

Preservação Digital no Brasil: construção de uma historiografia bibliométrica a partir da produção científica

Digital Preservation in Brazil: building a bibliometric historiography from scientific production

Edilma Maria Machado¹, Natanael Vitor Sobral², Thais Helen do Nascimento Santos³, Daniela Eugênia Moura de Albuquerque⁴

¹ Universidade Federal de Pernambuco (UFPE), Recife – PE, Brasil. Bacharela em Biblioteconomia. ORCID: <https://orcid.org/0009-0001-8872-9239>

² Programa de Pós-Graduação em Ciência da Informação (PPGCI/UFPE), Recife – PE, Brasil. Doutor em Ciência da Informação. ORCID: <https://orcid.org/0000-0003-2410-494X>

³ Programa de Pós-Graduação em Informação e Comunicação em Plataformas Digitais (PPGCI/UFPE), Recife – PE, Brasil. Doutora em Informação e Comunicação em Plataformas Digitais. ORCID: <https://orcid.org/0000-0001-7091-3225>

⁴ Programa de Pós-Graduação em Ciência da Informação (PPGCI/UFPE), Recife – PE, Brasil. Doutoranda em Ciência da Informação. ORCID: <https://orcid.org/0000-0002-1136-8965>

Autor para correspondência/Mail to: Natanael Vitor Sobral, natanvsobral@gmail.com

Recebido/Submitted: 04 de junho de 2025; **Aceito/Approved:** 16 de novembro de 2025



Copyright © 2025 Sobral et al. Todo o conteúdo da Revista (incluindo-se instruções, política editorial e modelos) está sob uma licença Creative Commons Atribuição 4.0 Internacional. Ao serem publicados por esta Revista, os artigos são de livre uso para compartilhar e adaptar e é preciso dar o crédito apropriado, prover um link para a licença e indicar se mudanças foram feitas. Mais informações em <http://revistas.ufpr.br/atoz/about/submissions#copyrightNotice>.

Resumo

Introdução: A preservação digital é fundamental para assegurar a integridade, a acessibilidade e a autenticidade de informações históricas, especialmente diante de desafios como a obsolescência tecnológica e a degradação dos suportes físicos. Nesse cenário, a pesquisa acadêmica tem se destacado como uma ferramenta crucial para a proteção do patrimônio cultural digital. Este artigo tem como objetivo analisar a historiografia bibliométrica da produção científica em periódicos nacionais sobre preservação digital no Brasil. **Método:** Adotou-se uma abordagem descritiva e bibliográfica, com aplicação de técnicas historiográficas e bibliométricas, baseada em dados indexados nas plataformas *Web of Science*, *Scopus* e *Brapi*, a partir de um corpus de 457 artigos. **Resultados:** Os dados indicam um crescimento contínuo da produção científica, com picos significativos nos anos de 2012, 2016 e 2020. Observa-se uma transição de abordagens predominantemente técnicas para temas relacionados à gestão do conhecimento e à formulação de políticas públicas. **Conclusão:** A análise historiográfica bibliométrica evidencia a consolidação da pesquisa em preservação digital no Brasil, sinalizando uma crescente maturidade temática na área.

Palavras-chave: Trajetória temática; produção científica brasileira; historiografia bibliométrica; preservação digital; informação digital.

Abstract

Introduction: Digital preservation is essential to ensure the integrity, accessibility, and authenticity of information, especially in the face of challenges such as technological obsolescence and the degradation of physical media. In this context, academic research has emerged as a crucial tool for protecting digital cultural heritage. This article aims to analyze the bibliometric historiography of scientific production on digital preservation in Brazil. **Method:** A descriptive and bibliographic approach was adopted, applying historiographic and bibliometric techniques based on data indexed in the *Web of Science*, *Scopus*, and *Brapi* platforms, using a corpus of 457 articles. **Results:** The data indicate a continuous growth in scientific production, with significant peaks in the years 2012, 2016, and 2020. A shift is observed from predominantly technical approaches to topics related to knowledge management and public policy development. **Conclusion:** The bibliometric historiographic analysis highlights the consolidation of research on digital preservation in Brazil, signaling a growing thematic maturity in the field.

Keywords: Thematic trajectory; Brazilian scientific production; bibliometric historiography; digital preservation; digital information.

INTRODUÇÃO

A Preservação Digital (PD) é uma área científica interdisciplinar considerada essencial na atualidade, ao se dedicar à garantia da autenticidade e acessibilidade de conteúdos digitais ao longo do tempo. Em um cenário marcado pela constante evolução tecnológica e pela obsolescência dos suportes digitais, a PD busca desenvolver estratégias e métodos que assegurem o acesso contínuo à objetos digitais, sejam eles de natureza administrativa, legal, fiscal ou histórica.

O avanço e a consolidação da PD como campo científico vêm sendo construídos por seus agentes por meio de publicações acadêmicas e técnicas. Compreender como essa produção se organiza, se transforma e se consolida ao longo do tempo é fundamental para entender a trajetória da área e seu papel dentro da ciência. Nesse contexto, a historiografia bibliométrica surge como um método eficaz para mapear essa evolução, permitindo observar tendências e compreender a dinâmica das publicações no campo.

Diante disso, este estudo propõe responder à seguinte questão: Como se estabelece a historiografia bibliométrica da produção científica sobre preservação digital no Brasil? Essa pergunta parte da necessidade de compreender como o conhecimento sobre PD tem sido construído e disseminado no país, especialmente em um contexto onde o acesso e a preservação da informação são cada vez mais desafiadores.

O objetivo desta pesquisa é analisar a historiografia bibliométrica da produção científica sobre preservação digital no Brasil. A partir dessa análise, busca-se oferecer uma visão histórica da evolução temática da área e destacar as transformações nas abordagens de pesquisa ao longo do tempo. Com isso, pretende-se contribuir para o fortalecimento da área, subsidiando o desenvolvimento de políticas, estratégias e tecnologias voltadas à preservação da informação digital.

PRESERVAÇÃO DIGITAL: CONCEITOS E APLICAÇÕES

Ao longo da história, a humanidade demonstrou uma constante preocupação em registrar seu passado, visando preservar memórias e sua identidade cultural. Desde antes da invenção da escrita, já era possível observar registros feitos em cavernas e pinturas rupestres, formas de comunicação utilizadas para expressar a existência humana em épocas distantes. Com o tempo, os métodos de registro evoluíram de desenhos rupestres às tecnologias computacionais atuais, refletindo a necessidade contínua de documentar o saber, adaptando-se aos recursos disponíveis em cada época Santos e Flores (2018).

