

Explorando intersecções entre Arquivamento da Web e Organização do Conhecimento

Jonas Ferrigolo Melo ¹; Moisés Rockembach ²

Resumo: Este artigo tem como objetivo explorar as relações entre a Organização do Conhecimento e o Arquivamento da Web, além de identificar critérios na preservação da Web. A metodologia adotada foi revisão de literatura, em conjunto com a análise de modelos para o Arquivamento da Web. Foram examinadas as pesquisas que abordam os métodos, técnicas e estratégias empregados na seleção, aquisição, organização, armazenamento, descrição e disponibilização de acesso ao conteúdo web arquivado. Conclui-se que os modelos selecionados podem ser aplicados por distintas iniciativas de Arquivamento da Web e apresentam relação estreita com os estudos de Organização do Conhecimento. No entanto, é necessário maiores estudos sobre a estrutura e as ferramentas utilizadas no campo da Organização do Conhecimento a fim de assegurar a preservação de longo prazo e o acesso eficaz as informações disponíveis na web.

Palavras chave: Arquivamento da web; Organização do Conhecimento; Websites; Preservação Digital;

Sumario:

1. Introdução
2. Intersecções do Arquivamento da Web e Organização do Conhecimento
3. Metodologias, técnicas e abordagens para o Arquivamento da Web
4. Conclusão
5. Referências

1. Introdução

A web, em razão da sua dinamicidade, exige que o processo de preservação seja pensado de forma sistêmica desde o princípio, incluindo metodologia de coleta dos dados, estabelecimento de políticas para seleção do conteúdo, técnicas e métodos de armazenamento, preservação digital, acesso e demais procedimentos necessários quando se trabalha com informação associada à tecnologia. A esse conjunto de atividades relativas à preservação do ambiente web é dado o nome de Arquivamento da Web que, de forma objetiva, é o processo que compreende coletar, armazenar e disponibilizar a informação retrospectiva da World Wide Web para futuros pesquisadores (Rockembach, 2018; Rockembach, Pavão, 2018).

Para a promoção do Arquivamento da Web são necessários outros recursos teóricos e práticos que possibilitam a efetiva concretização destes procedimentos. Ainda que tenha um relação estreita com a Ciência na Informação, é na área de

¹ Filiación Universidade do Porto
E-mail: jonasferrigolo@gmail.com
ORCID: - 0000-0002-7312-3509

² Filiación: Universidade de Coimbra
E-mail: moises.rockembach@gmail.com
ORCID: 0000-0001-9057-0602

estudo da Organização do Conhecimento em que muitas de suas técnicas, rotinas e procedimentos são apoiados, considerando a longa tradição de pesquisas desta área.

A Organização do Conhecimento é o campo que se dedica a indexação, classificação e representação de documentos para recuperação da informação, navegação e processos relacionados, sejam realizados por humanos ou computadores (Hjørland, 2021).

A Organização do Conhecimento, incluindo seus procedimentos (como catalogação, análise de assunto, indexação e classificação) e recursos (como sistemas de classificação, listas de cabeçalhos de assunto, tesouros, ontologias e sistemas de metadados), desempenha um papel essencial na Ciência da Informação. Isso ocorre porque a organização é um pré-requisito para qualquer forma de acesso e utilização eficaz do conhecimento no contexto dos serviços de informação (Mogollón et al., 2018). E, portanto, as linhas de estudo da Organização do Conhecimento são o cerne que compõe o fazer do Arquivamento da Web.

Na camada mais profunda, as relações entre Organização do Conhecimento e Arquivamento da Web baseiam-se em questões filosóficas melhor caracterizadas como “paradigmas”. No âmbito da migração massiva da produção da informação a partir da web, seja por meio de websites ou mídias sociais, emerge a necessidade de avaliar, classificar, preservar, recuperar e acessar esse conteúdo pelo tempo que for necessário. Para isso, diversas técnicas, procedimentos e metodologias podem ser utilizadas, sejam pautadas no paradigma custodial, patrimonialista, historicista e tecnicista; ou no pós-custodial, informacional e científico (Silva e Ribeiro, 2011, 2012). Independente do debate conceitual paradigmático, nenhum Arquivo da Web e nenhuma sistemática de Organização do Conhecimento podem ser neutros em relação a conflitos paradigmáticos e, portanto, tais problemas filosóficos representam a base para o estudo do Arquivamento da Web e Organização do Conhecimento.

