem e temático* (a partir dos metadados), têm uma presença pequena, de uma das três IGUs (ver Gráfico 4).

GRÁFICO 4. Dados globais da família *Busca* para todas as IGUs



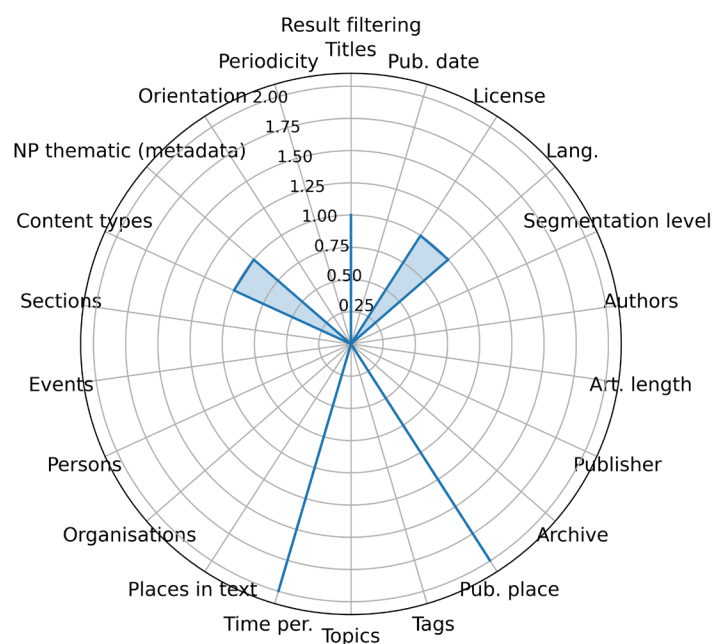
Fonte: Elaborado pelo autor

A família *Exibição dos resultados* apresenta oito funcionalidades, das quais apenas quatro possuem alguma ocorrência. A funcionalidade *Distribuição por Periódico* está presente em todas as IGUs, enquanto *Distribuição ao longo do tempo* aparece em duas delas. Já as funcionalidades *Snippetpreview* e *highlight da busca no fac-símile* ocorrem apenas uma vez. As demais funcionalidades – *Distribuição pelo local de publicação, Distribuição por nome de lugares nos artigos, Highlight da busca no texto* e *Ngrams* – estão ausentes (ver Gráfico 5).

GRÁFICO 5. Dados globais da família *Exibição de Resultados* para todas as IGUs

Fonte: Elaborado pelo autor

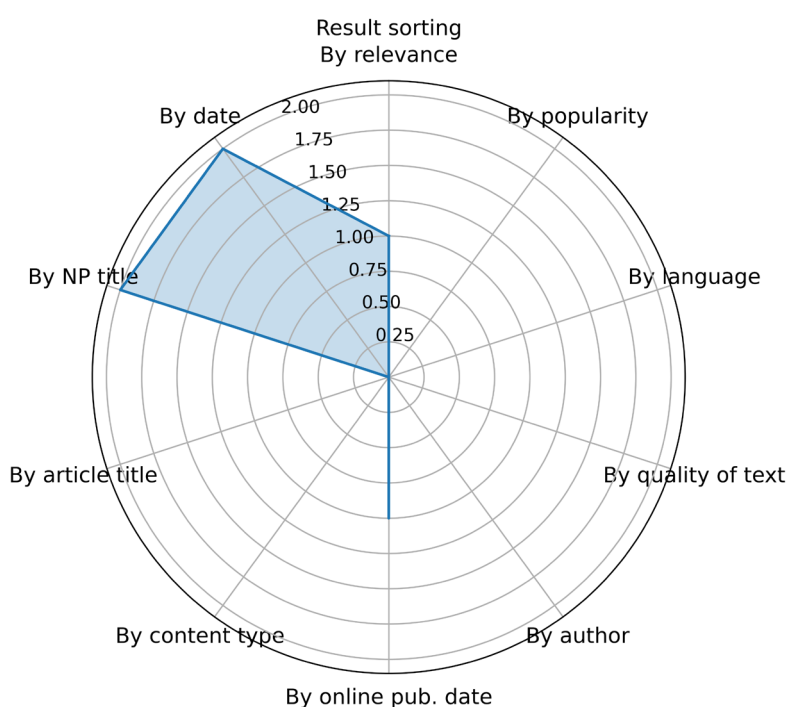
Os *Filtros de Resultados* apresentam uma implementação ainda mais limitada: das 22 funcionalidades listadas, apenas sete possuem ocorrências. Os filtros por *local de publicação* e *período temporal* são os mais comuns, estando presentes em duas das três IGUs. Já os filtros por *licença*, *idioma*, *tipo de conteúdo* e *temática* (a partir dos metadados) aparecem apenas uma vez cada. Por outro lado, a maioria das funcionalidades avaliadas está ausente, incluindo opções relevantes, como filtro por *autores*, *periodicidade*, *tópicos* e *editoras*, entre outros (ver Gráfico 6).