Nesse processo, os métodos de registro e preservação do conhecimento transformaram-se com os avanços tecnológicos e mudanças sociais. Segundo Lee, Slattey, Lu, Tang, e McCrary (2002), a evolução das civilizações trouxe novos meios de armazenamento, como a seda para escrita e a impressão em papel. Contudo, preservar informações não se resume a produzir registros, mas envolve estratégias que garantam sua autenticidade e acessibilidade. (Cassares, 2000, p. 12), define preservação como “um conjunto de medidas e estratégias de ordem administrativa, política e operacional que contribuem direta ou indiretamente para a preservação da integridade dos materiais”.

Com o surgimento das tecnologias digitais, como observam Lee et al. (2002), a maneira de registrar e preservar informações foi revolucionada. O avanço computacional e das redes de alta velocidade viabilizou a manipulação, disseminação e armazenamento de dados com mais eficiência. Innarelli (2011) aponta que, com a transição para o modelo digital, a informação digital assume papel central na preservação de saberes e tradições culturais. Boeres (2004) reforça a importância de cuidados específicos com a informação digital, assim como com os suportes analógicos.

Nesse cenário, a PD torna-se essencial para manter acessíveis ao longo do tempo os documentos históricos, tradições culturais, dados científicos e memórias coletivas. (Machado, 2015, p. 10), destaca que “uma das maiores preocupações das instituições que abrigam acervos arquivísticos, bibliográficos e museológicos é a preservação visando disponibilizá-los para acesso, pesquisa e deleite às gerações atuais e futuras”. Tais informações fundamentais à compreensão do passado e da identidade social podem se perder com o tempo, seja por desgaste físico ou obsolescência tecnológica.

Embora as tecnologias digitais facilitem a criação, pesquisa e transmissão de dados, também introduzem fragilidades. Claro e Castro-Grau (2023) afirmam que essas tecnologias moldam as sociedades, criando novas estruturas e tendências. Por isso, manter as informações digitais intactas é vital para garantir a continuidade do conhecimento e das tradições culturais.

A PD representa um desafio coletivo que envolve a sociedade, indústrias e a comunidade científica. O ritmo acelerado da inovação exige estratégias eficazes para garantir a autenticidade (identidade¹ e integridade²) e acesso às informações digitais. Gladney (2009) ressalta que as novas tecnologias trazem desafios que demandam soluções específicas, especialmente para preservar a informação acadêmica e cultural. Como apontam Formenton e Gracioso (2020), os objetos digitais enfrentam mudanças rápidas e dependem de tecnologias que podem se tornar obsoletas ou frágeis. De acordo com Ferreira (2006), a PD

[...] consiste na capacidade de garantir que a informação digital permaneça acessível e com qualidades de autenticidade suficientes para que possa ser interpretada no futuro recorrendo a uma plataforma tecnológica diferente da utilizada no momento da sua criação Ferreira (2006) p. 20.

Para mitigar esses riscos, é necessário adotar abordagens teóricas e práticas eficazes. Santos e Flores Santos e Flores (2018), p. 42 afirmam que “dentre estes procedimentos, podem-se destacar as estratégias de migração, emulação, encapsulamento e refrescamento”. Além disso, é fundamental que instituições desenvolvam políticas de PD robustas. Lira e Siebra (2021) identificam ameaças humanas, como a falta de políticas e profissionais especializados, e ameaças tecnológicas, como falhas de hardware e software. A acessibilidade e autenticidade das informações digitais dependem de infraestrutura sólida, profissionais capacitados e tecnologias apropriadas.

A escolha de formatos de arquivos e sistemas adequados também é essencial para garantir o acesso no futuro. A formação contínua de profissionais em PD assegura a aplicação eficaz dessas políticas. Lira e Siebra (2021)

¹Identidade: conjunto dos atributos de um documento arquivístico que o caracterizam como único e o diferenciam de outros documentos arquivísticos Conselho Nacional de Arquivos (2020a), p. 34.

²Integridade: estado dos documentos que se encontram completos e que não sofreram nenhum tipo de corrupção ou alteração não autorizada nem documentada Conselho Nacional de Arquivos (2020b), p. 35.

destacam que políticas eficazes devem considerar os custos de formação e manutenção tecnológica. A crescente produção de informação pode superar a capacidade das tecnologias de preservação, criando um descompasso entre produção e conservação.

A eficácia da preservação exige investigação de elementos técnicos e computacionais — como software, hardware, formatos, modelos, metadados e padrões — para evitar perdas na recuperação de dados. Em outros termos, é um processo que demanda atenção às ameaças do avanço tecnológico e ao armazenamento de longo prazo.

Ainda assim, os suportes digitais permanecem vulneráveis, exigindo estratégias como o armazenamento redundante. Formenton, Castro, Gracioso, Furnival, e Simões (2017) defendem que, além do acesso contínuo, é fundamental garantir a localizabilidade dos objetos digitais, preservando layout, funcionalidades, origem e contexto — elementos essenciais para comprovar a autenticidade no âmbito da PD. Diante desse cenário, torna-se imprescindível refletir sobre métodos eficazes de PD.

Para enfrentar esses desafios, faz-se necessário adotar abordagens teóricas e práticas que viabilizem a PD de forma eficaz ao longo do tempo. Dentre as estratégias mais recorrentes na área, destacam-se técnicas voltadas a assegurar o acesso contínuo aos arquivos digitais, mesmo diante da obsolescência tecnológica. Nesse contexto, o modelo *Open Archival Information System* (OAIS³), estabelecido pela norma ISO 14721, oferece uma estrutura conceitual sólida para o gerenciamento e a operacionalização da PD. Segundo Thomaz e Soares (2004), o objetivo do modelo é:

ampliar a consciência e a compreensão dos conceitos relevantes para a preservação de objetos digitais, especialmente entre instituições não arquivísticas; definir terminologias e conceitos para descrever e comparar modelos de dados e arquiteturas de arquivos; ampliar o consenso sobre os elementos e os processos relacionados à preservação e acesso à informação digital; e criar um esquema para orientar a identificação e o desenvolvimento de padrões Thomaz e Soares (2004), p. 9.

Instituições que mantêm acervos permanentes desempenham um papel central na preservação cultural, pois seus arquivos constituem fontes primárias para o conhecimento e a evolução das sociedades Innarelli (2011). Nesse contexto, a colaboração entre a iniciativa privada, universidades, bibliotecas, arquivos, museus e órgãos governamentais é fundamental para enfrentar os desafios da PD e garantir o acesso contínuo à informação em um mundo cada vez mais digital. Organizações voltadas ao patrimônio cultural e instituições de ensino superior vêm adotando estratégias para gerenciar e preservar materiais digitais de forma duradoura e acessível Formenton e Gracioso (2020).

PROCEDIMENTOS METODOLÓGICOS

O universo da pesquisa abrange a produção de artigos sobre PD no Brasil, com base em dados das plataformas *Web of Science* (WoS), *Scopus* e da Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (Brapci), que reúnem publicações em periódicos nacionais indexadas em bases nacionais e internacionais. A metodologia adotada é de natureza descritiva e bibliográfica, incorporando técnicas historiográficas e bibliométricas. Para alcançar o objetivo proposto, a pesquisa foi estruturada em cinco etapas essenciais, descritas a seguir.