Conscientes destes debates paradigmáticos pautados na área da Ciência da Informação, mais especificamente na Arquivologia, este artigo tem como objetivo apresentar as intersecções entre o Arquivamento da Web e a Organização do Conhecimento e identificar modelos para o Arquivamento da Web. A metodologia adotada neste estudo consistiu na revisão de literatura sobre Arquivamento da Web e Organização do Conhecimento. Foram analisados os trabalhos que discutem as metodologias, técnicas e abordagens utilizadas para selecionar, adquirir, organizar, armazenar, descrever e fornecer acesso ao conteúdo web arquivado. Destacam-se alguns modelos relevantes, como os propostos por Bragg e Hanna (2013), Maemura et al. (2018), Németh e Drótos (2019), Khan e Rahman (2019) e Rockembach (2021). Serão explorados os estágios desses modelos, demonstrando como podem ser aplicados por qualquer iniciativa de Arquivamento da Web.

A variedade de métodos utilizados no Arquivamento da Web é influenciada por diversos fatores, como o ambiente jurídico, o relacionamento entre produtores de conteúdo e os arquivos, a natureza do conteúdo arquivado, a escala do arquivo e a capacidade técnica e financeira da organização responsável pelo arquivamento (Masanès, 2006). Diferentes pesquisadores propuseram modelos e

fluxos de trabalho para o Arquivamento da Web, abrangendo desde a recolha de informação até a disponibilização para pesquisa.

2. Intersecções do Arquivamento da Web e Organização do Conhecimento

Com a rápida disseminação dos protocolos da web e a legalização das atividades comerciais pela Internet nos EUA no início da década de 1990 (Abbate, 1999; Guizzo, 1999), a Internet foi transformada, deixando de ser uma ferramenta especializada para cientistas e estudantes para se transformar em uma infraestrutura social globalmente acessível, com a possibilidade de uma produção massiva de informações. Junto disso, emerge uma onda caracterizada pela produção exponencial de dados, de novas formas de comunicação, disseminação e organização de conhecimento (Jenkins; Deuze, 2008; Finnemann, 2014).

Com a possibilidade de adicionar, modificar ou excluir qualquer fonte relevante, e de selecionar recursos a qualquer momento, a web como um todo parece qualificar-se como um sistema quase perfeito para a organização do conhecimento (Finnemann, 2019). No entanto, Finnemann (2019) três obstáculos que impedem que a web seja uma solução suficiente para a Organização do Conhecimento no século XXI: ela é muito grande para ser vista por qualquer observador, incluindo rastreadores automatizados; seu carácter efêmero, em que todos os dias novos materiais são inseridos ao passo que outros são apagados ou perdidos; e a grande vantagem da Internet, que permite a todos serem produtores de conteúdo. Essas razões fazem com que a web não seja, por si só, suficiente como um sistema de organização do conhecimento, ao passo que o conhecimento advindo desta área precisa estar inserido nas práticas e políticas de arquivamento da web, na medida em que os materiais sejam considerados dignos de preservação para acesso futuro.

Nesse sentido, as técnicas e práticas da Organização do Conhecimento começam a ingressar nos métodos e rotinas para Arquivamento da Web. A relação entre estes dois campos, ainda que pouco estudada, já foi explicitamente mencionada por Finnemann (2019), que aponta que as distinções entre coleções especiais, infraestruturas de pesquisa e arquivos gerais não são claras, porém ainda possuem significado, uma vez que cada uma dessas finalidades possui implicações no conjunto de métodos utilizados para seleção. Dessa forma, o sistema "ideal" para a organização do conhecimento acaba se tornando uma combinação cada vez mais complexa de materiais da web coletados e arquivados, devido a uma variedade de critérios.