GRÁFICO 6. Dados globais da família *Filtro de Resultados* para todas as IGUs

Fonte: Elaborado pelo autor

Das dez funcionalidades agrupadas na família *Classificação de resultados*, apenas quatro estão contempladas. As classificações por *data* e por *título* são as mais comuns, estando presentes em duas das três IGUs. Já as classificações por *relevância* e por *data de publicação online*, também aparecem, mas apenas uma vez entre as três possíveis. Por outro lado, as funcionalidades de classificação por *título do artigo*, *tipo de conteúdo*, *autor*, *qualidade do texto*, *idioma* e *popularidade* estão completamente ausentes (ver Gráfico 7).

GRÁFICO 7. Dados globais da família *Classificação de Resultados* para todas as IGUs



Fonte: Elaborado pelo autor

As próximas famílias de funcionalidades apresentam baixíssima presença nas IGUs estudadas; algumas, sem nenhuma ocorrência, o que revela um padrão significativo para nossa análise sobre como tais interfaces são concebidas e como podem influenciar os direcionamentos de acesso e as possibilidades de trabalho com seus dados por parte dos pesquisadores.

Na família *Informações sobre a digitalização*, apenas a funcionalidade *Informação sobre Copyright* está presente em duas das IGUs. As demais nove funcionalidades não estão contempladas, incluindo elementos essenciais para a avaliação heurística desses acervos, como *Reconhecimento de layout no nível de artigos*, *Pontuação de confiança de OCR e OLR*, *Documentação sobre vieses*, *Métodos de digitalização*, *Pontuação sobre relevância dos resultados*, *Ferramentas de OCR utilizadas* e *Resolução do scan*.

A ausência dessas informações compromete a transparência e a confiabilidade dos acervos, limitando as possibilidades de análise crítica dos pesquisadores e dificultando a compreensão sobre os processos técnicos que estruturam esses repositórios digitais.

De modo semelhante, a família *Interação do Usuário* revela que apenas a funcionalidade *Links Permanentes*, presente em duas das três IGUs, está contemplada, enquanto as demais treze não estão implementadas. Entre essas ausências, destacam-se funcionalidades essenciais para a personalização e gestão do conteúdo, como *Salvar artigos e buscas nos favoritos*, *Etiquetar artigos*, *Acompanhar materiais visualizados*, *Recomendações de artigos*, *Exportar citação*, *Opção para corrigir OCR*, *Adicionar/editar metadados*, *Ferramenta de captura de tela*, *Downloads em massa*, *Organizar artigos em coleções* e *Visão contrastiva de coleções*. A inexistência dessas opções limita significativamente o potencial de interação das interfaces, restringindo a capacidade do usuário de organizar, editar e acessar.

Dentre as funcionalidades de *Conectividade*, apenas a *Identificação de terceiros* está presente em uma interface, enquanto *Links para outros repositórios* e *Tecnologias de Web Semântica* não foram encontrados. A ausência de conectividade compromete o potencial de integração dos acervos com repositórios externos e ferramentas avançadas, reduzindo o alcance das interfaces para pesquisas colaborativas ou multidimensionais.

As famílias *Enriquecimento* e *APIs* estão completamente ausentes das IGUs das hemerotecas digitais em português. Entre as funcionalidades de *Enriquecimento*, não implementadas, estão *Reconhecimento e Classificação de Entidades Nomeadas*, *Vinculação de Entidades*, *Correção Automática de OCR*, *Correção de OCR Colaborativa*, *Modelagem de Tópicos*, *Reutilização de Texto*, *Análise de Sentimento*, *Recomendações* e *Deteção de Eventos*.

Já na família *APIs*, as funcionalidades *Link para o código-fonte da interface*, *API*, *IIIF Image API* e *IIIF Presentation API* não estão acessíveis diretamente na IGU. Essa ausência representa uma limitação significativa em termos de acessibilidade técnica e integração das interfaces com outras ferramentas e sistemas. APIs são componentes fundamentais para permitir o acesso automatizado aos dados, possibilitando a criação de aplicações personalizadas e o uso avançado dos acervos. A implementação de APIs, especialmente as baseadas no padrão *International Image Interoperability Framework*¹¹ (IIIF), traria benefícios consideráveis para pesquisadores e desenvolvedores, ao viabilizar a visualização e manipulação de imagens digitalizadas de maneira padronizada e interoperável.