1) Escolha e Acesso à Base de Dados As bases de dados foram selecionadas por sua abrangência, confiabilidade e relevância na análise da produção científica sobre PD. A WoS e a *Scopus* foram escolhidas por sua ampla cobertura multidisciplinar, fundamental para um tema que envolve CI, Tecnologia da informação, Arquivologia, Biblioteconomia e Gestão da informação. Ambas fornecem acesso à periódicos de alta qualidade, essenciais para garantir a credibilidade e profundidade da pesquisa.

Essas bases são reconhecidas por seus rigorosos critérios de indexação, assegurando a inclusão de periódicos de alto impacto, o que contribui para uma visão representativa da produção científica global. Além disso, contemplam publicações de diversos países, incluindo o Brasil, permitindo uma análise integrada dos cenários nacional e internacional. A partir delas, é possível gerar indicadores bibliométricos relevantes, enriquecendo a análise do impacto e da evolução da pesquisa em PD.

Para o contexto nacional, utilizou-se a Brapci, que reúne a principal produção científica brasileira em CI. Sua inclusão complementa a visão internacional, oferecendo uma perspectiva específica sobre as contribuições e particularidades da PD no Brasil. A combinação das três bases proporciona uma análise robusta e abrangente do tema.

O acesso às bases WoS e *Scopus* foi realizado via Portal Periódicos Capes, que disponibiliza importantes bases de dados científicas através de assinaturas realizadas pelo Governo Federal do Brasil, complementada pela indexação de bases de dados gratuitas e/ou de acesso livre. Já o acesso à Brapci foi feito diretamente por meio da plataforma online da base, que é de livre acesso ao público.

³ *Open Archival Information System* (OAIS) é um modelo conceitual que define um repositório digital, identificando o ambiente, os componentes funcionais, suas interfaces internas e externas, os objetos de dados e informações.

2) Recuperação da Informação através de Expressão de Busca Para a realização desta etapa foram utilizados termos, truncadores e operadores booleanos enquanto recursos de recuperação da informação para atingir o propósito estabelecido nesta pesquisa. Foi necessário utilizar a mesma expressão de busca nas três bases de dados analisadas.

- WoS: “*digital preservation*” (*Topic*) and *Article (Document Types)* and *BRAZIL (Countries/Regions)*, resultando em 46 documentos recuperados.
- Scopus: *TITLE-ABS-KEY (“digital preservation”) AND (LIMIT-TO (DOCTYPE, “ar”) AND (LIMIT-TO (AFFILCOUNTRY, “Brazil”)))*, que gerou 68 documentos.
- Brapci: A mesma expressão em português “preservação digital”, filtrando apenas em revistas científicas, recuperando um total de 477 documentos.

A pesquisa não estabeleceu um recorte temporal específico, pois era imprescindível mapear os primeiros artigos, no intuito de atender à dimensão histórica do estudo. Na Scopus e WoS, os filtros aplicados foram limitados ao país e à tipologia documental “artigo científico”, enquanto na Brapci foi utilizado apenas o filtro de publicações em periódicos científicos, sem necessidade de filtrar a nacionalidade.

3) Coleta de dados: Os dados foram coletados em 23 de dezembro de 2024, totalizando 591 registros: 46 da base WoS, 68 da Scopus e 477 da Brapci. No caso da Brapci, foram eliminadas as publicações de periódicos estrangeiros que não contavam com autores brasileiros, a fim de evitar distorções na análise. Para este estudo, considerou-se apenas a produção científica publicada em periódicos brasileiros. Também foram excluídos sete editoriais, inicialmente classificados como artigos de periódico, mas que não correspondiam a essa tipologia, além do “Manifesto para a Preservação Digital” Borbinha, Henriques, Lopes, e Sequeira (2002) e do artigo “A Ciência Aberta na Perspectiva de Especialistas Brasileiros: Proposta de Taxonomia” Silveira, Méndez-Solano, e Mora-Campos (2022), por se tratar de uma tradução de um trabalho já presente na amostra. Após esse processo, consolidou-se um corpus final de 457 registros.

O processo de exportação incluiu os formatos “*plain text file*” (*WoS* e *Scopus*) e “*xls*” (*Brapci*). Em seguida, os registros foram reunidos em planilha única para posterior padronização e criação das matrizes no software VantagePoint® Porter e Cunningham (2004). Ao serem exportadas, algumas publicações foram entregues sem o registro das palavras-chave, sendo a maioria delas da Brapci, com apenas uma ocorrência na WoS e na Scopus. Para resolver essa questão, foi necessário acessar os artigos individualmente e extrair as palavras-chave diretamente das publicações.

4) Padronização e Processamento dos dados: O tratamento dos metadados das publicações foi realizado utilizando o software de mineração de dados e texto VantagePoint Porter e Cunningham (2004). O software permite fazer a correção dos registros, o agrupamento dos dados e a padronização de nomes de pessoas, palavras-chave e outras variáveis presentes na base através do uso de tesouros especializados. Além disto, possibilita a construção de matrizes de coocorrência, permitindo cruzamentos entre diferentes campos da base de dados. Nesta pesquisa, a variável “ano” foi priorizada por sua importância nos estudos historiográficos, permitindo acompanhar a evolução da produção científica ao longo do tempo. O tratamento dos dados incluiu a eliminação de duplicatas e a unificação das palavras-chave. Inicialmente, foram identificados 1.020 registros de palavras-chave, reduzidos a 498 após o processo de junção. Para a tradução desses termos, utilizou-se a ferramenta ChatGPT®. Por fim, os dados foram organizados em matrizes, rankings e no formato .net, possibilitando a construção de redes de palavras-chave alinhadas ao objetivo da pesquisa.

5) Visualização e Análise dos Dados: Para a apresentação dos resultados, foi empregada a ferramenta VOSviewer Van Eck e Waltman (2010), que utilizou as matrizes geradas na etapa anterior. Nos grafos, foram analisadas as relações, a intensidade, o número de produções e a clusterização, representados por meio dos nós, vínculos e cores. Para garantir uma visualização clara, filtros foram aplicados com base no grau de cada nó. Além disso, para a criação de outros gráficos, utilizou-se a ferramenta Microsoft Excel®.

RESULTADOS E DISCUSSÃO

Os resultados e as discussões serão apresentados em duas partes: inicialmente, aborda-se a evolução temporal das publicações científicas; em seguida, discute-se a trajetória temática das publicações sobre PD. A Figura 1 apresenta a produção científica brasileira sobre PD por ano, abrangendo de 2004 a 2024, o que equivale a os 20 anos de pesquisas, bem como sua evolução temporal.