Para além da assertiva de Finnemann diretamente relacionada ao Arquivamento da Web, cabe destacar que a Organização do Conhecimento, de modo geral, colabora com o processo de formação de um Arquivo Web, especialmente quando apoiado por sua “vertente estreita” (Hjørland, 2008), que se refere às atividades como descrição, indexação e classificação de documentos e a preocupação com a

natureza e qualidade desses processos. Essas técnicas advindas da Organização do Conhecimento constitui um locus interdisciplinar eficaz (García Marco 1995; Hjørland 2008) no qual campos como ciência da informação, lógica, linguística, comunicação, tecnologia, entre outros, têm seus quadros teóricos e metodológicos integrados. Essa interdisciplinaridade tem o potencial de promover o desenvolvimento de processos, instrumentos, métodos e rotinas relativos à representação e organização de documentos e de seus conteúdos informativos, tanto por pessoas como por programas.

Após sua criação, um arquivo da web retém vestígios das conexões de links, funcionando como uma espécie de índice para entidades sociais, culturais e políticas, tanto de natureza cívica quanto profissional. Esses arquivos podem também ser concebidos para atuar como índices para categorias específicas de organização do conhecimento, por meio de links para coleções especiais, infraestruturas de pesquisa e sistemas de conhecimento oriundos de diversas fontes. Outro campo da Organização de Conhecimento que tem relevância basilar na composição de um Arquivo da Web é a indexação de um determinado documento. Essa etapa da composição de um arquivo deve refletir as necessidades de um determinado grupo de usuários ou de um propósito. Ou seja, qualquer descrição ou representação de um determinado documento precisa estar adequada ao seu contexto. Cabe destacar que o processo de descrição não será neutro, portanto, há de se estabelecer relações com o contexto de produção e o público-alvo sem, necessariamente, padronizar as descrições ou fazer uma descrição única para diferentes conjuntos documentais. A representação de um documento é feita para permitir aos usuários fazer discriminações relevantes. A recuperação da informação também entra neste rol de atividades advindas da Organização do Conhecimento que se tornam fundamentais para o Arquivamento da Web, ao passo que permite aos usuários encontrar as informações que buscam.

Essas atividades sempre partirão de um “lugar teórico” no âmbito da Ciência da Informação e das técnicas da Organização do Conhecimento e, portanto, não existe a maneira “ideal” para sua execução. Isso vai depender do alicerce teórico que está sendo utilizado e, também, dos objetivos de composição de um arquivo da web. Essa afirmativa que defendemos encontra críticas por parte de outros pesquisadores, que defendem que uma perspectiva única deve ser estabelecida sob o risco de se criar um campo difuso. Entendemos, por um lado, que um campo científico pode encontrar certa dificuldade em se desenvolver sem uma clareza teórica, entretanto, por outro lado, quando se insere uma nova perspectiva de objeto, é importante que diferentes abordagens sejam debatidas para serem distinguidas umas das outras e comparadas entre si. Em outras palavras: temos que examinar e interpretar as diferentes perspectivas de abordagens com muita honestidade e cuidado para que a composição de um Arquivo da Web não se torne algo automatizado e sem que seja pensado na perspectiva do usuário.

Uma maneira de fazer isso tem sido estabelecer uma distinção entre o “paradigma custodial” (ou “paradigma orientado ao sistema”), de um lado, e o “paradigma informacional e científico” (ou “paradigma orientado ao usuário”), do outro. A Organização do Conhecimento, especialmente quando articulada com as técnicas de Arquivamento da Web, pode encontrar um campo mais vantajoso quando apoiada em uma perspectiva orientada ao usuário. Em certo sentido, todas as abordagens da Organização do Conhecimento podem concordar sobre o objetivo de que os sistemas e processos visam satisfazer as necessidades de informação dos usuários. Portanto, pode-se afirmar que os usuários se beneficiam de sistemas bem estruturados, razão pela qual a abordagem vai ser orientada ou “amigável” para o usuário. A melhor forma de definir esta abordagem é provavelmente através do método: os sistemas baseados em abordagens orientadas para o usuário devem especificar como é feito o design de um sistema com base em estudos empíricos dos utilizadores.

