

2023

RELATÓRIO PARCIAL 1: METAS 1 A 5

PROJETO DE PESQUISA: ECOSSISTEMA DE INFORMAÇÃO GOVERNAMENTAL

RELATÓRIO PARCIAL 1: METAS 1 A 5

PROJETO DE PESQUISA: ECOSSISTEMA DE INFORMAÇÃO GOVERNAMENTAL

PRESIDENCIA DA REPÚBLICA

Luiz Inácio Lula da Silva

Presidente da República

Geraldo José Rodrigues Alckmin Filho

Vice-Presidente da República

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO

Luciana Santos

Ministra da Ciência, Tecnologia e Inovação

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT)

Cecília Leite Oliveira

Diretora

Reginaldo de Araújo Silva

Coordenação de Administração - COADM

Ricardo Pimenta

Coordenação de Ensino e Pesquisa em Informação para a Ciência e Tecnologia - COEPI

José Luis dos Santos Nascimento

Coordenação de Planejamento, Acompanhamento e Avaliação - COPAV

Marcel Garcia de Souza

Coordenação-Geral de Informação Tecnológica e Informação para a Sociedade - CGIT

Bianca Amaro de Melo

Coordenação-Geral de Informação Científica e Técnica - CGIC

Tiago Emmanuel Nunes Braga

Coordenação-Geral de Tecnologias de Informação e Informática - CGTI

Milton Shintaku

Coordenação de Tecnologias para Informação - COTEC

Alexandre Faria de Oliveira

Coordenação de Governança em Tecnologias para Informação e Comunicação - COTIC



Ministério da Ciência, Tecnologia e Inovação
Instituto Brasileiro de Informação em Ciência e Tecnologia

RELATÓRIO PARCIAL 1: METAS 1 A 5

PROJETO DE PESQUISA: ECOSSISTEMA DE INFORMAÇÃO GOVERNAMENTAL



Brasília
2023

© 2023 Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict).

Esta obra é licenciada sob uma licença Creative Commons - Atribuição CC BY 4.0, sendo permitida a reprodução parcial ou total desde que mencionada a fonte.



EQUIPE TÉCNICA

Diretora do Instituto Brasileiro de Informação em Ciência e Tecnologia

Cecília Leite Oliveira

Coordenador-Geral de Tecnologias de Informação e Informática – CGTI

Tiago Emmanuel Nunes Braga

Coordenador do Projeto

Tiago Emmanuel Nunes Braga

Diagramação e projeto gráfico

Rafael Fernandez Gomes

Organizadores do relatório

Tiago Braga, Milton Shintaku, Alexandre Oliveira e Valéria Paiva.

Normalização

Larissa Alves, Marilete Silva e Ingrid Schiessl

Pesquisadores do projeto

Adirce Juliana Alves de Sena, Adriano Pellizzaro da Motta, Alecsander Cancio Sena, Alessandra de Oliveira Silva Rocha, Alisson Eugênio Costa, Amanda Cavalcante Pereira, Angelica Alves da Cunha Marques, Arthur Coelho Bezerra, Bruno Leonardo Costa Alves Silva, Celso Moreira Ferro Júnior, Deivdy William Silva, Diego Ramalho Lopes, Douglas André Muller, Elias Suaiden Neto, Elizabete Cristina de Souza de Aguiar Monteiro, Emir Jose Suaiden, Erelisiane Cristine Vieira, Fabiano Couto Corrêa da Silva, Felipe de Castro Teixeira, Felipe Moura da Silva, Fernanda Maciel Rufino, Fernando de Jesus Pereira, Flavia Furlan Granato, Flavia Karla Ribeiro Santos, Flavio Castro de Sousa, Flor de Maria Silvestre Estela, Frederico Ramos Oliveira, Graziela Barros Gomes, Heliomar Medeiros de Lima, Henrique Denes Hilgenberg Fernandes, Hesley da Silva Py, Ingrid Torres Schiessl, Jane Fontes Gadelha, João Victor Santos Aguiar, Jordana Peres Padovani, José Carlos Abbud Grácio, José Paulo Pereira das Dores Savioli, Jose Wilson da Costa, Josir Cardoso Gomes, Karen Martins Vitoria, Karoline Leonardi de Sousa, Ketlen Stueber, Larissa Couto, Larissa de Araújo Alves, Laura Vilela Rodrigues Rezende, Lena Vania Ribeiro Pinheiro, Lilian Maria Araujo de Rezende Alvares, Luana Farias Sales Marques, Lucas Ângelo da Silveira, Lucas Rodrigues Costa, Luciano Heitor Gallegos Marín, Marcelle Costal de Castro dos Santos, Marcello Mundim Rodrigues, Márcio Bezerra da Silva, Marco André Feldman Schneider, Miguel Angel Mardero Arellano, Mirele Carolina Souza Ferreira Costa, Mirella de Souza Balestero, Moisés Lobo de Souza Choas, Natália Nakano, Nathaly Cristine Leite Rocha, Nikole Barros Mota, Nuielle Cristine de Medeiros da Silva, Pâmela Isabela Ferreira Lima, Pâmela Raquel Silva da Cruz, Patrícia Boquady Barros de Sousa, Patricia Veronica Moreira, Paulo Cesar Gonçalves Egler, Priscila Rodrigues dos Santos, Rafael de Souza Nobre, Rafael Fernandez Gomes, Rafael Teixeira de Souza, Raíssa da Veiga de Meneses, Rebeca dos Santos de Moura, Renata Monteiro Rodrigues, Renata Muhlebeier, Richard William Valdivia, Ronaldo de Castro Ianotti, Rondineli Gama Saad, Rossana Coely de Oliveira Moura, Sarah Costa Schmidt, Sérgio Luiz de Menezes Filho, Silvana Aparecida Borsetti Gregorio Vidotti, Sonia Araújo de Assis Boeres, Valéria Gameleira da Mota, Victor Baessa de Queiroz, Victor Guerresi de Mello Braga, Virgínia Ferreira da Silva Castro, Vitor Mateus Costa do Rego, Vivian Iwane, Welber Amaro Santos de Souza, Wesley Lopes Ferreira Santos

Este Relatório de Técnico é um produto do Projeto: Ecossistema de Informação Governamental

Ref. IBICT - Processo SEI 01302.000382/2021-72

Ref. FUNDEP 29178

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto Brasileiro de Informação em Ciência e Tecnologia ou do Ministério da Ciência, Tecnologia e Inovação.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.



Setor de Autarquias Sul (SAUS) Quadra 05 Lote 06, Bloco H – 5º andar Cep:70.070-912 – Brasília, DF Telefones: 55 (61) 3217-6360/55 / (61) 3217-6350 - <https://www.gov.br/ibict/pt-br>