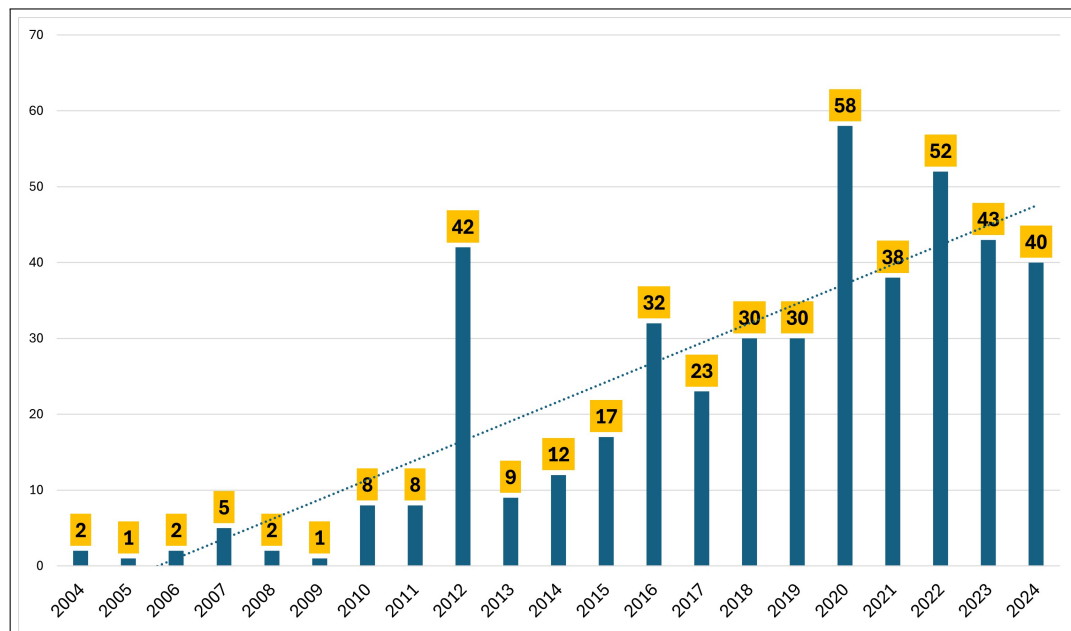


Figura 1. Produção por ano

Em 2004, ano seminal para a PD, duas publicações pioneiras lançaram as bases para essa área emergente. Ambos os artigos, representam os primeiros avanços na discussão sobre PD. O primeiro, escrito por Katia P. Thomaz em parceria com Antônio José Soares, foi publicado em fevereiro na revista *DataGramZero*. Intitulado “A preservação digital e o modelo de referência *Open Archival Information System (OAIS)*”, o artigo discute os desafios relacionados ao armazenamento e ao transporte de informações digitais, que se intensificaram com o aumento significativo de dados nas últimas décadas.

O estudo teve como objetivo fornecer uma visão geral das diversas iniciativas que surgiram para enfrentar esses desafios, além de discutir os requisitos e as principais estratégias adotadas. De acordo com Thomaz e Soares (2004), a elaboração e o desenvolvimento de metadados de preservação representam um componente essencial na maioria das estratégias de PD. Os autores destacam que os métodos atuais são instáveis e as tecnologias de acesso rapidamente se tornam obsoletas. Thomaz e Soares (2004) também observaram que, até recentemente, os profissionais da área de informação concentravam-se principalmente na durabilidade do suporte físico utilizado para armazenar os dados.

O segundo artigo, publicado na edição de maio/agosto de 2004 na revista *Ciência da Informação*, foi escrito por Miguel Ángel Márdero Arellano. Intitulado “Preservação de documentos digitais”, o estudo teve como objetivo apresentar os resultados de uma pesquisa bibliográfica sobre PD, analisando as práticas contemporâneas voltadas para a conservação de documentos digitais.

Esta pesquisa tornou-se substancial para a área, especialmente porque, no início dos anos 2000, a internet estava em plena expansão. A crescente digitalização de documentos, aliada à rápida transformação digital, tornou a questão da PD ainda mais urgente. Como destaca Arellano (2004), garantir a preservação dos documentos digitais exige a adoção de medidas específicas, pois sem essas ações não é possível manter o acesso, a qualidade e autenticidade desses registros com o passar do tempo.

A pesquisa de Arellano (2004) baseou-se em estudos de especialistas que apresentam diversos argumentos sobre a importância e a urgência das atividades de preservação de objetos digitais. Segundo o autor, inicialmente as práticas de PD focavam principalmente na garantia da longevidade dos arquivos. No entanto, como aponta Arellano (2004), a preocupação atual deslocou-se para a falta de conhecimento sobre estratégias de PD e suas implicações na manutenção da durabilidade dos arquivos digitais.

O ano de 2012 destacou-se como o primeiro grande aumento na produção e marcou o início de um ciclo de crescimento. Com um total de 42 publicações, esse ano apresentou um aumento relevante em relação aos anos anteriores, que registraram uma média de 1 a 8 publicações. Refletindo o fortalecimento da produção científica na área e o início de um período mais produtivo. Esse avanço pode ser atribuído, em parte, ao impacto da publicação de normas relacionadas à PD no Brasil. Um exemplo é a Lei nº 12.682, de 9 de julho de 2012, que estabelece que documentos digitalizados, produzidos e armazenados conforme os requisitos legais, possuem o mesmo valor probatório dos documentos originais, permitindo inclusive a eliminação dos originais físicos, exceto para documentos de valor histórico (Lei nº 12.682, 2012). Essa foi a primeira legislação a abordar diretamente o valor legal dos documentos digitalizados e seu armazenamento digital, lançando as bases para a modernização da gestão documental no país.

Outro exemplo é a Resolução nº 37, de 19 de dezembro de 2012, do Conselho Nacional de Arquivos (CONARQ), a qual aprovou as “Diretrizes para a Presunção de Autenticidade de Documentos Arquivísticos Digitais”. Essa norma complementa o arcabouço legal ao fornecer orientações técnicas e conceituais que visam assegurar a autenticidade e o valor probatório dos documentos arquivísticos digitais ao longo do tempo [Conselho Nacional de Arquivos \(2012\)](#). [Grácio \(2012\)](#) reforça a necessidade da existência de leis brasileiras que regulamentem os documentos em formato eletrônico, possibilitando sua legalização e oferecendo suporte às instituições comprometidas com a PD.

Em 2016, houve um novo pico na produção acadêmica, consolidando uma tendência de crescimento. Com 32 publicações, esse ano registrou o maior volume desde 2012, quando os números foram de 9, 12 e 17 publicações, representando um momento significativo. O aumento na produção reflete um período de maior engajamento e aprofundamento na área.