A BND-PT adota o padrão *IIIF* em seus objetos digitais, oferecendo vantagens importantes em relação à Hemeroteca Digital Brasileira (HDB), que ainda não implementa esse recurso. No entanto, a BND-PT não oferece, de formas explícita e direta, acesso aos dados relacionados ao *IIIF* em sua interface. Portanto, não há um caminho detalhado e documentação para

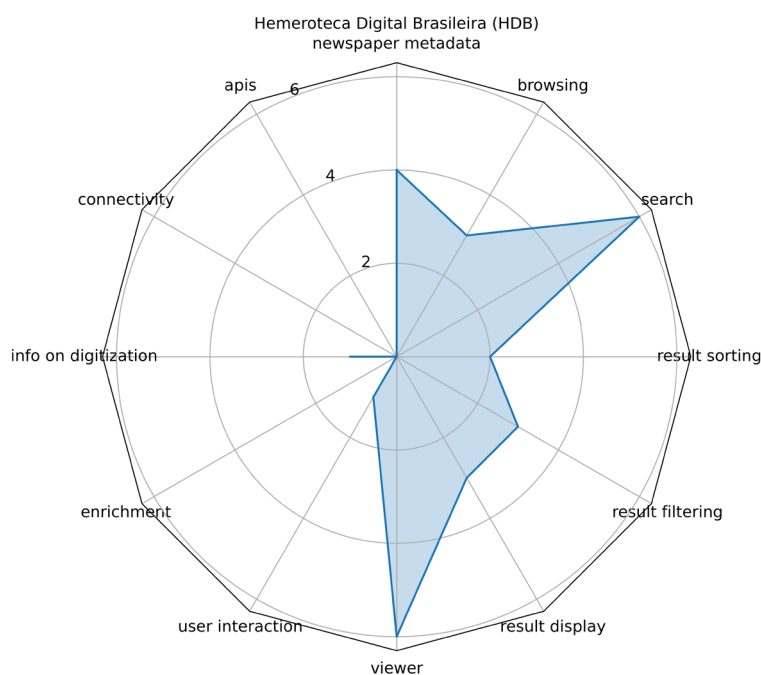
¹¹ O *International Image Interoperability Framework* (IIIF) é um conjunto de padrões abertos que permite o acesso, a exibição e a manipulação avançada de imagens digitalizadas em alta resolução. Desenvolvido para facilitar a interoperabilidade entre repositórios e aprimorar a experiência do usuário, o *IIIF* possibilita que acervos digitais compartilhem imagens de forma estruturada, permitindo sua integração com diversas ferramentas e plataformas sem a necessidade de adaptação específica para cada sistema. Esse padrão é amplamente adotado por bibliotecas, arquivos e museus ao redor do mundo, pois viabiliza funcionalidades, como *zoom* dinâmico, comparação lado a lado de documentos de diferentes instituições e anotação colaborativa. Ao adotar o *IIIF*, acervos digitais ganham maior flexibilidade e acessibilidade, promovendo novas possibilidades para pesquisas em grande escala e a análise computacional de documentos históricos.



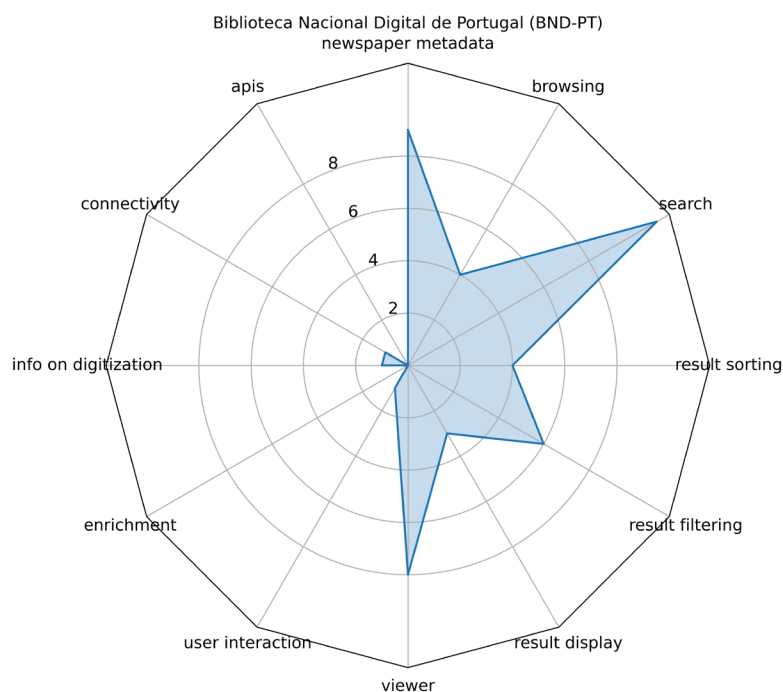
utilização desses recursos, nem acesso ao manifesto da imagem, tampouco um método eficiente para a coleta massiva desses dados. Essa característica compromete a exploração plena do potencial do *IIIF*, restringindo as possibilidades de pesquisa automatizada e de análises que dependem da integração de grandes volumes de dados. Assim, embora a BND-PT avance em conectividade técnica, ainda enfrenta desafios na acessibilidade prática de seu acervo em larga escala.

Ao analisarmos cada interface, percebemos uma acentuada semelhança entre a Hemeroteca Digital Brasileira (HDB) (Gráfico 8) e a Biblioteca Nacional Digital de Portugal (BND-PT) (Gráfico 9), com o formato do gráfico radial das famílias de funcionalidades apresentando padrões próximos. Ambas possuem um forte desenvolvimento em funções de busca, navegação, visualização e informação sobre metadados dos periódicos, além de algumas funcionalidades relativas aos resultados, como classificação, exibição e filtros. Já a interface da Hemeroteca Digital de Lisboa (HDL) é a mais limitada em recursos, apresentando apenas implementações voltadas para metadados dos jornais, navegação e visualização (Gráfico 10).