Sumário

1.	INTRODUÇÃO	9
2.	RESULTADOS META 1	11
2.1	Objetivo geral da Meta 1	11
2.2	Objetivos específicos da Meta 1	11
2.3	Descrição de resultados da Meta 1	12
2.3.1	Definir amostra de sistemas de informação a ser contemplado pelo estudo	12
2.3.2	Levantar, registrar e analisar tecnologias utilizadas ou apropriadas para implementação de sistemas de informação no governo	13
2.3.3	Desenvolver modelos de sistemas de informação governamentais	14
2.3.4	Propor formas de integração de sistemas que utilizam diferentes tecnologias	16
2.4	Considerações finais sobre a Meta 1	16
3.	RESULTADOS META 2	17
3.1	Objetivo geral da Meta 2	17
3.2	Objetivos específicos da Meta 2	17
3.3	Descrição de resultados da Meta 2	18
3.3.1	Prospecção de modelos de preservação digital	18
3.3.1.1	Iniciativas de arquivamento: exemplos internacionais	18
3.3.1.2	Iniciativas de arquivamento: exemplos nacionais	21
3.3.2	Análise e avaliação de soluções tecnológicas para preservação a longo prazo de documentos oficiais governamentais	22
3.3.2.1	Legislação: panorama mundial	22
3.3.2.2	Legislação: panorama nacional	24
3.3.2.3	Um pouco sobre repositórios e política de arquivamento	29
3.3.2.4	Tecnologia para arquivamento <i>Web</i>	30
3.3.2.5	Modelos de ciclo de vida de dados na <i>Web</i>	32
3.3.3	Modelo de preservação de sites governamentais	33
3.3.3.1	Arquivamento da <i>Web</i> e o modelo OAIS	35
3.3.3.2	Descarte de conteúdos <i>Web</i>	37
3.3.3.3	Iniciativas e estratégias de arquivamento e preservação digital	38
3.3.4	Modelos de formatos adequados para preservação de dados e documentos digitais governamentais	40
3.3.4.1	Políticas de arquivamento de páginas governamentais	43

3.3.5	Estudos para o aprimoramento de modelos de preservação digital arquivística	44
3.4	Considerações finais sobre a Meta 2	46
4.	RESULTADOS META 3	49
4.1	Objetivo geral da Meta 3	49
4.2	Objetivo específico da Meta 3	50
4.3	Descrição de resultados da Meta 3	50
4.3.1	Desenvolver modelo e orientações para criação ou conversão para documentos digitais acessíveis	50
4.3.2	Criar ou converter documentos para o modelo	52
4.3.2.1	O Tesouro de Documentos Acessíveis	52
4.3.2.2	Descrição padrão do termo em linguagem simples	53
4.3.2.3	Gravação e edição dos vídeos traduzidos em linguagem brasileira de sinais	53
4.3.2.4	Criação do website e apresentação da informação acessível, e avaliação do website	53
4.3.3	Avaliar e ajustar modelos e orientações	54
4.4	Considerações finais sobre a Meta 3	54
5.	RESULTADOS META 4	55
5.1	Objetivo geral da Meta 4	55
5.2	Objetivo específico da Meta 4	55
5.3	Descrição de resultados da Meta 4	56
5.3.1	Estudos para verificar as relações entre ciência de serviços, informação estratégica, Business Intelligence, e-GOV, entre outros e informação e dados geolocalizados	56
5.3.2	Realizar estudos para implementação de melhorias no Visão para atendimento às novas funcionalidades	56
5.3.2.1	Objetivos principais e secundários do sistema Visão definidos	57
5.3.2.2	Requisitos da versão 2.0 do sistema Visão definidos	57
5.3.2.3	Protótipo da versão 2.0 do sistema validado	59
5.3.2.4	Definição e implementação de catálogo de metadados do sistema Visão	59
5.3.2.5	Ferramentas para <i>upload</i> dos dados para a plataforma	61
5.3.2.6	Desenvolvimento de área para gerenciamento de repositório pessoal de dados	61
5.3.2.7	Grupo de usuários	62
5.3.2.8	Visualizações secundárias	62
5.3.2.9	Integração de ferramentas	62
5.3.3	Desenvolver modelo de integração de dados de sistemas para o Visão	62

5.3.3.1	Perfil de metadados, baseado no ISO 19115/19139 e no Perfil Brasileiro de Metadados Geoespaciais (PMGB)	62
5.3.3.2	Disseminação de conhecimentos e novas funcionalidades do Visão junto a potenciais usuários/parceiros	70
5.4	Considerações finais sobre a Meta 4	70
6.	RESULTADOS META 5	72
6.1	Objetivo geral da Meta 5	72
6.2	Objetivo específico da Meta 5	72
6.3	Descrição de resultados da Meta 5	73
6.3.1	Desenvolvimento de documentação técnica sobre os estudos	73
6.3.1.1	Livro - Ecossistemas e inovações tecnológicas [recurso eletrônico]: da construção às boas práticas	73
6.3.1.2	<i>Livro - Hipátia: modelo de preservação para repositórios arquivísticos digitais confiáveis</i>	74
6.3.2	Desenvolvimento de documentação científica sobre os estudos	76
6.3.2.1	Artigo de periódico publicado em periódicos nacionais	76
6.3.3	Apresentação dos resultados de estudo em eventos	76
6.3.3.1	Conversa com especialistas: Tecnologias de Informação e Informática apoiadas pelo Instituto Brasileiro de Informação em Ciência e Tecnologia	76
6.3.3.2	Apresentação de trabalhos no evento XXII Encontro Nacional de Pesquisa em Ciência da Informação (Enancib)	77
6.3.3.3	Apresentação de trabalhos no evento XIII Encuentro EDICIC	77
6.3.3.4	Apresentação de trabalhos no evento Workshop de Informação, Dados e Tecnologia (V WIDaT)	78
6.3.3.5	Apresentação de trabalho no evento Congresso Nacional de Arquivologia (CNA 2022)	78
6.3.4	Reconhecimento pelo Selo Nacional de Modernização do Estado - 2022	78
6.4	Considerações finais sobre a Meta 5	79
7.	CONSIDERAÇÕES FINAIS	80
	REFERÊNCIAS	81

1. INTRODUÇÃO

A presença do governo é constante na vida dos cidadãos e se manifesta por meio de diversas fontes de informação, seja de forma direta ou por intervenção de intermediários, especialmente os jornalistas. Nesse aspecto, a comunicação governamental é semelhante à comunicação científica, uma vez que ambas envolvem dois processos: disseminação e divulgação. No caso da comunicação governamental, a disseminação ocorre de forma oficial, direcionada aos membros do próprio governo, enquanto a divulgação é realizada por jornalistas, muitos deles especializados em política.

Com a evolução da tecnologia e a adoção do governo eletrônico (e-GOV), ultimamente, a informação governamental tornou-se mais acessível. Isso se deve, em grande parte, à publicação da Lei de Acesso à Informação (LAI). Ações de transparência, como os dados abertos de governo e a Lei Orçamentária Anual (LOA), por exemplo, possibilitam o acesso a dados e informações governamentais, mas sempre de forma particionada. No entanto, em nível estadual, municipal e distrital, muitas vezes, ainda há barreiras técnicas e tecnológicas a serem transpassadas no livre acesso à informação e dados governamentais.

Independentemente da esfera ou poder, ocorrem possivelmente dois grandes desafios no governo. O primeiro é a ausência de alguns sistemas informatizados para gerir a informação, assim como o isolamento dos sistemas. O segundo é a possibilidade de problemas relacionados à acessibilidade dos documentos tramitados e sua preservação por longos períodos.

Os desafios relacionados à apresentação de informações e dados governamentais em formato geolocalizado são crescentes. Essa modalidade de apresentação oferece facilidades, mas também exige precisão nas informações. Em muitos casos, a geolocalização é fundamental para contextualizar ou compor indicadores e informações sobre entidades, entre outros.

A pesquisa aplicada é fundamental para o avanço da ciência e da tecnologia, pois permite a geração de conhecimento e a criação de modelos, aplicações, *softwares* e outros resultados práticos que melhoram a vida das pessoas. No entanto, a Ciência da Informação ainda apresenta poucos estudos e linhas de pesquisa relacionadas à informação governamental. Por isso, estudos inovadores nessa área são importantes para contribuir com o desenvolvimento da pesquisa e do conhecimento. Nesse sentido, foi proposto o projeto de pesquisa *Ecossistema de Informação Governamental*, que tem como principal objeto de investigação a realização de *estudos voltados à estruturação de modelo para criação e gestão de ecossistema de informação governamental*.