O problema não está em apenas formular e utilizar uma teoria, mas descobrir pressupostos teóricos em diferentes práticas, formular esses pressupostos claramente a fim de tornar possível a comparação de abordagens. Nesta perspectiva, as lições advindas da Organização do Conhecimento associadas aquelas advindas das metodologias, técnicas e abordagens para o Arquivamento da Web podem trazer caminhos positivos para a preservação destes conteúdos informacionais.

3. Metodologias, técnicas e abordagens para o Arquivamento da Web

O processo de Arquivamento da Web envolve uma série de etapas e modelos que visam garantir a preservação a longo prazo dos recursos online. Diferentes pesquisadores têm proposto abordagens sistemáticas e modelos para orientar as organizações nesse processo complexo e algumas instituições já desenvolvem programas para preservar websites em seus países, encarando de forma efetiva o desafio de preservar o conteúdo produzido na Internet. No caso do Brasil, o Conselho Nacional de Arquivos aprovou em dezembro de 2021 a Política de Preservação de Websites e Mídias Sociais para o país (Melo et al., 2023; Melo; Rockembach, 2023). O desafio em estabelecer uma política com os procedimentos para a formação de um Arquivo Web perpassa tanto as questões técnicas e tecnológicas, quanto a coleta destes dados que mudam, somem e se alteram rapidamente, considerando a natureza dinâmica da Internet.

Ao longo de mais de uma década de estudos em relação à preservação da Internet foram criadas algumas iniciativas de Arquivamento da Web que definem os padrões, políticas e tecnologias para efetivação deste arquivamento. Um dos modelos utilizados é o proposto por Brown (2006), que apresenta um fluxo que começa na política de seleção, passa pela seleção, lista de coleta, coleta, garantia de qualidade, armazenamento, preservação e entrega (acesso). Esse modelo fornece uma estrutura abrangente que aborda desde a definição dos critérios de seleção até a disponibilização dos recursos preservados para acesso.

A rotina de preservação de *websites*, segundo Gomes (2010), é dividida em três etapas que envolve a recolha de informação proveniente da *web*, indexação e disponibilização de serviços de pesquisa e acesso, sendo que a primeira etapa se subdivide em coletar o arquivo, armazená-lo, extrair os endereços a partir dos *hiperlinks*, e inserir os endereços para a recolha. Com a intenção de especificar um pouco mais as fases, Niu (2012) diz que com a diversidade de recursos informacionais, o gerenciamento e o fluxo de trabalho do Arquivamento da Web passa a exigir o estabelecimento de passos, rotinas e técnicas que incluem procedimento referentes à avaliação e seleção, aquisição, organização e armazenamento, descrição e acesso, sendo este, basicamente, o fluxo de trabalho do núcleo do Arquivamento da Web.

Da mesma forma, em 2013, um grupo de trabalho do *Archive-It*, serviço de arquivamento ligado ao *Internet Archive*, publicou um *White Paper* intitulado “*The web archiving life cycle model*”, em que fala sobre o ciclo de vida do Arquivamento da Web. O modelo baseia-se nas experiências da equipe, bem como nas lições aprendidas pelas instituições parceiras, incluindo estudos de caso de seis instituições. Trata-se de um esforço voltado com a intenção de representar fluxos de trabalho comuns e criar um modelo mensurável que possa ser utilizado como referência para organizações que queiram criar ou melhorar seus programas de arquivamento da *web* (Bragg; Hanna, 2013).

O modelo é uma tentativa de incorporar os braços tecnológicos e programáticos do arquivamento da *web* em uma estrutura que será relevante para qualquer organização que deseje arquivar a *web*, independentemente do tamanho da organização, orçamento ou métodos técnicos de arquivamento da *web*. (Bragg; Hanna, 2013, p. 28, *tradução nossa*).