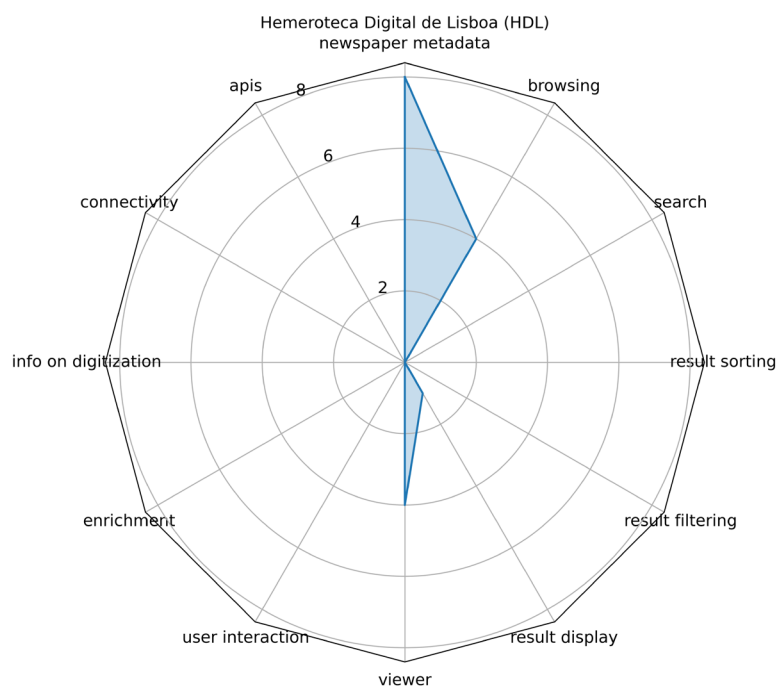
GRÁFICO 8. Gráfico radial das famílias de funcionalidades da HDB



Fonte: Elaborado pelo autor

GRÁFICO 9. Gráfico radial das famílias de funcionalidades da BND-PT

Fonte: Elaborado pelo autor

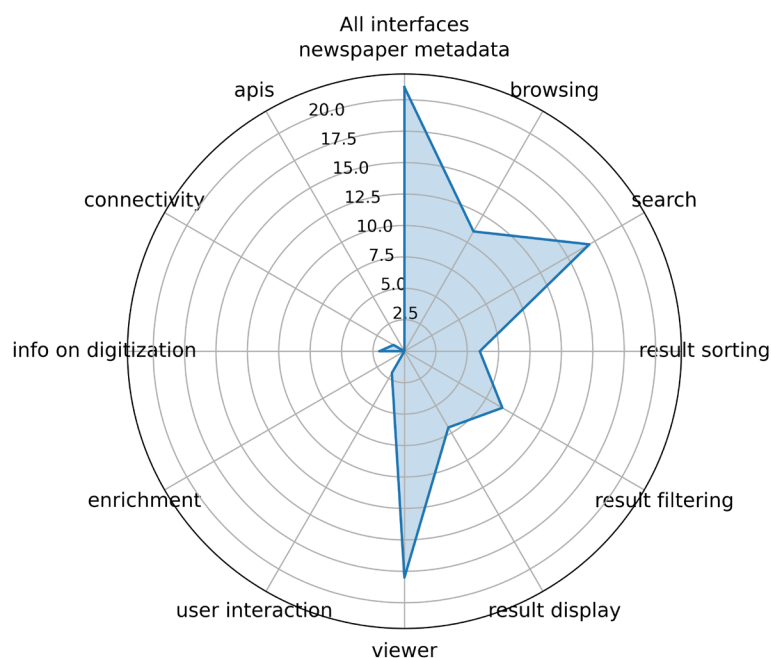
GRÁFICO 10. Gráfico radial das famílias de funcionalidades da HDL

Fonte: Elaborado pelo autor

É importante destacar, também, as ausências comuns às três interfaces. Recursos relativos à conectividade, APIs e enriquecimento constituem lacunas acentuadas nos casos estudados, assim como informações sobre o processo de digitalização e interações com os usuários. Essa

ausência é sintomática, pois essas categorias são importantes para a realização de pesquisas complexas, a aplicação de técnicas avançadas de análise computacional, a viabilização de ações de interoperabilidade e reprodutibilidade e o estabelecimento de um registro metodológico rigoroso dos processos de pesquisa.