Além do objeto, o escopo da pesquisa desenvolvida no contexto do referido projeto se consolida por meio do objetivo geral, que trata sobre a *criação de modelo de ecossistema de informação governamental, incluindo estratégias para sua implementação, gestão e manutenção*. Tendo em vista o alcance de tal objetivo geral, o plano de trabalho do projeto explicita as seguintes metas:

Meta 1 - Desenvolver estudos para criação de modelo de integração de sistemas informacionais governamentais considerando as diversas tecnologias utilizadas na sua implementação.

Meta 2 - Estudos para criação de modelo de preservação de dados e informações governamentais alinhado às normas de padrões existentes.

Meta 3 - Desenvolver modelos e orientações para criar ou converter documentos digitais, de forma a torná-los acessíveis, contribuindo com a democratização da informação.

Meta 4 - Desenvolver estudos para criação de novas funcionalidades informatizadas para apresentação de dados e informações geolocalizadas amparada na ciência da informação.

Meta 5 - Disseminar resultados de pesquisas.

O presente relatório busca apresentar, de forma consolidada, os resultados parciais dos estudos realizados pela equipe de pesquisadores desde a implementação do projeto de pesquisa até março de 2023. O conteúdo do relatório encontra-se estruturado seguindo a ordem das cinco metas definidas para a pesquisa no plano de trabalho oficial. Logo, para cada meta, é elencado um capítulo de resultados correspondentes.

O Quadro 1, abaixo, demonstra a associação entre o enunciado das metas e a porcentagem de andamento concluído para cada uma delas. Os resultados explicitados nas seções e capítulos subsequentes desse relatório encontram-se correlacionados ao *status* de desenvolvimento explicitado no Quadro 1.

META	DESENVOLVIMENTO (%)
Desenvolver estudos para criação de modelo de integração de sistemas informacionais governamentais considerando as diversas tecnologias utilizadas na sua implementação	50%
Estudos para criação de modelo de preservação de dados e informações governamentais alinhado às normas de padrões existentes	85%
Desenvolver modelos e orientações para criar ou converter documentos digitais, de forma a torná-los acessíveis, contribuindo com a democratização da informação	90%
Desenvolver estudos para criação de novas funcionalidades informatizadas para apresentação de dados e informações geolocalizadas amparada na ciência da informação	75%
Disseminar resultados de pesquisas	60%

Quadro 1 - Andamento das metas em porcentagem de desenvolvimento de atividades.

2. RESULTADOS META 1

Desde o surgimento da informática, ainda na década de 1940, durante os esforços de guerra, seguiu-se com a informatização dos sistemas de informação. Entretanto, até a década de 1980, ainda com a tecnologia dos grandes computadores (*mainframes*), havia poucas iniciativas de integração devido às diferentes tecnologias utilizadas e às incapacidades impostas à época. Desde então, cada vez mais procura-se integrar sistemas de informação por meio de seus sistemas informatizados, com compartilhamento de dados e informações.

Assim, padrões e técnicas foram desenvolvidos para possibilitar o intercâmbio de informações e dados. Dentre os quais pode-se citar o *eXtensible Markup Language* (XML) e o *JavaScript Object Notation* (JSON), que aliam orientações sintáticas e semânticas de forma que humanos e máquinas possam compreender, facilitando a possibilidade de integração entre sistemas informatizados. Entretanto, parte das ferramentas atuais ainda não apresentam formas automáticas de integração, possibilitando apenas a exportação de dados nesses padrões mais comuns.

Nesse sentido, são necessários estudos que visem a integração de dados e informações de sistemas informatizados, de forma que possam compartilhar dados e informações para melhoria na execução de atividades de uma instituição ou organização. Nota-se que dados ou informações de um sistema informatizado muitas vezes são requeridos em outros, sendo replicados, o que pode levar a falhas ou diminuição da eficiência.

Nesse contexto, fazem-se necessários os estudos voltados para integrar os sistemas de informação voltados para gestão da informação, na medida em que atendem a toda instituição ou organização. A informação, como capital, requer sistemas informatizados para serem geridos, demandando integração para melhor atender às necessidades.

2.1 Objetivo geral da Meta 1

Desenvolver um modelo de integração de sistemas informacionais governamentais considerando as diversas tecnologias utilizadas.

2.2 Objetivos específicos da Meta 1

- a. Definir amostra de sistemas de informação a ser contemplado pelo estudo;
- b. Levantar, registrar e analisar tecnologias utilizadas ou apropriadas para implementação de sistemas de informação no governo;
- c. Desenvolver modelos de sistemas de informação governamentais;
- d. Propor formas de integração de sistemas que utilizam diferentes tecnologias.

2.3 Descrição de resultados da Meta 1

Os estudos preliminares apontam alguns resultados que vão orientar as direções das metas seguintes. Com o projeto em andamento, alguns desses resultados ainda podem sofrer mudanças, na medida em que outros resultados mostram-se relevantes e divergem do que já está estabelecido. Assim, aqui, apresentam-se os resultados preliminares do projeto.

2.3.1 Definir amostra de sistemas de informação a ser contemplado pelo estudo

O Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), unidade de pesquisa do Ministério da Ciência, Tecnologia e Inovação (MCTI), em especial a Coordenação de Tecnologias para a Informação (Cotec), atua no apoio à disseminação de tecnologias para gestão da informação. Dessa forma, o projeto inicia com esses sistemas, que atendem a sistemas de informação importantes para a missão do instituto.

Sistemas de publicação da informação atendem, em grande parte, aos usuários que querem criar revistas e editoras, podendo ser técnicas ou científicas. Nesse caso, o Ibict apoia as ferramentas mantidas pelo *Public Knowledge Project* (PKP), como o *Open Monograph Press* (OMP), para editoras, e o *Open Journal Systems* (OJS), para criação de portais de revistas técnicas e científicas. Importante ressaltar, que em 2018, com o PKP completando duas décadas, Alperin, Willinsky, Owen, Macgregor, Smecher e Stranack (2018) apresentam algumas premissas que orientam as suas atividades, como uma organização sem fins lucrativos, em constante interação com a comunidade usuária.

Bibliotecas digitais e repositório são sistemas de informação importantes para a gestão da informação. Tanto que, grande parte dos órgãos de governo tem criado suas bibliotecas digitais para atendimento à Lei de Acesso à Informação (LAI) de forma ativa, disponibilizando documentos on-line, mesmo sem a requisição, que não estão protegidos por questões legais. Nesse caso, tem-se algumas ferramentas que podem ser utilizadas para criação de bibliotecas digitais ou repositórios, como DSpace, Omeka, Tainacan, entre outros. Pode-se afirmar que as bibliotecas digitais ganharam maior espaço nas organizações, principalmente junto aos órgãos de governo, muitas por meio do modelo desenvolvido pela Biblioteca Digital Jurídica (BDJur), apresentada por Basevi (2005), voltado a disseminar a memória técnica do Superior Tribunal de Justiça (STJ).

Assim como os repositórios de documentos, os dados requerem sistemas de informação para sua gestão. Seguindo a orientação do dados.gov.br, selecionou-se o *Comprehensive Knowledge Archive Network* (CKAN) para ser a tecnologia apoiada pelo instituto e ter estudos sobre integração de dados e informações. Grande parte dos órgãos de governo, incluindo o Ibict, possui repositórios de dados implementados com o CKAN.

Outro tipo de informação que merece atenção são as imagens, em grande parte fotografias, que com a digitalização teve um grande aumento, visto a facilidade de produção. Nesse caso, os bancos de imagens são sistemas de informação que merecem estudos. Para tanto, o software selecionado é o Piwigo, que atende grande parte das necessidades de criação de banco de imagens.