O ano de 2016 foi marcado por importantes normativas voltadas à modernização da gestão pública e ao fortalecimento da confiabilidade documental no ambiente digital. Um dos destaques foi a promulgação do Decreto nº 8.638, em 15 de janeiro, que instituiu a Política de Governança Digital no âmbito da administração pública federal direta, autárquica e fundacional. Essa política visou ampliar a eficiência e a transparência na prestação de serviços públicos, por meio da utilização estratégica da informação e das tecnologias de informação e comunicação (TICs) (Decreto nº 8.638, 2016). Complementando esse cenário, a Resolução nº 4.474, publicada pelo Banco Central do Brasil em 31 de março do mesmo ano, estabeleceu diretrizes para a digitalização e a gestão de documentos relacionados às operações financeiras realizadas por instituições autorizadas a funcionar pelo órgão regulador. A norma também regulamentou o descarte das versões físicas dos documentos digitalizados, reforçando a validade legal dos arquivos eletrônicos e contribuindo para a consolidação de práticas seguras e eficazes de PD no setor financeiro (Resolução nº 4.474, 2016). O ano de 2020 registrou o maior número de publicações, totalizando 58 trabalhos. Em comparação com o ano anterior, que contabilizou 39 publicações, esse aumento representa o pico da produção científica no período analisado. Esse resultado evidencia um crescimento expressivo, consolidando 2020 como o ano de maior destaque na série de publicações.

Esse aumento ocorreu em um contexto crucial: a vigência da LGPD, estabelecida em agosto de 2018, mas que entrou em vigor de forma escalonada. A aplicação dos dispositivos da lei teve início em setembro de 2020, momento em que passaram a valer as regras gerais de proteção de dados pessoais. Além disso, outro fator relevante foi a pandemia de covid-19, que acelerou a digitalização em diversos setores da sociedade. A rápida transição para o trabalho remoto, o ensino à distância e o aumento no uso de serviços online destacaram a importância da segurança digital para garantir o acesso à informação e a continuidade dos processos, alterando as dinâmicas de produção e gerando uma maior reflexão sobre a PD.

Observa-se que 2020 foi um ano importante para a área, pois destacou a relevância da PD e impulsionou discussões sobre a necessidade de infraestruturas adequadas para a gestão e preservação de dados em um mundo cada vez mais digital. Após o pico de 2020, a tendência de crescimento continuou, com 38 publicações em 2021 e 52 em 2022, mantendo uma trajetória positiva. No entanto, a produção registrou uma leve queda em 2023 (n=43) e 2024 (n=40). Apesar dessa redução, o número de publicações ainda se mantém significativamente superior ao observado nos anos iniciais, evidenciando uma tendência geral de crescimento ao longo do tempo.

Buscou-se também analisar a trajetória temática das publicações, permitindo uma compreensão dos conceitos adotados pelos pesquisadores ao longo do tempo. A figura 2 apresenta a trajetória temática das publicações sobre PD, relacionando as palavras-chave atribuídas pelos autores às décadas de suas publicações. Vale ressaltar que para esta análise optou-se por excluir a palavra “preservação digital”, com o intuito de evidenciar elementos de maior representatividade semântica e interpretativa, bem como explorar a profundidade e a diversidade dos temas adjacentes. Essa estratégia permite uma análise mais detalhada e favorece a identificação de temas emergentes que orbitam os conceitos centrais da área.

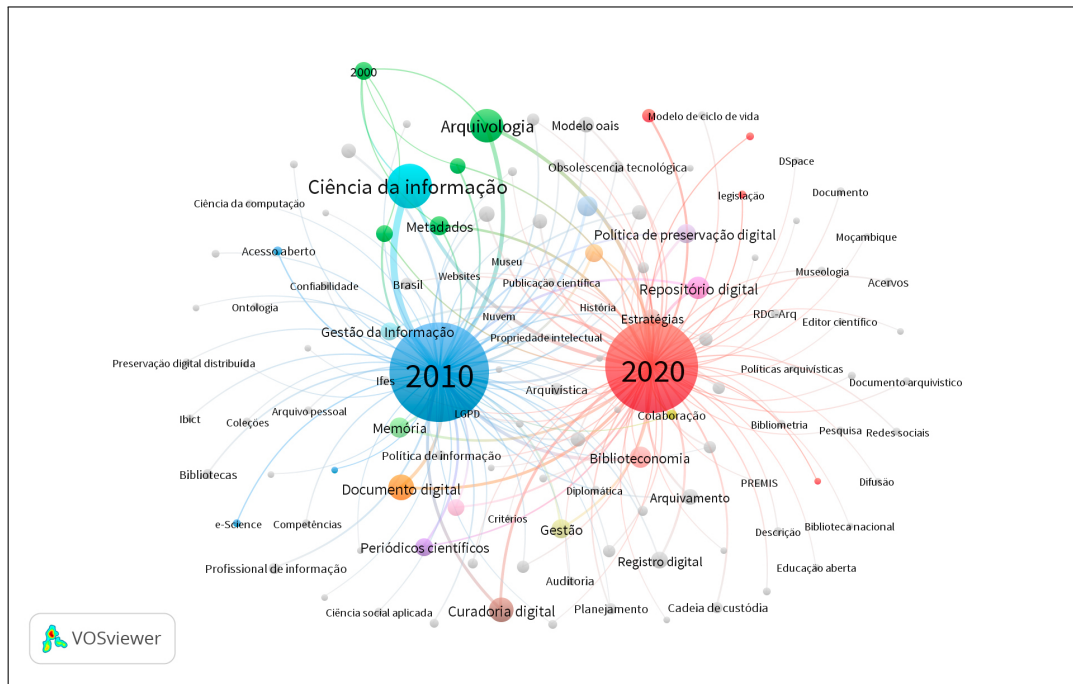


Figura 2. Trajetória temática das publicações sobre preservação digital ao longo das décadas (com mais de 3 relações)

Os temas mais abordados nos anos 2000 foram: CI ($n=5$), Arquivologia ($n=5$), Bibliotecas Digitais ($n=4$), Metadados ($n=4$) e Digitalização ($n=3$). O conceito de “CI” se destacou ao fornecer a base teórica necessária para lidar com a crescente complexidade dos dados digitais. Segundo ? , a CI trata dos desafios relacionados à comunicação eficiente do conhecimento e de seus registros, considerando as necessidades informacionais de indivíduos, organizações e da sociedade. Essa área busca responder a questões científicas e práticas que envolvem o fluxo de informação em diversos contextos, como o social e o institucional. Assim, a CI contribuiu para o desenvolvimento de diretrizes voltadas à gestão e preservação da informação digital. No entanto, sob uma perspectiva bibliométrica, a palavra-chave “CI” apresenta baixa relevância contextual, pois refere-se à própria área em que os artigos foram publicados, tornando sua presença uma obviedade.

O tema “Arquivologia” também se destacou nas discussões, visto que é um campo de estudo voltado para a organização, gestão e preservação de documentos e informações, tratando do ciclo de vida dos documentos e garantindo sua preservação, tem seus princípios alinhados à PD.