O modelo apresenta uma forma de visualizar as diferentes etapas e fases que instituições enfrentam ao desenvolver e gerenciar um programa de arquivamento da *web*, tendo sua forma circular para sugerir a natureza repetitiva dos passos no ciclo de vida (Bragg; Hanna, 2013). O *Modelo de Ciclo de Vida do Arquivamento da Web* pode ser visto na Figura 1.

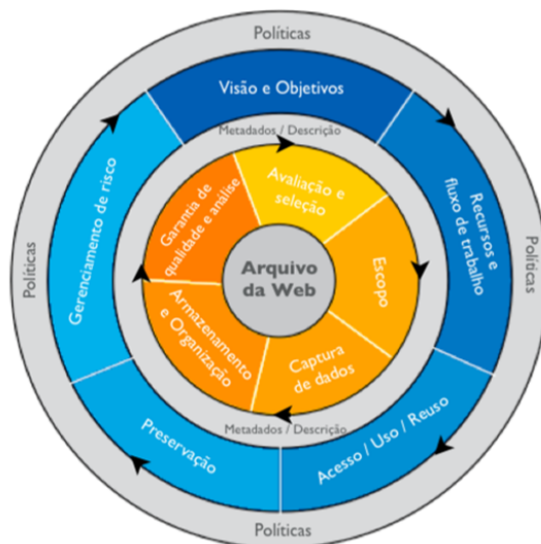


Figura 1 - Modelo de Ciclo de Vida do Arquivamento da *web*
 Fonte: Bragg; Hanna (2013, p. 3 apud ROCKEMBACH, 2018, p. 26).

Considerando que quase todas as rotinas do arquivamento da *web* envolvem algum tipo de decisão política, estes aspectos são representados pelo nível mais externo do ciclo de vida, a esfera política. Ao envolver todos os passos do ciclo de vida nesta esfera, o modelo representa a necessidade da constante elaboração de políticas nas rotinas de preservação destes conteúdos. Em uma segunda esfera, estão representados os metadados e a descrição: o serviço *Arquivo-It* escolheu incorporá-los como uma esfera completa, em vez de um segmento da esfera, para enfatizar que criar, importar e exportar metadados é um processo contínuo que ocorre em conjunto com outras atividades que compõem o ciclo de vida (Bragg; Hanna, 2013).

O círculo azul dentro da banda de políticas representa as decisões de alto nível que uma instituição enfrenta ao configurar e gerenciar seu programa de arquivamento da *web*; por sua vez, o círculo laranja descreve as tarefas do dia-a-dia envolvidas no negócio de Arquivamento da Web (Bragg; Hanna, 2013). O centro do modelo representa a própria coleção, o conteúdo da web arquivado, considerando que esses dados são o resultado final de todas as etapas anteriores.

Outra abordagem que se destaca pela sua acessibilidade e clareza é a Abordagem Sistemática para a Preservação da Web, proposta por Khan e Rahman (2019). Essa abordagem divide o processo de preservação em estágios autoexplicativos e projeta um modelo que pode ser utilizado por qualquer iniciativa de arquivamento da web. Khan e Rahman (2019) estabeleceram o modelo começando pela definição do escopo do arquivo, ou seja, determinando o tipo de site, tópico ou domínio que será preservado.

Nesta abordagem é importante reconhecer os diferentes aspectos que precisam ser observados durante o processo de preservação e arquivamento da web: gestão dos objetos digitais, formato e armazenamento do objeto digital,

gestão do arquivamento, questões administrativas, acesso e segurança ao arquivo, planejamento de preservação (Khan e Rahman, 2019), e demais questões que precisam ser entendidas para uma eficaz gestão do arquivo *web* e que ajudarão a tratar os desafios que ocorrem durante o processo de preservação.