Sistemas para representação da informação são importantes por possibilitarem atuar com a terminologia de uma linguagem de especialidade. Esses sistemas são utilizados na biblioteconomia muito antes da informatização, principalmente para bases de autoridade. Assim, pode-se ter estrutura simples como listas de termos até as mais complexas, como as ontologias. No caso, a ferramenta selecionada para atuar nesses sistemas de informação é o Tematres. Tal software possui potencial interoperabilidade, essa capacidade de comunicação de um sistema com outro, que pode ser classificada em níveis e/ou tipos. Andrade e Cervantes (2012) trazem, em sua pesquisa, o conceito de interoperabilidade técnica, semântica, organizacional, política e humana, intercomunitária, legal e internacional. Zeng (2019), por sua vez, apresenta um modelo de interoperabilidade que engloba a sistêmica, a semântica, a sintática e a estrutural. Já Moreiro González *et. al* (2012) tratam da interoperabilidade técnica, semântica e organizacional. A exploração da temática e a prática da interoperabilidade são evidentes e tornam instigante, e ao mesmo tempo desafiador, o processo de implementação em instrumentos de controle de vocabulário, ainda pouco desenvolvido, assim como o uso do software para a sua gestão, o Tematres.

O Ibict, historicamente, tem atuado com bibliotecas desde o seu nascimento. Com isso, tem promovido os Sistemas de Gestão de Bibliotecas, que tem no Koha a ferramenta mais utilizada no mundo. Assim, atende o legado do Instituto, com estudos voltados a integrar sistemas de bibliotecas que custodiam o acervo físico com sistemas de informação de tratamento digital.

Por fim, os Gestores de conteúdo têm sido muito utilizados para a criação de sites e portais, com ferramentas conhecidas como *Content Management System* (CMS). Dentre todas as opções existentes, o Ibict tem optado por trabalhar com o WordPress, pela sua facilidade de uso e simplicidade nos códigos. Em muitos casos, utiliza-se o CMS como camada de apresentação para outros sistemas de informação, ou mesmo ecossistemas de informação.

2.3.2 Levantar, registrar e analisar tecnologias utilizadas ou apropriadas para implementação de sistemas de informação no governo

Essa etapa do projeto ainda está em andamento, mesmo que já tenhamos alguns resultados, principalmente com novas tecnologias selecionadas para atuação do instituto. Assim, de forma inicial, temos o seguinte quadro de tecnologias utilizadas para implementação de sistemas de informação no governo (Quadro 2):

Item	Sistema de informação	Software	Observação
1	Repositório de documentos	DSpace	Utilizado em vários órgãos de governo, principalmente no Judiciário, com o modelo da Biblioteca Digital Jurídica (BDjur)
2	Repositório de dados	CKAN	Utilizada pelo dado.gov.br, além do próprio Ibict para os dados de governo
3	Repositório de código	GitLab	Utilizada para distribuição de códigos, incluindo o Ibict
4	Repositório de imagens	Piwigo	Ainda pouco utilizado no Brasil

Item	Sistema de informação	Software	Observação
5	Repositório de arquivo	AToM	Utilizada para disseminar documentos junto com o archivematica
6	Repositório de museu	Omeka, Tainacan	Utilizada pelo Ibram, Iphan, TJDFT e outros
7	Gestão de conteúdo	WordPress	Grande parte dos órgãos de governo
8	Gestão de acervo	Koha	Utilizada pelo Ibict, Ibram e Enap
9	Sistema de publicação	OJS e OMP	Grande parte das universidades, além das revistas mantidas pelos órgãos de governo
10	Terminologia	Tematres	Utilizada pela Enap e outros
11	Visualização	Visão	Utilizada pela Enap e Ibict

Quadro 2 - Lista de tecnologias utilizadas em sistemas de informação governamental

Fonte: Elaborado pelos autores (2023).

Com o desenvolvimento dos estudos, pode ser que outras ferramentas sejam adicionadas à lista atual, na medida em que necessidades sobre a informatização de sistemas de informação no governo possam surgir, visto que a gestão da informação governamental está sempre em construção. Assim, o levantamento inicial do projeto ainda é tímido, mas consistente.

2.3.3 Desenvolver modelos de sistemas de informação governamentais

Sistemas de informação governamentais nascem da necessidade de atender a atividades, na medida em que, na definição clássica, envolve processos, atividades, padrões, métodos, pessoas e um sistema informatizado. Assim, requer o entendimento dos processos governamentais para a proposição de sistemas de informação. Entretanto, para gerir informações, sabe-se que requer sistemas que garantam a seleção, armazenamento e disseminação da informação registrada por meio de documentos que compõem a chamada memória técnica.

Nesse sentido, memória técnica como relatado por Andrade, Shintaku e Barros (2018), assume grande variedade documental, requerendo cuidado para a sua organização e descrição, tendo em vista a complexidade. O termo memória técnica, para indicar a documentação produzida pelos colaboradores do governo que não passou por processo editorial, não é um consenso, visto que pode encontrar-se outras denominações, como produção intelectual, memória institucional, memória organizacional etc. Assim, fez-se uso do termo memória técnica para designar essa documentação, conforme a definição de González de Gómez e Machado (2007), descrevendo-a com um conteúdo que explicita procedimentos, muitas vezes considerados como literatura cinzenta, com forte relação administrativa.

Assim, alguns modelos para desenvolvimento de sistemas de informação têm sido pensados com base nos resultados de estudo, como no caso de bibliotecas digitais, no qual o uso de cadeia de valor se mostrou um instrumento valioso para a sua organização. Esses resultados foram apresentados em artigo científico¹, no qual Shintaku (2023) defende que o uso da cadeia de valor é utilizado para levantar a documentação e o fluxo informacional que culminará na biblioteca digital. Assim, modelos de sistemas de informação precisam identificar formas de implementação.

Nesse sentido de cadeia de valor, como uma forma orientadora para desenvolvimento de sistema de informação, destaca-se a percepção de Albano (2013), para os dados abertos de governo, o autor discute a cadeia de valor especificamente na questão das informações de forma generalizada. Assim, as atividades primárias estariam relacionadas com: geração, processamento, armazenamento, oferta de serviços, distribuição e apresentação da informação. O governo, em grande parte, é uma organização que oferta serviços com vasta quantidade de informações, geradas em maior número internamente pelas suas mais diversas atividades.

Kaplinsky e Morris (2000) apresentam as variações em torno das definições da cadeia de valor, diferenciando os seguintes tipos:

- **Cadeia de valor simples:** descreve as vastas atividades requeridas para trazer um produto ou serviço, desde o início ao seu final, incluindo as atividades após o uso.
- **Cadeia de valor estendida:** descreve também o contexto e as atividades relacionadas à produção do produto ou serviço.

Um dos pontos que se pode destacar nas definições dos autores é a adição de serviços nas análises da cadeia de valor. Com eles, há a clara indicação de que a cadeia de valor pode ser utilizada para análise de empresas que ofertam serviços, transcendendo os conceitos iniciais voltados à manufatura de produtos. Logo, entende-se que a cadeia de valor pode ser utilizada para análise de órgãos de governo.

Numa aplicação do ciclo de vida da informação, Moresi (2000) propõe uma cadeia de valor para sistemas de informação e, em cada etapa dos processos, agrega-se valor ao dado, incluindo as realimentações necessárias. De certo modo, tal cadeia de valor da informação para a tomada de decisão ainda figura nas atividades de suporte, mesmo que colete dados das atividades principais. Assim, a fonte de dados está nas atividades primárias, mas todas as atividades de agregação de valor da cadeia ocorrem em atividades de suporte.