O tema “Bibliotecas Digitais” surgiu em um momento imprescindível para o armazenamento e acesso às informações digitais, refletindo a necessidade de criar ambientes organizacionais e tecnológicos que possibilitem o acesso remoto a acervos de documentos e obras digitalizadas, além de garantir a preservação desses recursos para as gerações futuras. De acordo com Marcondes, Kuramoto, Toutain, e Sayão (2005), a biblioteca digital é um sistema que organiza e disponibiliza conteúdos digitais, como livros e artigos, acessíveis via rede e armazenados em servidores próprios ou compartilhados.

Outro conceito que ganhou destaque foi o de “metadados”, emergindo da necessidade de descrever, organizar e recuperar informações digitais. Grácio (2012) define metadados como um conjunto de dados que descrevem um determinado item, facilitando seu acesso e recuperação por usuários ou sistemas de busca. O autor enfatiza a importância do desenvolvimento de modelos e padrões capazes de atender às novas demandas informacionais, mesmo diante das constantes transformações tecnológicas, assegurando o acesso e a preservação em longo prazo.

Embora o tema “Digitalização” não tenha sido o mais destacado na década de 2000, os conteúdos mais recorrentes são interligados e fundamentais no processo de digitalização, já que a transição do formato físico para o digital exige uma reconfiguração das formas de organização, gestão e preservação de objetos digitais. Para Marcondes et al. (2005), p. 16., a digitalização é o “processo de conversão de um documento analógico para um formato digital, convertendo-o em sinais binários, por meio de dispositivo apropriado, como um *scanner* ou câmera fotográfica digital”. Nesse contexto, a digitalização impulsionou o aumento da produção e da necessidade de gestão de informações em formato digital, estabelecendo as bases para os estudos e práticas de PD.

Observa-se que, em 2000, temas como “Curadoria Digital” e “Política de PD” ainda não figuravam entre os mais discutidos, o que indica uma lacuna nas publicações iniciais da área. Isso sugere que, naquele momento, a principal preocupação estava centrada na organização e no acesso à informação, com uma abordagem predominantemente técnica e estratégica, sem um aprofundamento nas políticas institucionais e normativas. Por outro lado, houve ênfase em temas relacionados à Arquivologia e Bibliotecas digitais.

Nos anos 2010, os temas mais representativos foram: CI (n=101), Arquivologia (n=39), Biblioteconomia (n=26), Documento Digital (n=25), Curadoria Digital (n=21) e Documento Arquivístico Digital (n=19). O conceito CI, voltou a ganhar destaque por se referir à área em que os artigos foram publicados. Além disso, observa-se a relevância do tema “Arquivologia”, evidenciada por sua função essencial no gerenciamento e na preservação de documentos digitais, cuja produção se intensificou com os avanços tecnológicos. Ademais, esse campo incorporou conceitos fundamentais para a organização eficiente desses materiais, contribuindo para a introdução e consolidação da temática “Curadoria Digital”, que também ganhou destaque nesta década. Segundo ? , a curadoria digital é o processo contínuo de selecionar, interpretar e divulgar informações relevantes, com o objetivo de atender às necessidades informacionais de um público específico sobre determinado tema.

O tema “Biblioteconomia” ganhou relevância em um contexto de ampla expansão da informação digital. Nesse cenário, as bibliotecas passaram a ser reconhecidas como instituições centrais na preservação do patrimônio digital, devido à sua importância na seleção, organização, recuperação e conservação de documentos. Segundo Souza (1996), a Biblioteconomia é uma ciência moldada por práticas sociais, voltada para a mediação da informação e o acesso ao conhecimento, sendo caracterizada por sua flexibilidade e capacidade de adaptação às transformações culturais e tecnológicas. No entanto, apesar dessa natureza flexível e interdisciplinar, a transição para o ambiente digital exigiu que as bibliotecas — assim como os profissionais da informação — revisassem e atualizassem seus princípios e práticas tradicionais para atender às novas demandas tecnológicas e informacionais. A discussão sobre o tema “Documento Digital” emergiu na década de 2010, refletindo a transição do ambiente físico para o digital — processo que vai além da simples digitalização. De acordo com o Conselho Nacional de Arquivos (2020a), documento digital é a informação armazenada em código binário. Esse formato surgiu para facilitar o acesso, a disseminação e a preservação de informações, tornando-se um elemento central no campo da PD.

A análise dos dados revela que, na década de 2010, tanto os “Metadados” quanto a “Política de PD” não ocuparam posição de destaque no debate acadêmico, evidenciando uma importante lacuna temática no período. Embora os metadados tenham sido um dos temas centrais nas discussões dos anos 2000, seu protagonismo reduziu-se significativamente na década seguinte — fato que sugere menor ênfase nas questões de descrição, recuperação e interoperabilidade de objetos digitais.

No mesmo intervalo, observa-se que o debate sobre políticas de PD tampouco assumiu centralidade, mesmo diante do crescente interesse por curadoria digital, repositórios institucionais e da promulgação de normativas específicas no contexto brasileiro. Essa dimensão normativa e política, embora indispensável para a consolidação institucional da PD, não alcançou o protagonismo esperado nas discussões do período.

Nos anos 2020, os temas mais significativos foram: CI (n=29), Arquivologia (n=29), Documento Digital (n=20), Repositório Digital (n=19), Curadoria Digital (n=18), Memória (n=16) e Política de PD (n=16). Após o tema “CI”, os temas “Arquivologia” e “Documento Digital”, que já haviam se destacado na década anterior, continuaram a manter sua relevância nesta nova década.

O tema “Arquivologia” consolidou-se como eixo central da PD desde suas origens. Inicialmente, as discussões sobre PD surgiram da necessidade de proteger documentos físicos e gerenciar sua migração para o formato digital. Com o amadurecimento das pesquisas, a Arquivologia tornou-se ainda mais relevante. Conforme Heredia Herrera (1987), a arquivologia é a área do conhecimento voltada ao estudo da natureza dos arquivos, abrangendo os princípios de sua conservação, organização e os métodos que garantem seu acesso e uso adequados.

A crescente complexidade dos dados digitais e a necessidade de políticas de preservação mantiveram a Arquivologia em destaque, sobretudo ao incorporar o conceito de “Política de PD”. De acordo com Holanda (2019), trata-se de uma ferramenta institucional que permite aos órgãos e entidades definirem como preservar seus documentos. Essa política compreende princípios, diretrizes e responsabilidades que orientam a criação de programas, projetos, planos e procedimentos com vistas a garantir a conservação e o acesso a documentos arquivísticos digitais autênticos.

O tema “Políticas de PD” emergiu de forma tardia no debate sobre PD, apesar da existência, desde o início dos anos 2010, de leis e atos normativos brasileiros relacionados à digitalização e ao armazenamento em mídias eletrônicas. No entanto, somente a partir dos anos 2020 esse tema passou a ser amplamente discutido.