GRÁFICO 11. Gráfico global das famílias de funcionalidades de todas as IGUs



Fonte: Elaborado pelo autor

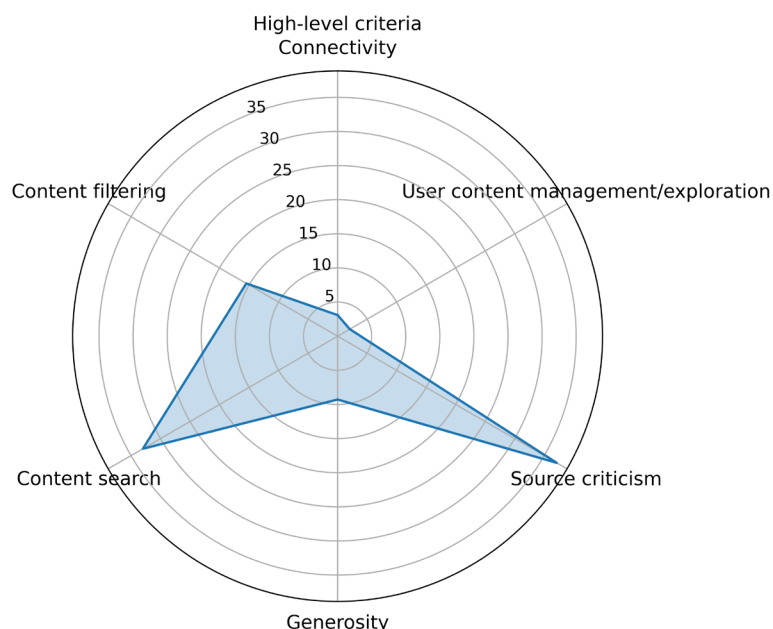
A análise evidencia que, embora as interfaces atendam bem às funcionalidades básicas de busca, elas apresentam limitações significativas nas opções mais avançadas, comprometendo a possibilidade de explorar plenamente os acervos por meio de ferramentas mais sofisticadas. Além disso, há uma carência de recursos voltados para a compreensão da própria estrutura da interface, de seus métodos de busca, OCR e apresentação dos resultados, dificultando uma navegação mais informada e eficiente para os pesquisadores.

5.1. Critérios de avaliação de alto nível

Ao analisarmos os Critérios de Avaliação de Alto Nível, que incluem Crítica das Fontes (avalia a transparência das informações sobre documentos digitalizados, metadados e processos de digitalização), Busca de Conteúdo (mensura a amplitude e sofisticação das ferramentas de pesquisa), Filtro de Conteúdo (examina os mecanismos de refinamento de resultados), Conectividade (verifica a integração com outros repositórios), Gerenciamento/Exploração de Conteúdo pelo Usuário (permite ao pesquisador organizar e anotar material) e Generosidade (oferece caminhos alternativos e mais ricos de descoberta de conteúdos), constatamos, a partir do gráfico radial, uma forte implementação de funcionalidades associadas aos três primeiros critérios. Isso

se manifesta na variedade de recursos disponíveis para checar a procedência dos documentos, realizar pesquisas em grandes volumes de dados e refinar resultados (Gráfico 12).

GRÁFICO 12. Gráfico radial dos Critérios de Avaliação de Alto Nível

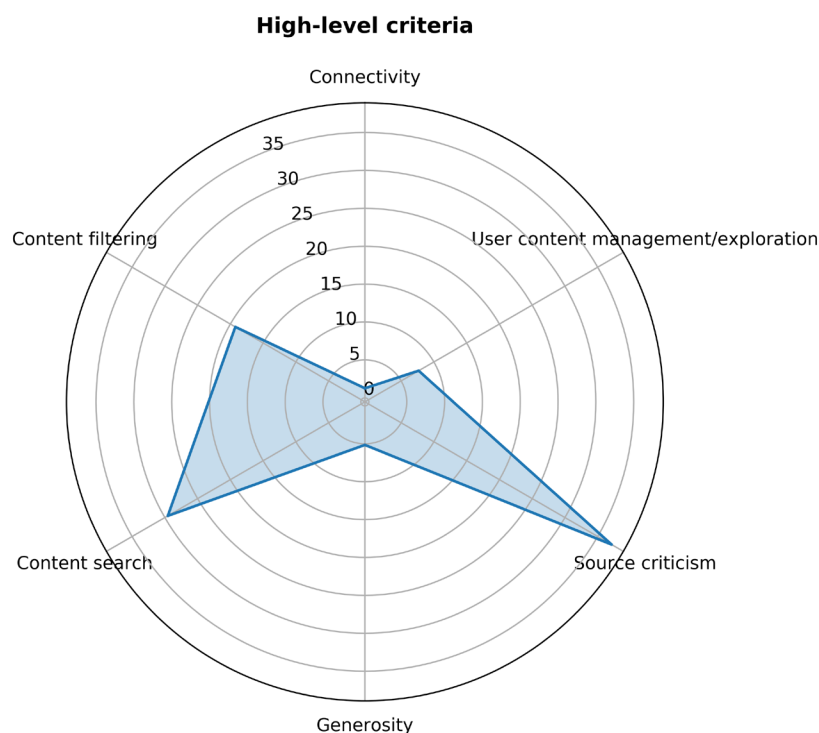


Fonte: Elaborado pelo autor

Em contrapartida, as dimensões de Conectividade, Gerenciamento/Exploração e Generosidade apresentam-se substancialmente limitadas. O mapeamento detalhado das famílias de funcionalidades indica que essas lacunas impactam diretamente a experiência do usuário, restringindo o compartilhamento e a interação (Conectividade), a curadoria pessoal do acervo (Gerenciamento/Exploração) e as possibilidades de apropriação autônoma e inovadora dos materiais (Generosidade). Esse cenário evidencia a necessidade de aprimoramentos estruturais, capazes de fomentar maior colaboração, personalização e integração das hemerotecas digitais em ecossistemas de pesquisa mais amplos.