1 Disponível em: <https://revistas.ufpr.br/atoz/article/view/84584/48447>

2.3.4 Propor formas de integração de sistemas que utilizam diferentes tecnologias

O compartilhamento de dados e informações provenientes de sistemas de informação distintos é um desafio, resultado do projeto. Entretanto, em alguns casos, pode-se pensar em ecossistemas de informação muito complexos em que haja a necessidade de um hub de informações. Em outros casos, uma formulação em rede, do formato em grafo, deve atender, pois o compartilhamento de dados e informações é pontual. Entretanto, o estudo ainda não chegou a esse ponto de ter resultados consistentes para um modelo, ou modelos.

2.4 Considerações finais sobre a Meta 1

A evolução da gestão pública e o surgimento e implantação de técnica de governança requerem o uso de sistemas que apoiem a gestão da informação e promovam a transparência. Fortini e Shermam (2017) reforçam que a gestão pública requer medidas de controle interno que possibilitem a auditoria com vistas a precaver problemas. Entretanto, isso só é possível se a documentação estiver organizada e disponível para acesso, dando maior transparência, só sendo possível por meio de bibliotecas digitais, sistemas de acesso a documentos arquivísticos, entre outros.

O estudo em andamento tem apresentado resultados parciais, no qual ratifica o caminho do Ibict em tratar de sistemas de informação e formas de integrá-los para ter melhores resultados no cumprimento da missão da instituição e organização. É correto afirmar que: a informatização serve para dar maior eficiência às atividades humanas. Entretanto, nota-se não haver ilhas isoladas nessas atividades. Com isso, faz-se necessário a integração.

3. RESULTADOS META 2

O governo está presente na vida cotidiana dos cidadãos, principalmente por intermédio da disseminação de informações direta ou indiretamente veiculadas por canais midiáticos. Nesse ponto, assemelha-se à comunicação científica, no qual possui dois processos, o de disseminação e o de divulgação. Com isso, na comunicação governamental, a disseminação é feita no formato oficial para os pares, ou seja, para outros membros do governo, mas a divulgação é feita por jornalistas, grande parte especializados em política. Além de elaborar, comunicar e promover acesso às informações, é de extrema importância que o acesso possa ser garantido, tendo em vista a manutenção da integridade e da veracidade dessas informações a longo prazo. É nesse contexto que se destaca a relevância da preservação digital na proposição de um ecossistema de informação governamental.

Muito mais do que um *backup*, a preservação digital se preocupa não só em manter íntegro e acessível o documento em si, mas também o seu contexto e os metadados responsáveis por descrevê-lo e identificá-lo. Na literatura de Ciência da Informação, quando se trata de preservação digital, é apresentado o Modelo de referência Open Archival Information System (OAIS), que recomenda as seguintes etapas base para preservação: recolhimento dos dados no sistema produtor de documentos, extração, empacotamento, preservação e disseminação. Diante de tal referência, a Meta 2 do projeto procura desenvolver investigações em busca da elaboração de um modelo voltado para a preservação de dados e informações governamentais.

3.1 Objetivo geral da Meta 2

Estudos para criação de modelo de preservação de dados e informações governamentais.

3.2 Objetivos específicos da Meta 2

- Prospecção de modelos de preservação digital;
- Análise e avaliação de soluções tecnológicas para preservação a longo prazo de documentos oficiais governamentais;
- Desenvolver modelo de preservação de sites governamentais;
- Desenvolver modelos de formatos adequados para preservação de dados e documentos digitais governamentais;
- Estudos para o aprimoramento de modelos de preservação digital arquivística.

3.3 Descrição de resultados da Meta 2

Os conteúdos apresentados nas próximas subseções descrevem os resultados alcançados organizados pelos objetivos específicos da Meta 2.

3.3.1 Prospecção de modelos de preservação digital

Como avanço tecnológico, cada vez mais as publicações estão disponíveis exclusivamente na *Web*, entretanto, a informação é mais volátil do que a publicada em documentos impressos. Os arquivos da *Web* são sistemas informatizados criados para realizar o arquivamento automático da informação digital e partilham o mesmo objetivo que os arquivos ou bibliotecas: a preservação de conhecimento para as gerações futuras. Nesse sentido, várias iniciativas de arquivamento da *Web* são idealizadas em todo o mundo com o objetivo de enfrentar os desafios da preservação digital.

3.3.1.1 Iniciativas de arquivamento: exemplos internacionais

Alguns dos pioneiros nas iniciativas internacionais de preservação da *web* são o *Internet Archive*, o projeto da Biblioteca Nacional da Austrália - PANDORA (do inglês, *Preserving and Accessing Networked Documentary Resources of Australia*), e o *Kulturarw3*, da Suécia, iniciados em 1996. Merece destaque o trabalho do Consórcio Internacional de Preservação da Internet (IIPC, do inglês *International Internet Preservation Consortium*), criado em 2003 e que se dedica ao desenvolvimento de padrões e ferramentas que auxiliam no processo de arquivamento da *Web*. (INTERNATIONAL INTERNET PRESERVATION CONSORTIUM, 2021). Os métodos empregados nos arquivos da *Web* vêm sendo estabelecidos por ele por meio da troca de experiências entre as instituições membros. A maioria delas se situa em países europeus, sendo que a única sul-americana, a Biblioteca Nacional do Chile, que se filiou em 2013 (CARRETA, 2019).

Day (2003) publicou uma pesquisa sobre iniciativas de preservação da *Web*, destacando a relevância de algumas organizações: *Internet Archive* (sediado nos Estados Unidos, mas de abrangência Internacional), *Kulturarw3* (Suécia), *Bibliothèque Nationale de France* (França), AOLA (Áustria), PANDORA (Austrália), *Helsinki University Library* (Finlândia), *Britain on the Web* (Reino Unido) e MINERVA (Estados Unidos). Em pesquisa realizada por Gomes, Miranda e Costa (2011), foram levantadas e analisadas 42 iniciativas de arquivamento da *Web* ao redor do mundo.

No Brasil, Rockembach (2018, p. 17) apontou a iniciativa do Núcleo de Pesquisa em Arquivamento da *Web* e Preservação Digital (NUAWEB) da Universidade Federal do Rio Grande do Sul (UFRGS), criado em 2017. O autor ainda cita as iniciativas *The National Archives*; a da Biblioteca de Catalunha (Espanha); a da *Bibliothèque Nationale de France* (França); a da *Harvard Library* (Estados Unidos); o provedor de serviço *HanzoArchives Limited* (Reino Unido); o *UK Web Archive* (UKWA) e o sistema distribuído Sistema de Replicação Colaborativa (rARC) que permite a qualquer utilizador da Internet colaborar para a preservação da *Web*. Cada contribuidor instala uma aplicação e cede espaço, em seu computador, para guardar parte da informação arquivada no repositório central. Em caso de perda destes conteúdos (arquivados no repositório), a informação do arquivo da *Web* será recuperada a partir das cópias guardadas nos computadores dos contribuidores. (GOMES, 2010, p. 6).

De acordo com Ferreira (2019), alguns exemplos de iniciativas são as plataformas *Internet Archive*, dos Estados Unidos, *Library and Archives of Canada* e o *Arquivo.pt*, de Portugal. Este último tem crescido em importância no mundo, a nível de comparação, o Brasil é o terceiro país que mais utiliza o *Arquivo.pt*, depois de Portugal e dos Estados Unidos. Em termos econômicos, o próprio serviço fez uma estimativa de quanto poderia valer, em dinheiro, seu patrimônio, e o resultado, a partir de dados de 2017, foi de que ele seria superior ao PIB de Portugal no mesmo ano (CARRETA, 2019).