A presença recorrente da temática “Arquivologia” ao longo das décadas reforça sua centralidade na evolução da PD, demonstrando sua capacidade de adaptar princípios e práticas à preservação de documentos não apenas físicos, mas também digitais, à medida que surgem novas tecnologias.

O tema “Documento Digital” manteve-se em evidência nas duas últimas décadas, impulsionado pela expansão da internet e pelo avanço tecnológico. Tornou-se fundamental compreender as especificidades, os desafios e as necessidades associadas aos diferentes tipos de documentos digitais (como texto, imagem, áudio, vídeo, entre outros). Cordeiro, Parga, Barbosa, e Menezes (2016) observam que, diante da constante evolução tecnológica, a PD torna-se essencial, sendo os documentos digitais um dos maiores desafios, já que sua conservação ao longo do tempo ainda é uma questão em aberto para a sociedade atual.

Por fim, nota-se que, em 2020, os temas “Digitalização” e “Biblioteconomia” não figuraram entre os mais abordados. A digitalização, que teve proeminência nos anos 2000, deixou de ser tratada como questão técnica e passou a ser considerada uma prática consolidada. Por sua vez, o tema Biblioteconomia, que ganhou relevância nos anos 2010, perdeu força nas discussões seguintes, indicando uma mudança de foco para temáticas relacionadas à memória e às políticas de preservação.

CONSIDERAÇÕES FINAIS

Este estudo analisou a historiografia bibliométrica da produção científica sobre PD no Brasil, oferecendo uma visão atualizada de sua evolução temporal e temática. Os dados evidenciam crescimento contínuo da produção desde 2004, com marcos significativos em 2012, 2016 e 2020, associados a eventos como a promulgação de atos normativos, a LGPD e a pandemia de Covid-19. Observou-se também uma mudança no foco temático: de abordagens técnicas, como metadados e digitalização, para questões mais amplas e estratégicas, como curadoria digital, políticas de preservação e memória institucional. Essa transição confirma tendências apontadas por autores como Thomaz e Soares (2004), Grácio (2012) e Arellano (2004). A pesquisa evidencia a presença consolidada da PD na CI, que, ao apoiar-se em sua natureza interdisciplinar, tem desempenhado um papel central na discussão de diretrizes técnicas, políticas e metodológicas. Essa contribuição reforça a relevância da CI na consolidação de práticas voltadas à gestão e à preservação da memória informacional em ambientes digitais. Como limitação, destaca-se o recorte exclusivo em artigos científicos, que excluiu livros, anais, dissertações, teses e literatura cinzenta, possivelmente deixando de captar parte relevante da produção. Recomenda-se que estudos futuros ampliem as bases de dados, incluindo fontes como o Oasisbr, e considerem diferentes tipos de documentos, além de aperfeiçoar as estratégias de busca. Espera-se que esta investigação contribua para o avanço do conhecimento científico ao sistematizar informações dispersas, identificar lacunas, como a baixa discussão sobre políticas de PD na década de 2010, e apontar caminhos para novos estudos. Ao evidenciar o papel institucional, normativo e estratégico da PD, o estudo oferece subsídios valiosos para pesquisadores, profissionais da informação e formuladores de políticas públicas.