O gráfico radial da pesquisa de Ehrmann; Bunout e Düring (2019, p. 17), em comparação com o gráfico do nosso estudo, revela uma estrutura análoga (Gráfico 13). Em ambos os casos, funcionalidades, como busca por texto completo, filtros básicos (por título ou data) e visualização de fac-símiles aparecem como as mais consolidadas, enquanto aspectos relacionados à conectividade (incluindo integração entre coleções e uso de APIs) e ao enriquecimento semântico permanecem menos desenvolvidos. A transparência dos processos de digitalização e OCR, bem como a documentação de metadados e potenciais vieses, também se mostra ausente com frequência nas duas amostras, dificultando a avaliação crítica das fontes. Essas tendências recorrentes – ainda que decorram de contextos institucionais distintos – reforçam a conclusão de que hemerotecas digitais, de maneira geral, precisam superar barreiras tecnológicas e epistemológicas para oferecer uma experiência de pesquisa mais robusta e confiável.

GRÁFICO 13. Gráfico radial dos Critérios de Avaliação de Alto Nível da pesquisa de Ehrmann; Bunout e Düring (2019)



Fonte: Ehrmann; Bunout e Düring (2019)

Os resultados do estudo indicam que as interfaces gráficas de usuário das hemerotecas digitais em português situam-se na primeira geração de interfaces, conforme a categorização proposta por Ehrmann; Bunout e Düring (2019, p. 12). O foco recai na disponibilização de material textual bruto e metadados básicos já existentes em registros bibliográficos, oferecendo apenas recursos essenciais de busca e organização de resultados. Embora essas interfaces representem um avanço relevante para o acesso e preservação dos acervos, ainda carecem de funcionalidades que permitam análises historiográficas aprofundadas.

A segunda geração de interfaces amplia a interação com os usuários, incorporando recursos, como *links* permanentes (*permalinks*), salvamento de favoritos, agrupamento de itens em coleções e, em alguns casos, correção colaborativa de OCR. Já a terceira geração, destaca-se pela incorporação de funcionalidades de enriquecimento semântico, como reconhecimento de entidades, além da oferta de APIs para consultas programáticas, o que expande as possibilidades de pesquisa automatizada e avançada. Por fim, a quarta geração, ainda em desenvolvimento em diversos projetos, busca promover maior transparência sobre o processo de digitalização e enriquecer a experiência do usuário por meio de mecanismos de recomendação e sugestões de pesquisa personalizadas, criando interfaces que atendam tanto ao público geral quanto a especialistas em história digital (2019, p. 12-15).

6. Pelo desenvolvimento de hemerotecas digitais 2.0

A análise das interfaces de hemerotecas digitais em português evidencia como as escolhas técnicas, muitas vezes subestimadas, impactam diretamente as possibilidades de pesquisa em história (Gil, 2024). Funcionalidades, como filtros avançados, integração com APIs e ferramentas de enriquecimento semântico, poderiam ampliar significativamente as abordagens metodológicas, mas permanecem insuficientes em grande parte dessas plataformas.

As limitações técnicas das hemerotecas digitais em português acabam reforçando, paradoxalmente, o uso de práticas analógicas em ambientes digitais. Em vez de estimular a adoção de metodologias inovadoras, essas plataformas frequentemente engendram a lógica tradicional de consulta, na qual o pesquisador acessa documentos de forma individualizada, com pouca ou nenhuma mediação por recursos tecnológicos avançados. A ausência de ferramentas de enriquecimento, conectividade com outros repositórios e suporte à análise automatizada de dados faz com que os historiadores recorram a métodos manuais, replicando estratégias analógicas no meio digital, em vez de desenvolverem abordagens criativas e metodologicamente inovadoras diante do potencial dos dados digitais.

A ausência de mecanismos de reconhecimento automático de padrões ou de modelagem de tópicos exige que os pesquisadores organizem, categorizem e interpretem grandes volumes de dados manualmente, resultando em um processo lento e trabalhoso, que retarda a inovação na pesquisa histórica. As limitações nos recursos de interação do usuário com as interfaces – como a ausência de ferramentas para gerenciamento de coleções ou marcação de artigos – impossibilitam a personalização e a exploração aprofundada dos acervos, ao mesmo tempo em que dificultam a reprodutibilidade das pesquisas. Esse cenário reforça a percepção equivocada de que acervos digitais são meras extensões dos arquivos físicos, em vez de plataformas transformadoras, capazes de ampliar significativamente as possibilidades de investigação histórica.