Importante destacar que a vida útil da mídia de armazenamento digital pode ser surpreendentemente curta, e a rápida evolução das tecnologias de renderização pode impedir o acesso futuro (Lavoie, 2000). Neste sentido, o Modelo *Open Archival Information System* (OAIS) se propõe a responder o que é necessário para preservar e manter o acesso às informações digitais a longo prazo, ao passo que o modelo de referência OAIS é uma tentativa de fornecer uma estrutura de alto nível para o desenvolvimento de arquivos digitais (Khan e Rahman, 2019). Ainda que seja uma questão de difícil solução, o modelo de referência OAIS apresenta uma estrutura conceitual para um sistema de arquivo dedicado a preservar e manter o acesso às informações digitais a longo prazo (Lavoie, 2000). As dez etapas que compreendem a abordagem sistemática para preservação da web estão ilustradas na Figura 2, a seguir:

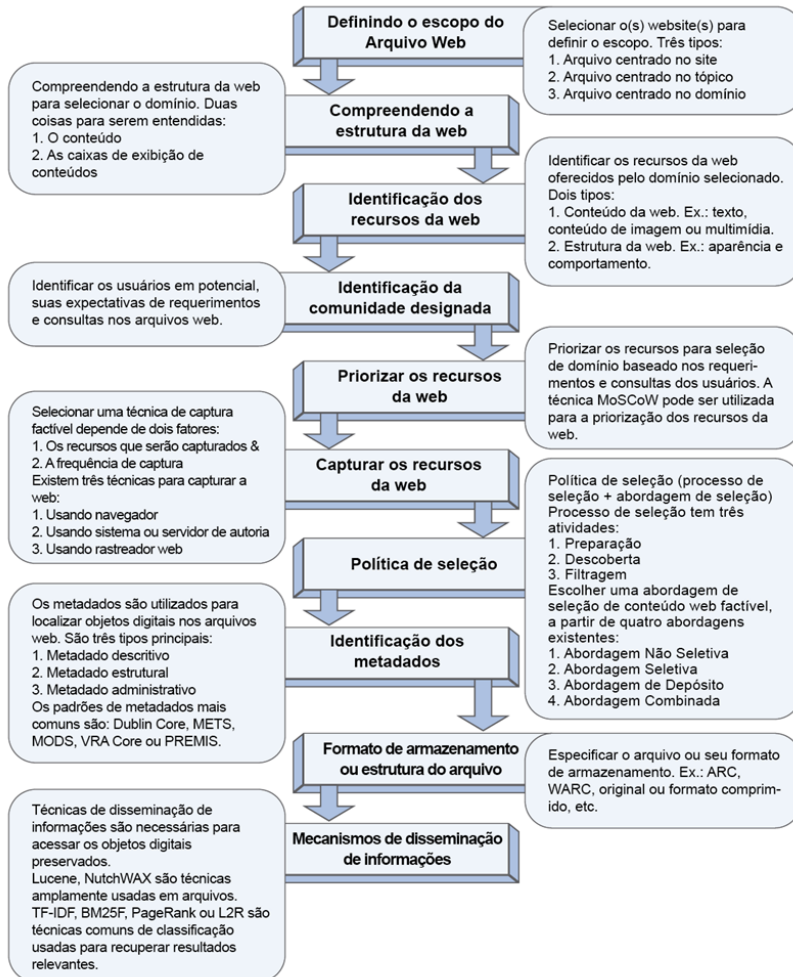


Figura 2 – Abordagem sistemática para o processo de preservação da web
 Fonte: Khan; Rahmam (2019, p. 74 apud Melo, 2020, p. 72).

A abordagem para arquivamento da web de Maemura et al. (2018) propõem um processo composto por três etapas de seleção e captura: pré-produção, rastreamentos de produção e garantia de qualidade de pós-produção. Na fase de pré-produção, são realizados rastreamentos de teste para definir o tamanho e a abrangência da coleção. Os rastreamentos de produção são responsáveis por capturar todos os recursos da web da coleção, enquanto a garantia de qualidade na pós-produção pode exigir rastreamentos adicionais. Geralmente, são necessárias várias iterações para cada uma dessas etapas, garantindo a qualidade e a integridade dos recursos preservados.