É de importância universal que os conteúdos postados somente na *Web* e nas mídias sociais sejam preservados, pois são uma parte fundamental da história da sociedade atual, que pode ser modificada ou excluída como se nunca tivesse existido. Para Rockembach (2018, p. 17) a maior das iniciativas é a *Internet Archive*, sem fins lucrativos, fundada em 1996, é considerada uma das mais antigas iniciativas de arquivamento da *Web*. Dados de 2017 mostravam que o *Internet Archive* havia coletado e mantido arquivados cerca de 286 bilhões de páginas e mais de 361 milhões de *websites*. A iniciativa *Internet Archive* é um dos membros fundadores do Consórcio Internacional de Preservação da Internet, e dispõe de uma ferramenta bastante popular, o *Wayback Machine*, que permite acessar como um determinado *site* era em determinada época (*snapshot*). Para tanto é necessário inserir o domínio a ser pesquisado e, com os resultados da pesquisa, selecionar alguma data específica a partir dos dados capturados.

Por meio do *Wayback Machine* pode-se identificar a coleta de dados da *Web* de forma exaustiva e a coleta por domínios de primeiro nível, nacionais e regionais, seletivo, por eventos, temático, coleta em massa, *survey*, *sites* que serão encerrados (em fim de vida) e *sites* específicos. A coleta também pode ser feita de forma ativa pelo usuário, na página da ferramenta, onde é possível inserir o endereço eletrônico que se pretende capturar e salvar a página no banco de dados da iniciativa *Internet Archive*, permitindo, futuramente, fazer uma citação de tal referência. Ressalta-se que essa ferramenta só está acessível para *sites* que permitem o uso de rastreadores da *Web* (*crawlers*) para a indexação desses *sites*, mesma tecnologia utilizada por motores de busca. Diferentes tipos de documentos são arquivados pela *Internet Archive* não só *sites* da *Web*, mas também livros, textos, filmes, softwares, música e imagens e coleções das mais diversas temáticas. (ROCKEMBACH, 2018, p. 17).

Ainda sobre essa ferramenta, Alam *et al.* (2019) afirmam que a última fase no fluxo de preservação dos *sites* arquivados é permitir que o conteúdo esteja disponível para as partes interessadas. Por meio de softwares como o *Open Wayback* é possível reproduzir os arquivos *WARCs* armazenados. Nessa fase os responsáveis pela curadoria conseguirão validar se a estrutura e o conteúdo do *website* foram mantidos. Além de utilizar a ferramenta *WARC*, os autores sugerem o Protocolo Memento (financiado pelo Programa Nacional de Infra-estrutura e Preservação de Informação Digital (NDIIPP) para negociação de conteúdo *datetime* na reprodução. Ele especifica a negociação de conteúdo baseada em tempo (*time-based content negotiation*) para que um agente de usuário possa solicitar a representação de um recurso original URI (ou URI-R) em ou próximo a um determinado momento, por meio da intermediação de um recurso *TimeGate* (ou URI-G).

De acordo com Gomes (2010, p. 3 - 5), o projeto de Arquivo da *Web* Portuguesa (AWP) foi iniciado em 2007 e faz recolhas exaustivas quatro (4) vezes por ano, incluindo arquivos HTML, imagens, PDF e com o formato ARC, mas aceita conteúdos fornecidos em qualquer formato, convertendo-os para que possam ser indexados. A maneira de preservar conteúdos anteriormente publicados na *Web* é obtendo-os por entidades externas que os tenham. Assim, o AWP tem buscado fortemente obter conteúdos históricos. Apenas é arquivada a informação pública e

não são preenchidos formulários, todas as páginas protegidas por senhas ou outros mecanismos de restrição de acesso não são recolhidas. De acordo com o autor, perto de 90% dos arquivos da *Web* portuguesa são recolhidos ao fim de 7 dias, no entanto, a recolha é contínua para os sítios mais lentos ou com maior número de conteúdos.

Entre 1996 e 2004 existiu pouca colaboração entre iniciativas de arquivo da *Web*. A principal conclusão desses esforços foi notar que esse tipo de arquivamento é um desafio que demanda esforço conjunto, em escala mundial. Em 2004, foi criado o projeto *Archive-Access*, liderado pelo *Internet Archive*, que reuniu e desenvolveu ferramentas de software gratuitas para arquivar a *Web*. O Arquivo da *Web* Portuguesa (AWP) adotou três sistemas pertencentes ao projeto *Archive-Access*: o sistema de recolha *Heritrix*; o sistema de pesquisa *NutchWAX* e o sistema de acesso a conteúdo arquivado *Wayback*. O sistema de recolha foi adotado sem alterações significativas, já os sistemas de pesquisa e acesso precisaram ser modificados para que os objetivos do AWP fossem alcançados, satisfazendo as necessidades da comunidade portuguesa que eram traduzir para português a interface de utilização; melhorar a usabilidade e grafismo das páginas; reestruturação dos índices para permitirem respostas mais rápidas e melhorar a ordenação dos resultados para fornecer resultados mais relevantes (GOMES, 2010, p.4).

No AWP as páginas arquivadas ficam publicamente acessíveis um ano após o arquivamento. O objetivo do período de embargo de acesso é evitar a concorrência com os sítios *Web* que os publicaram. Os autores podem dar indicações para que uma página não seja arquivada por meio do *Robots Exclusion Protocol*. Após esse processo, os autores podem solicitar o bloqueio do seu acesso pelo AWP. Os endereços de todos os arquivos a serem bloqueados são necessários pois os conteúdos de uma página podem ter vários autores, por exemplo, o texto pode pertencer a um escritor, mas a imagem, a um fotógrafo (GOMES, 2010, p. 4 - 6). O AWP permite que qualquer pessoa sugira um sítio para arquivo, mas é necessário validar, utilizando critérios de seleção concisos, as sugestões feitas antes de incluí-las. Os especialistas das áreas de Ciências da Informação poderiam definir diretrizes que permitiriam que um não-especialista identificasse se um determinado sítio deveria ser arquivado.

Segundo Gomes (2010, p. 6) a experiência da AWP contribui para planejar o desenvolvimento futuro do serviço, permitindo (grifo nosso):

- Identificar perfis de utilizadores de um Arquivo da *Web* e estimar a sua prevalência no universo de utilizadores;
- Identificar os casos de uso para cada perfil e estimar a percentagem dos casos de uso entre os perfis;
- Identificar necessidades de informação para cada caso de uso;
- Identificar casos de uso e necessidades de informação que não são satisfeitas pelos motores de busca e arquivos da *Web* atuais.

Por fim, para atender às necessidades de arquivamento da *Web* de instituições menores, várias ferramentas foram lançadas no mercado nos últimos anos. Em 2005, veio o *Archive-it*, um serviço fundamentado em assinatura e baseado nas principais tecnologias de arquivamento da *Web*, desenvolvido para instituições de memória e arquivos estatais para criar coleções não capturadas por arquivamento da *Web*. O *Web Archiving Service*, da *California Digital Library*, usa o *Heritrix*, *Wayback* e *NutchWAX*, sendo usado por muitas instituições, incluindo as Universidades de

Stanford e de Berkeley. O *ArchivetheNet* é oferecido pela *Internet Memory Foundation*, e usado, entre outros, pela *British Library* para criar o arquivo *Web* do governo do Reino Unido (DAVIS, 2014).