Diante desse quadro, torna-se urgente um salto qualitativo rumo às Hemerotecas Digitais 2.0. Essa nova geração de plataformas deve ultrapassar a simples disponibilização de fontes digitalizadas com buscas por palavras-chave para incorporar avanços em conectividade, enriquecimento semântico e experiência do usuário. A emergência de plataformas, como o *Chronicling America* (EUA), o projeto *Europeana Newspapers* (União Europeia) e o *impresso* (Suíça/França), ilustra alguns caminhos possíveis.

A coleção *Chronicling America*, mantida pela *Library of Congress* em parceria com o *National Endowment for the Humanities* (NEH), é um exemplo de interface que está se encaminhando para se tornar uma *Hemeroteca Digital 2.0*, com avanços tanto em infraestrutura técnica quanto em usabilidade e abertura à pesquisa computacional. Em 2024, foi lançada uma nova documentação voltada ao uso de sua API, integrada à arquitetura do *site LoC.gov*, que permite acesso automatizado e em larga escala a metadados e páginas digitalizadas de jornais históricos norte-americanos. Essa API pública, sem necessidade de autenticação, é acompanhada por tutoriais detalhados e *notebooks* Jupyter voltados a pesquisadores e desenvolvedores, disponíveis no repositório do GitHub da própria instituição. Além disso, a plataforma está em processo de migração para uma nova interface com infraestrutura de busca redesenhada, cuja versão preli-



minar já se encontra disponível, prometendo maior eficiência na recuperação da informação e melhor experiência de navegação.¹²

A *Europeana*, além de um repositório de dados, atua como uma plataforma dinâmica e colaborativa que agrega acervos de milhares de instituições culturais europeias, estruturando um ecossistema digital que articula padrões técnicos, ferramentas de interoperabilidade (como APIs e metadados padronizados) e políticas públicas de digitalização do patrimônio. Especificamente no caso da coleção de jornais, esses princípios se concretizam por meio da oferta de visualizações interativas, mecanismos de enriquecimento semântico, integração com projetos agregadores nacionais e domínio público, além de políticas voltadas ao reuso criativo e científico dos dados (Capurro e Severo, 2023).

Já discutido anteriormente no artigo, o projeto *Impresso - Media Monitoring of the Past* constitui um exemplo emblemático de Hemeroteca Digital 2.0, ao integrar técnicas avançadas de mineração de textos históricos com uma interface projetada para o trabalho exploratório em humanidades digitais. Desenvolvido por um consórcio de instituições suíças, o *impresso* se destaca por disponibilizar *corpora* de jornais históricos acompanhados de ferramentas analíticas baseadas em enriquecimento semântico (como reconhecimento de entidades nomeadas e análise de tópicos), além de visualizações interativas e acesso programático via API.

Essas iniciativas têm investido em APIs, enriquecimento semântico, interoperabilidade e transparência nos processos técnicos, oferecendo modelos relevantes para pensar o futuro das hemerotecas em português. Sob essa perspectiva, a noção de Hemeroteca Digital 2.0 pressupõe o desenvolvimento colaborativo e multidisciplinar de ferramentas e funcionalidades que vão desde APIs interoperáveis e tecnologias da *web* semântica até mecanismos de enriquecimento automatizado de conteúdo e transparência técnica.

No entanto, essas transformações não dependem apenas de inovações tecnológicas. É fundamental promover debates sobre políticas de financiamento público que incentivem a colaboração entre historiadores, profissionais das instituições arquivísticas e bibliotecas, linguistas e especialistas em tecnologia. A valorização de iniciativas abertas e gratuitas reforça o compromisso com a democratização do acesso e a preservação do patrimônio documental, ao mesmo tempo em que amplia o potencial das ciências humanas no meio digital. Em última instância, a modernização das hemerotecas digitais em português constitui não apenas um desafio técnico, mas também uma oportunidade para repensar e aprofundar o papel dessas plataformas na pesquisa histórica, consolidando-as como ferramentas indispensáveis para o avanço das humanidades no século XXI.

CONFLITO DE INTERESSES

O autor declara não haver conflito de interesses relacionado a este artigo.

¹² Disponível em: <<https://www.loc.gov/collections/chronicling-america/about-this-collection/>>. Acessado em: 08 de abril de 2025.

FINANCIAMENTO

Este trabalho é um dos resultados da pesquisa de pós-doutorado desenvolvida no Laboratório de Humanidades Digitais do Instituto de História Contemporânea (IHC), da Faculdade de Ciências Sociais e Humanas da Universidade NOVA de Lisboa, integrante do IN2PAST – Laboratório Associado para a Investigação e Inovação em Património, Artes, Sustentabilidade e Território. A pesquisa, entretanto, não contou com financiamento específico de agência de fomento.