Como se pode perceber, cada fluxo de trabalho possui suas próprias rotinas, critérios, processos seletivos e softwares estabelecidos conforme sua necessidade de composição do arquivo web. A responsabilidade pela coleta dos

conteúdos web têm se estabelecido como uma importante ferramenta para preservar a memória oficial destes países (Bingham; Byrne, 2021). Ainda que as iniciativas tenham escopos distintos, no geral elas apresentam as etapas levantadas por Rockembach (2021) e ilustradas na Figura 3:

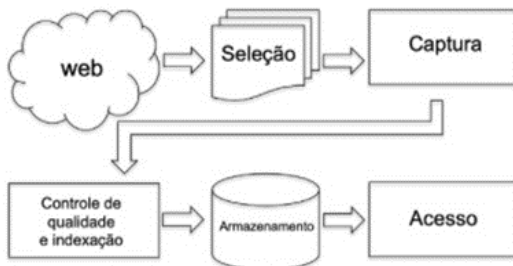


Figura 3 – Modelo de Arquivamento da Web

Fonte: Rockembach (2021, p. 7).

4. Conclusão

Esses diferentes modelos e abordagens destacam a complexidade e a importância de seguir um processo estruturado para o Arquivamento da Web. A definição de critérios de seleção, a compreensão da estrutura da web, a escolha de técnicas de captura adequadas e a garantia de qualidade são aspectos fundamentais para o sucesso das iniciativas de preservação da web. E, portanto, é posto que a inserção dos princípios de Organização do Conhecimento ajudam a estabelecer relacionamentos entre diferentes recursos da web, permitindo que os usuários naveguem por conteúdos relacionados, organizados com base em assuntos, tópicos ou relações semânticas. A partir disso, os usuários podem explorar e descobrir informações interconectadas dentro da coleção da web arquivada.

O Arquivamento da Web é uma área em constante evolução, impulsionada pela necessidade de preservar o conteúdo produzido na Internet, que é dinâmico e suscetível a mudanças e exclusões rápidas. E, sendo assim, conclui-se que as abordagens e metodologias levantadas atendem a este desafio. Entretanto, ainda há que se explorar a estrutura e as ferramentas fornecidas pela Organização do Conhecimento para que sejam garantidas a preservação a longo prazo e o acesso efetivo a informações e conhecimentos baseados na web. Nesta pesquisa, ainda que seja evidente a correlação entre essas duas áreas do conhecimento, não foram encontrados estudos específicos em Arquivamento da Web que tratam especificamente de técnicas e ferramentas advindas de estudos de Organização do Conhecimento.

É importante ressaltar que o Arquivamento da Web não é apenas uma questão técnica, mas também envolve aspectos políticos, jurídicos e sociais. As instituições responsáveis pelo Arquivamento da Web devem considerar esses diferentes fatores ao desenvolver suas estratégias e fluxos de trabalho. Além disso, é fundamental que haja uma colaboração entre os produtores de conteúdo, os

arquivos e a comunidade em geral para garantir o sucesso dessas iniciativas. O avanço das tecnologias e a evolução das práticas de preservação digital certamente contribuirão para aprimorar ainda mais a forma como o conteúdo web é selecionado, adquirido, organizado, armazenado, descrito e disponibilizado para acesso.