3.3.1.2 Iniciativas de arquivamento: exemplos nacionais

Embora no exterior sejam encontradas algumas iniciativas que buscam coletar e armazenar toda a *Web*, até recentemente, segundo Rockembach (2018, p. 20) não havia iniciativas de arquivamento da *Web* identificadas no Brasil, nem mesmo nos demais países da América Latina até aquela data. Em 2019 o Chile ingressou como membro do Consórcio Internacional de Preservação da Internet, a partir da sua Biblioteca Nacional. O Programa Permanente de Preservação e Acesso a Documentos Arquivísticos Digitais do Arquivo Nacional do Brasil (AN Digital), iniciado em 2010, disponibilizou em sua página a Política de Preservação Digital, com versões de 2012 e atualizada em 2016. Ambos os documentos ressaltam a necessidade de preservação dos documentos digitais, contudo, dispõem que a preservação incidirá sobre os tipos documentais “texto estruturado com formatação, imagem matricial, imagem vetorial, áudio, audiovisual, mensagens de correio eletrônico, apresentações (*slides*), planilha e base de dados relacional” (ARQUIVO NACIONAL DO BRASIL, 2016, p.11) e, futuramente, outros tipos, como multimídia e páginas *Web*.

Como dito anteriormente, no Brasil, Rockembach (2019) aponta como iniciativas de arquivamento da internet o Núcleo de Pesquisa em Arquivamento da *Web* e Preservação Digital (NUAWEB) da Universidade Federal do Rio Grande do Sul (UFRGS), criado em 2017. E, em estudo mais recente, Rockembach (2021, p. 6) mostra que diversos conteúdos do domínio .br e da internet brasileira podem passar a ser identificados e capturados para acesso retrospectivo por meio do trabalho do Arquivo da *Web* Brasileira², com a colaboração daquele Núcleo de Pesquisa em Arquivamento da *Web* e Preservação Digital³ e voluntários. Os alvos atingidos são diversos, passando pelas áreas institucional, organizacional, esfera federal governamental e das eleições.

Uma iniciativa que possui relação com a *Web* brasileira é a *Latin American Web Archiving Project*, hospedada em endereço eletrônico da Universidade do Texas, com foco nos documentos governamentais e de expressão política. Outra iniciativa internacional que buscou colaborar com o Brasil foi a do Grupo de Trabalho de Desenvolvimento de Conteúdo (CDG, do inglês, *Content Development Working Group*), vinculado ao Consórcio Internacional de Preservação da Internet, que coletou sites, artigos, notícias, blogs e mídias sociais sobre as Olimpíadas do Rio 2016. Em tese, Dantas (2015 apud ROCKEMBACH, 2018, p. 20) utilizou a *hashtag* #RIO2016WA (no *Twitter*) como forma de marcar e acompanhar informações do arquivamento das Olimpíadas 2016, e conectar pessoas para contribuir com o arquivamento da *Web* do evento, e Dantas constatou não haver coleções de páginas *Web* em instituições brasileiras. Pesquisa no site do *Internet Archive* mostrou algumas coleções que têm relação com conteúdo produzido no Brasil, de forma dispersa, sem uma linha definida ou políticas de seleção e arquivamento estabelecidos. Para Brayner (2019), países como o Brasil ainda não têm uma preocupação em armazenar e

2 Disponível em: www.arquivo.org.br

3 Disponível em: www.ufrgs.br/nuaweb

preservar a *Web* em nível nacional, e deveriam ter políticas de arquivamento da *Web* como objetivo, a fim de resguardar o patrimônio digital brasileiro.

Hoje a *Web* é um ambiente com informações sobre diversos campos do conhecimento, que pode contar com diversos formatos de arquivo. Enquanto o arquivamento digital for aplicado a diferentes contextos e países, mais pesquisas poderão ser efetivadas. Caso não haja responsabilidades atribuídas ao Brasil e às suas Instituições governamentais e de pesquisa, muito do que já foi e vem sendo gerado digitalmente poderá desaparecer definitivamente. Isso compromete as pesquisas, que não serão aproveitadas por futuros pesquisadores de diversas áreas de conhecimento (ROCKEMBACH, 2018, p. 20).

3.3.2 Análise e avaliação de soluções tecnológicas para preservação a longo prazo de documentos oficiais governamentais

Mediante a necessidade de analisar e avaliar as soluções tecnológicas que permeiam a preservação digital, uma vez que as leis diferem de país para país, pretendeu-se analisar a legislação de páginas *web* governamentais, com a finalidade de obter subsídios para fundamentar projetos brasileiros. O levantamento teve foco na legislação nacional, mas abrangeu, a priori, normativas do panorama mundial, destacando os países onde o assunto já está sendo tratado.

3.3.2.1 Legislação: panorama mundial

Para conservar os recursos de informação analógica, no geral, é suficiente hospedá-los em um ambiente apropriado. Algumas vezes pode ser necessário fazer um microfilme ou uma cópia xerográfica do original, mas a cópia é a exceção, e não a regra. A preservação digital, entretanto, já começa com uma cópia. No mínimo, os arquivos precisam ser copiados para um resgate futuro, ou para ter uma guarda compartilhada (múltiplas cópias) e localizada em mais de um local, ou ainda para tirá-los de mídias obsoletas. Preservação digital e acesso estão diretamente ligados a copiar. (MINOW, 2003).

Arquivar a *Web* não é um projeto simples, cada vez mais se percebe que a mídia na qual as informações estão armazenadas se torna ilegível, ou o *hardware*, e o software necessário para ler a obra se torna obsoleto. Há legislação que versa sobre os temas ligados ao arquivamento de páginas *Web*, por exemplo, o ligado ao direito autoral (*copyright*), nela, copiar significa reproduzir, que é um dos direitos do autor. Quase todos os documentos possuem autoria, mas há exceções como no caso de obras criadas pelo governo federal, que são de domínio público, ou então os fatos, as ideias e as obras antigas (MINOW, 2003). Para a autora, existem pelo menos três seções específicas da lei de direitos autorais que podem ser úteis, e ela mostra como a lei americana as trata:

- **Arquivando programas de computador:** de acordo com a lei americana, apesar dos direitos exclusivos do proprietário dos direitos autorais, é permitido que se faça uma cópia para fins de arquivamento de um programa de computador protegido por direitos autorais. A lei permite que se faça uma cópia de um programa (do qual se é o possuidor legal), e até mesmo é permitido adaptá-lo para “rodar” em outra máquina, mas não compartilhar o arquivo com ninguém. A seção se aplica apenas ao próprio programa de

computador, não autoriza a reprodução ou adaptação de documentos criados quando os direitos autorais destes forem de propriedade de outra pessoa;

- **Bibliotecas e arquivos que fazem cópias de preservação:** de acordo como a autora, uma das poucas vantagens incluídas no *Digital Millenium Copyright Act* ("DMCA") foi a cláusula que permite explicitamente às bibliotecas e arquivos fazer até três cópias de uma obra para fins de preservação, os itens podem estar em qualquer formato (texto, imagens, som etc.) e as cópias podem ser digitais, desde que não sejam distribuídas nem colocadas à disposição do público fora das instalações da biblioteca ou arquivo. A lei estipula outros detalhes, como que as bibliotecas e arquivos devem ser abertos ao público ou permitir o acesso a pesquisadores não afiliados; se o trabalho não for publicado, cópias para preservação podem ser feitas para fins de preservação ou segurança, se o trabalho for publicado, cópias de preservação podem ser feitas para substituir um original que está "danificado, deteriorado, perdido ou roubado, ou se o formato existente no qual o trabalho está armazenado tornou-se obsoleto;
- **Preservação de "uso justo" (Fair Use) por indivíduos e bibliotecas:** no cerne da isenção de "uso justo" está a avaliação dos quatro fatores que o constituem: *objetivo* do uso, natureza da obra, *quantidade ou substancialidade* usada e *impacto* no mercado, no original, em inglês, seria: **Purpose of the use**, **Nature of the work**, **A**mount or substantiality used, e **M**arket impact (PNAM).