REFERÊNCIAS

- Arellano, M. A. (2004). Preservação de documentos digitais. *Ciência da Informação*, 33(2), 15–27. Recuperado de <https://www.scielo.br/j/ci/a/FLfgJvpH3PZKf3HbpKYchZr/?format=pdf&lang=pt>
- Boeres, S. A. d. A. (2004). *Política de preservação da informação digital em bibliotecas universitárias brasileiras* (Dissertação de mestrado, Universidade de Brasília). Recuperado de https://repositorio.umb.br/bitstream/10482/1693/1/Dissertacao_Sonia.pdf (Dissertação de Mestrado)
- Borbinha, J., Henriques, C., Lopes, B., & Sequeira, J. D. F. (2002). Manifesto para a preservação digital. *Cadernos BAD*(2). Recuperado de <https://doi.org/10.48798/cadernosbad.867> doi: 10.48798/cadernosbad.867
- Cassares, N. C. (2000). *Como fazer conservação preventiva em arquivos e bibliotecas*. Arquivo Público do Estado e Imprensa Oficial. Recuperado de https://www.arqsp.org.br/arquivos/oficinas_colecao_como_fazer/cf5.pdf
- Claro, M., & Castro-Grau, C. (2023). *O papel das tecnologias digitais na aprendizagem do século xxi* (Relatório Técnico). UNESCO. Recuperado de https://unesdoc.unesco.org/ark:/48223/pf0000386981_por (Fórum Regional de Política Educacional, Vol. 7)
- Conselho Nacional de Arquivos. (2012). *Resolução nº 37, de 19 de dezembro de 2012: Aprova as diretrizes para a presunção de autenticidade de documentos arquivísticos digitais*. Recuperado de <https://www.gov.br/conarq/pt-br/legislacao-arquivistica/resolucoes-do-conarq/resolucao-no-37-de-19-de-dezembro-de-2012> (Diário Oficial da União, seção 1, n. 249, p. 152–153)
- Conselho Nacional de Arquivos. (2020a). *Identidade. in glossário: documentos arquivísticos digitais (8. ed.)*. Recuperado de https://www.gov.br/conarq/pt-br/assuntos/camaras-tecnicas-setoriais-inativas/camara-tecnica-de-documentos-eletronicos-ctde/glosctde_2020_08_07.pdf (Câmara Técnica de Documentos Eletrônicos (CTDE))
- Conselho Nacional de Arquivos. (2020b). *Integridade. in glossário: documentos arquivísticos digitais (8. ed.)*. Recuperado de https://www.gov.br/conarq/pt-br/assuntos/camaras-tecnicas-setoriais-inativas/camara-tecnica-de-documentos-eletronicos-ctde/glosctde_2020_08_07.pdf (Câmara Técnica de Documentos Eletrônicos (CTDE))
- Cordeiro, L. S., Parga, M. d. F. A. S., Barbosa, N. d. S., & Menezes, S. C. F. (2016). Preservação digital e a biblioteconomia. *Revista Bibliomar*, 15(esp), 36–49. Recuperado de <https://periodicoseletronicos.ufma.br/index.php/bibliomar/article/view/9823/5776>
- Ferreira, M. (2006). *Introdução à preservação digital: conceitos, estratégias e atuais consensos*. Escola de Engenharia da Universidade do Minho. Recuperado de <https://repositorium.uminho.pt/bitstream/1822/5820/1/livro.pdf>
- Formenton, D., Castro, F. F. d., Gracioso, L. d. S., Furnival, A. C. M., & Simões, M. d. G. d. M. (2017). Los estándares de metadatos como recursos tecnológicos para la garantía de la preservación digital. *Biblios*(68), 82–95. Recuperado de <https://doi.org/10.5195/biblios.2017.414> doi: 10.5195/biblios.2017.414
- Formenton, D., & Gracioso, L. d. S. (2020). Preservação digital: desafios, requisitos, estratégias e produção científica. *RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação*, 18, e020012. Recuperado de <https://doi.org/10.20396/rdbci.v018i0.8658868> doi: 10.20396/rdbci.v018i0.8658868
- Gladney, H. M. (2009). Long-term preservation of digital records: trustworthy digital objects. *The American Archivist*, 72(2), 401–435. Recuperado de <https://doi.org/10.17723/aarc.72.2.g513766100731832> doi: 10.17723/aarc.72.2.g513766100731832
- Grácio, J. C. A. (2012). *Preservação digital na gestão da informação: um modelo processual para as instituições de ensino superior*. Cultura Acadêmica. Recuperado de <https://repositorio.unesp.br/server/api/core/bitstreams/c38a6f96-1eab-4731-8e88-9c0f04a51094/content>
- Heredia Herrera, A. (1987). *Archivística general: teoría y práctica* (2a. ed.). Diputación Provincial. Recuperado de https://www.anabad.org/wp-content/uploads/2024/04/1987_1-Archivistica-general-teoria-y-practica-2a-edicion.pdf
- Holanda, A. P. d. (2019). *Recomendações para elaboração de política de preservação digital*. Recuperado de https://www.gov.br/arquivonacional/pt-br/servicos/gestao-de-documentos/orientacao-tecnica-1/recomendacoes-tecnicas-1/politica_presercacao_digital.pdf
- Innarelli, H. C. (2011). Preservação digital: a influência da gestão dos documentos digitais na preservação da informação e da cultura. *Revista Digital de Biblioteconomia e Ciência da Informação*, 9(1), 72–87. Recuperado de <https://doi.org/10.20396/rdbci.v8i2.1934> doi: 10.20396/rdbci.v8i2.1934
- Lee, K.-H., Slattey, O. T., Lu, R., Tang, X., & McCrary, V. (2002). The state of the art and practice in digital preservation. *Journal of Research of the National Institute of Standards and Technology*, 107(1), 93–106. Recuperado de <https://doi.org/10.6028/jres.107.010> doi: 10.6028/jres.107.010
- Lira, J., & Siebra, A. S. d. (2021). Preservação digital: revisitando o essencial. In S. d. A. Siebra & V. R. Borba (Eds.), *Preservação digital e suas facetas* (p. 31–83). Pedro & João. Recuperado de https://www.ufpe.br/documents/39626/3547884/Ebook_Preservac%C2%B8a_o_Digital_e_suas_Facetas_2021.pdf/128192be-71fe-4b9b-b685-28fbee9d928d
- Machado, B. P. (2015). *A importância do diagnóstico de conservação para nortear as ações de preservação em arquivos, bibliotecas e museus*. Recuperado de https://bdm.umb.br/bitstream/10483/11204/1/2015_BrunaPereiraMachado.pdf (Monografia de Graduação, Universidade de Brasília)
- Marcondes, C. H., Kuramoto, H., Toutain, L. B., & Sayão, L. F. (Eds.). (2005). *Bibliotecas digitais: saberes e práticas*. EDUFBA and IBICT. Recuperado de <https://livroaberto.ibict.br/bitstream/123456789/1013/1/Bibliotecas%20Digitais.pdf>
- Porter, A. L., & Cunningham, S. W. (2004). *Tech mining: exploiting new technologies for competitive advantage*. John Wiley & Sons. Recuperado de http://repo.darmajaya.ac.id/4054/1/Alan%20L.%20Porter%20C%20Scott%20W.%20Cunningham%20-%20Tech%20Mining_%20Technology%20Management%20through%20Information%20Mining-Wiley-Interscience%20%282004%29.pdf
- Santos, H. M., & Flores, D. (2018). A obsolescência do conhecimento em preservação digital. *Ciência da Informação em Revista*, 5(1), 41–58. Recuperado de <https://doi.org/10.28998/cirev.2018v5n1d> doi: 10.28998/cirev.2018v5n1d
- Silveira, L., Méndez-Solano, A., & Mora-Campos, A. (2022). Ciência aberta na perspectiva de especialistas brasileiros:

proposta de taxonomia. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, 27(1), 1–31. Recuperado de <https://brapci.inf.br/v/194199> (Tradução)

Souza, F. C. d. (1996). Os paradigmas da biblioteconomia e suas implicações no ensino desta ciência. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*(2). Recuperado de <https://periodicos.ufsc.br/index.php/eb/article/view/5/8>

Thomaz, K. P., & Soares, A. J. (2004). A preservação digital e o modelo de referência Open Archival Information System (OAIS). *DataGramaZero*, 5(1). Recuperado de <https://pt.scribd.com/document/96453309/A-preservacao-digital-e-o-modelo-de-referencia-Open-Archival-Information-System>

Van Eck, N., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. Recuperado de <https://doi.org/10.1007/s11192-009-0146-3> doi: 10.1007/s11192-009-0146-3

Como citar este artigo (APA):

Machado, E. M., Sobral, N. V., Santos, T. H. N., & Albuquerque, D. E. M. (2025). Preservação Digital no Brasil: construção de uma historiografia bibliométrica a partir da produção científica. *AtoZ: novas práticas em informação e conhecimento*, 14, 1 – 12. Recuperado de: <http://dx.doi.org/10.5380/atoz.v14.99934>

NOTAS DA OBRA E CONFORMIDADE COM A CIÊNCIA ABERTA

CONTRIBUIÇÃO DE AUTORIA

Papéis e contribuições	Edilma Maria Machado	Natanael Vitor Sobral	Thais Helen do Nascimento Santos	Daniela Eugênia Moura de Albuquerque
Concepção do manuscrito	X	X		
Escrita do manuscrito	X	X		
Metodologia	X	X	X	X
Curadoria dos dados	X	X		
Discussão dos resultados	X	X	X	X
Análise dos dados	X	X	X	X

EQUIPE EDITORIAL

Editora/Editor Chefe

Paula Carina de Araújo (<https://orcid.org/0000-0003-4608-752X>)

Editora/Editor Associada/Associado Júnior

Karolayne Costa Rodrigues de Lima (<https://orcid.org/0000-0002-6311-8482>)

Editora/Editor de Texto Responsável

Suzana Zulpo (<https://orcid.org/0000-0002-6311-8482>)

Seção de Apoio às Publicações Científicas Periódicas - Sistema de Bibliotecas (SiBi) da Universidade Federal do Paraná - UFPR

Editora/Editor de Layout

Tiago Batista Pedra (<https://orcid.org/0009-0000-7385-7273>)