REFERÊNCIAS

- BETTENCOURT, Angela Maria Monteiro; PINTO, Monica Rizzo Soares. A hemeroteca digital brasileira. **Anais do Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação – FEBAB**, v. 25, n. 0, p. 1028-1038, 2013.
- BRASIL, Eric; NASCIMENTO, Leonardo Fernandes. História digital: reflexões a partir da Hemeroteca Digital Brasileira e do uso de CAQDAS na reelaboração da pesquisa histórica. **Revista Estudos Históricos**, v. 33, n. 69, p. 196-219, 2020.
- BRASIL, Eric; VALVERDE, Priscila; VELOSO, Ana Carolina. Desafios e expectativas da prática da história na Era Digital: percepções de docentes de história nas instituições de ensino superior na Bahia. **Boletim do Tempo Presente**, v. 13, n. 2, p. 140-171, 2024.
- BRASIL, Eric. pyHDB – Ferramenta Heurística para a Hemeroteca Digital Brasileira: utilizando técnicas de *web scraping* para a pesquisa em História. **História da Historiografia: International Journal of Theory and History of Historiography**, v. 15, n. 40, 2022.
- BUNOUT, Estelle; EHRMANN, Maud; CLAVERT, Frédéric. Digitised Historical Newspapers: A Changing Research Landscape – Introduction. In: **Digitised Newspapers – A New Eldorado for Historians?: Reflections on Tools, Methods and Epistemology**. Berlin: De Gruyter Oldenbourg, 2022. (Studies in Digital History e Hermeneutics, v. 3).
- CAMPOS, Fernanda Maria; CABRAL, Rosa Maria Galvão Maria Luísa; BORBINHA, José Luis. The National Library of Portugal. **Alexandria: The Journal of National and International Library and Information Issues**, v. 14, n. 2, p. 83-101, 2002.
- CAPURRO, Carlotta; SEVERO, Marta. Mapping European Digital Heritage Politics: An Empirical Study of Europeana as a Web-based Network. **Heritage&Society**, v. 0, n. 0, p. 1-21, 2023.
- EHRMANN, Maud; ROMANELLO, Matteo; CLEMATIDE, Simon; STRÖBEL, Phillip Benjamin; BARMAN, Raphaël. Language Resources for Historical Newspapers: the Impresso Collection. In: **Proceedings of the Twelfth Language Resources and Evaluation Conference**, p. 958-968, Marseille, France. European Language Resources Association, 2020.
- EHRMANN, Maud; BUNOUT, Estelle; DÜRING, Marten. Historical Newspaper User Interfaces: A Review. In: **Historical Newspaper User Interfaces**, 2019, Athens, Greece. **Proceedings of the 85th IFLA General Conference and Assembly (IFLA)**. Athens, Greece: IFLA Library, 2019.
- GIL, Tiago. Entre teoria, técnica e imaginação: consumo e produção de motores de busca por historiadores, através do caso do projeto “Oxossi”. **Acervo**, v. 37, n. 3, p. 1-20, 2024.



GIORDANO, Rafaela Boeira. **Do jornal à ciência: a hemeroteca digital brasileira como fonte de informação para a pesquisa científica**. 2016. Doutorado - Instituto Brasileiro de Informação em Ciência e Tecnologia/ Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.

MARINO, Ian Kisil et al. Como contar a história da COVID-19? Reflexões a partir dos arquivos digitais no Brasil. **Esboços: histórias em contextos globais**, v. 28, n. 48, p. 558-583, 2021.

MATOS, Álvaro Costa de. **A Digitalização do Acervo Documental da Hemeroteca Municipal de Lisboa: Uma Primeira Abordagem Ao Suporte Electrónico, a Partir Do Jornal Os Ridículos**. Lisboa, 2001.

MEIRELLES, Raissa Félix. **A Biblioteca Nacional Digital e a premência de realização de estudo de usuários**. Orientadora: Luana Farias Sales. Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2020.

MILLIGAN, Ian. Illusionary Order: Online Databases, Optical Character Recognition, and Canadian History, 1997-2010. **The Canadian Historical Review**, v. 94, n. 4, p. 540-569, 2013.

OLIVEIRA, João Carlos Salvador da Silva de; MATOS, Álvaro Costa de. Hemeroteca Municipal de Lisboa: Património Projetado No Futuro. **Cadernos de Biblioteconomia, Arquivística e Documentação**, n. 1, 2015.

PATRÍCIO, Helena. Portugal: Enhancing Digital Services at the National Library of Portugal. **Uncommon Culture**, p. 136-143, 2011.

PFANZELTER, Eva et al. Digital Interfaces of Historical Newspapers: Opportunities, Restrictions and Recommendations. **Journal of Data Mining & Digital Humanities**, v. HistoInformatics, n. HistoInformatics, 2021.

VELOSO, Ana Carolina de Oliveira. **A Hemeroteca Digital Brasileira na pesquisa histórica: uma análise dos simpósios nacionais de História da ANPUH (2013-2019)**. 2022. (Trabalho de Conclusão de Curso) – Universidade da Integração da Lusofonia Afro-brasileira, São Francisco do Conde, BA, 2022.