5. Referencias

- Abbate, J. (1999). *Inventing the internet*. MIT press.
- Bingham, N. J., & Byrne, H. (2021). Archival strategies for contemporary collecting in a world of big data: Challenges and opportunities with curating the UK web archive. *Big Data & Society*, 8(1), 2053951721990409.
- Bragg, M., & Hanna, K. (2013). The web archiving life cycle model. *Archive-It*. Recuperado de https://archive-it.org/static/files/archiveit_life_cycle_model.pdf.
- Brown, A. (2006). *Archiving websites: a practical guide for information management professionals*. facet publishing.
- Finnemann, N. O. (2014). Digitization: New trajectories of mediatization?.
- Finnemann, N. O. (2019). Web archive. *KO KNOWLEDGE ORGANIZATION*, 46(1), 47-70.
- Garcia Marco, Francisco Javier. 1995. "Fundamentos de organización y representación documental: aportaciones de la archivística." *Revista general de información y documentación* 5: 91-148.
- Gomes, D. (2010). Preservar a Web: um desafio ao alcance de todos. In *Actas do Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas* (pp. 1-9).
- Guizzo, É. M. (1999). *Internet, o que é, o que oferece, como conectar-se*. Editora Atica.
- Hjørland, B. (2008). What is knowledge organization (KO)? *KO Knowledge Organization*, 35(2-3), 86-101.
- Hjørland, B. Information Retrieval and Knowledge Organization: A Perspective from the Philosophy of Science. *Information* 2021, 12, 135. <https://doi.org/10.3390/info12030135>
- Jenkins, H., & Deuze, M. (2008). Convergence culture. *Convergence*, 14(1), 5-12.
- Khan, M., & Rahman, A. U. (2019). A systematic approach towards web preservation. *Information Technology and Libraries*, 38(1), 71-90.
- Lavoie, B. (2000). Meeting the challenges of digital preservation: The OAIS reference model. *OCIC Newsletter*, 243, 26-30.
- Lavoie, B. (2004). The open archival information system reference model: Introductory guide.
- Maemura, E., Worby, N., Milligan, I., & Becker, C. (2018). If these crawls could talk: Studying and documenting web archives provenance. *Journal of the Association for Information Science and Technology*, 69(10), 1223-1233.
- Masanés, J., & Masanés, J. (2006). *Web archiving: issues and methods* (pp. 1-53). Springer Berlin Heidelberg.
- Melo, J. F. (2020). Arquivamento dos websites do governo federal brasileiro: preservação do domínio GOV. BR. Dissertação de mestrado - Programa de Pós-graduação em Comunicação e Informação - Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil. 133 f.
- Melo, J. F., de Oliveira, C., & Rockembach, M. (2023, June 11). História do arquivamento da web no Brasil: um percurso entre a academia, o legislativo e executivo brasileiro. <https://doi.org/10.31235/osf.io/u4ydh>

- Melo, Jonas Ferrigolo & Rockembach, Moisés. Public Policies for Governmental Web Archiving in Brazil, presentation, May 12, 2023; (<https://digital.library.unt.edu/ark:/67531/metadc2143896/>: accessed September 20, 2023), University of North Texas Libraries, UNT Digital Library, <https://digital.library.unt.edu>; crediting International Internet Preservation Consortium.
- Mogollón, J. B. M., Dalessandro, R. C., & Tognoli, N. B. (2018). Metateoria, ciência da informação e organização do conhecimento: um estudo a partir da literatura científica da área. *Brazilian Journal of Information Science*, 12(2), 6-12.
- Németh, M., & Drótos, L. (2019). A blended learning-based curriculum on Web archiving in the national Széchényi library. *Digital Library Perspectives*, 35(2), 97-114.
- Niu, J. (2012). An overview of web archiving. *D-Lib magazine*, 18(3/4).
- Rockembach, M. (2018). Arquivamento da web: estudos de caso internacionais e o caso brasileiro. *RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação*, 16(1), 7-24.
- Rockembach, M. (2021). A web brasileira na Covid-19: arquivamento da web e preservação digital. *Liinc em Revista*, 17(1), e5713-e5713.
- Rockembach, M., & Pavão, C. M. G. (2018). Políticas e tecnologias de preservação digital no arquivamento da web. *Revista Ibero-Americana de Ciência da Informação*, 11(1), 168-182.
- Silva, A. M. da, & Ribeiro, F. (2011). Paradigmas, serviços e mediações em Ciência da Informação. *Recife: Néctar*, 122-138.
- Silva, A. M. da, & Ribeiro, F. (2012). Documentation/Information and their paradigms: characterization and importance in research, education, and professional practice. *KO KNOWLEDGE ORGANIZATION*, 39(2), 111-124.