A maioria das informações encontradas na *Web* são automaticamente protegidas pelo direito do autor (*copyright*) quando criadas. Porém, para (MINOW, 2003), os grupos que desejam preservar as páginas da *Web* geralmente não são os proprietários dos direitos autorais dessas páginas. Pode-se presumir que o proprietário concedeu uma licença implícita para permitir que se copie e exiba uma página da *Web* para uma máquina local. A tentativa mais ambiciosa de preservar a *Web*, o *Internet Archive* e sua *Wayback Machine*, permite recuperar páginas virtuais desatualizadas de vários momentos (INTERNET ARCHIVE, 2021?).

A disposição do *Internet Archive* de respeitar os desejos dos proprietários do *copyright* que desejam limitar e controlar a reprodução de suas obras protegidas, reduz o risco de processos por violação do arquivo, e diminui sua utilidade (arquivo) como um todo, excluindo partes importantes da *Web*. Deve-se ler cuidadosamente todos os acordos de licenciamento a serem assinados antes de fazer cópias para preservação dos recursos licenciados. De acordo com o DMCA, se o recurso digital é protegido por uma tecnologia que controla o acesso ao recurso, não se pode legalmente contornar o mecanismo de controle de acesso, nem mesmo para preservá-lo.

Minow (2003) diz que, nos Estados Unidos, boas práticas de preservação sempre existiram em uma área jurídica "cinzenta". As bibliotecas já faziam três cópias durante a microfilmagem, e muitos programas de rádio foram salvos porque os indivíduos os gravaram diretamente enquanto estavam no ar, algumas vezes até sem a permissão do proprietário dos direitos autorais. A preservação digital reside na necessidade fundamental de copiar informações digitais para preservá-las. Para ela, o dono da informação tem um interesse maior em preservá-la, e a falta de segurança jurídica não deve impedir que se preserve a informação digital. A lei americana autoriza explicitamente algumas ações de preservação (especialmente se os materiais não forem disponibilizados digitalmente para terceiros), e uma forte defesa do *Fair Use* pode ser construída fora da biblioteca ou dos arquivos.

Na Biblioteca do Congresso americano, basicamente existem duas abordagens para lidar com a questão dos direitos autorais: *opt-in* e *opt-out*, de acordo com (DAVIS, 2014). Com a abordagem *opt-in*, a instituição que realiza o arquivamento da *Web* entra em contato com os proprietários de conteúdo apropriados para obter permissão explícita para capturar seu site. A abordagem *opt-out* simplesmente captura sites de interesse, respeitando o protocolo do *robots.txt*, e a retirada. Outras solicitações são tratadas após o fato, caso a caso. Para o autor, a abordagem *opt-in* fica insustentável porque significaria a captura de uma parcela mínima dos sites. O *Internet Archive* (IA) usa uma abordagem *opt-out* e, se solicitado pelo proprietário do conteúdo, o IA removerá um site do acesso público por meio do *Wayback Machine* - mas não, o que é importante, de seu índice central.

A Biblioteca do Congresso Americano encontrou um equilíbrio entre as duas perspectivas. Embora possam buscar permissão explícita para alguns sites, “algum aviso, no mínimo, deve ser fornecido ao proprietário” antes que ele seja rastreado e capturado (LIBRARY OF CONGRESS COLLECTION DEVELOPMENT OFFICE, 2017). Por causa da falta de precedência legal explícita e das abordagens variadas para arquivamento da *Web* em várias instituições e em várias jurisdições, é muito importante lidar com a questão das permissões o mais cedo possível, muitas vezes em consulta a um advogado.

O *Publications Office* do serviço de arquivamento da *Web* da União Europeia (UE) realiza o arquivo para preservar os sítios *Web* da União Europeia. A maior parte do conteúdo arquivado dos sites que estão acessíveis no Arquivo da *Web* da União Europeia (em inglês EU *Web Archive* - EUWA) está sob seu direito autoral, ou de instituições, agências ou organismos da UE. (EU *WEB ARCHIVE*, 2021?). A propriedade e os direitos autorais desses sites continuam sendo responsabilidade dos proprietários (dos sites). Salvo indicação em contrário, o material obtido do EUWA pode ser reproduzido livremente. Esse princípio geral pode estar sujeito a condições que podem ser especificadas em avisos de direitos autorais individuais. Não se aplica a fotografias, vídeos, peças musicais ou outro material sujeito a direitos de propriedade intelectual de terceiros (fora da UE). Nesses casos, a permissão para usar o material deve ser solicitada diretamente aos detentores dos direitos autorais. O *Publications Office* não garante que todo o conteúdo de terceiros seja devidamente marcado.

Aquela foi a orientação geral para a UE, mas outros países tomaram atitudes diferentes. Na Suécia, uma das restrições encontradas na página da iniciativa de arquivamento da *Web* da sua Biblioteca Nacional é que só se pode pesquisar os arquivos da *Web* se estiver pessoalmente na biblioteca, onde há computadores especiais para essa finalidade. De acordo com a lei sueca, tais arquivos só podem ser exibidos dentro da biblioteca. Na França, as restrições encontradas na página da sua Biblioteca Nacional mostram que os arquivos da *Web* são acessíveis a usuários autorizados da biblioteca apenas nas salas de leitura e de pesquisa. Essa restrição é a mesma que se aplica a todas as coleções de depósitos legais (ROCKEMBACH, 2018, p. 15).

3.3.2.2 Legislação: panorama nacional

O Brasil tem cerca de 16 mil leis, mais que o dobro desse número na forma de medidas provisórias (decretos-lei), 220 mil decretos do executivo e mais de 500 mil portarias e instruções normativas. (PORTAL NORMAS.LEG.BR, 2021). Com números tão grandes quanto a extensão territorial do país, percebe-se que uma análise de suas leis não é algo trivial. Para auxiliar na pesquisa e acesso à legislação brasileira, um projeto liderado pelo Senado Federal, o portal Lexml, pretende reunir leis, decretos, acórdãos, súmulas e projetos de leis, entre outros, das

esferas federal, estadual e municipal dos três poderes brasileiros (Executivo, Legislativo e Judiciário). É uma rede de informação legislativa e jurídica que almeja organizar, integrar e dar acesso às informações disponibilizadas nos diversos portais de órgãos do governo na Internet (LIMA; LIMA; KRAUSS, 2022?). É fruto de 20 anos de trabalho, e tem como subproduto o portal “*normas.leg.br*” (NORMAS.LEG.BR., 2022?), uma iniciativa do Congresso Nacional para apresentar a evolução das normas de forma simples, informativa e transparente. O portal exibe a compilação estruturada das normas constitucionais, possibilitando consultar qualquer versão da norma na linha do tempo, selecionar o tipo de anotação que será apresentada no texto/sumário, e visualizar infográficos (LIMA; LIMA; KRAUSS, 2022?). O Portal Lexml permite consultar artigos remotamente, compilando normas (compilação estruturada).

A partir de pesquisas feitas em bases de dados da Câmara dos Deputados, da legislação federal em vigor, e de proposições legislativas relacionadas direta ou indiretamente com o tema de arquivamento e preservação de sítios de instituições públicas na Internet, foi elaborado o trabalho de Santos (2020). Ele é dos mais recentes e específicos sobre o tema, e traz uma visão geral sobre a legislação brasileira, em nível federal.

Da legislação brasileira, o Projeto de Lei 2431/2015 (SANTOS, 2020, p. 4 - 5) tem pontos importantes, com destaques nossos, com relação às definições:

Art. 2º Para efeitos desta lei considera-se:

I - *patrimônio público digital institucional*: todo o conteúdo hospedado em sítios oficiais na rede mundial de computadores, qualquer produção comunicacional em sua formatação e estilo original, incluindo acervo fotográfico, audiovisual, de áudio e de texto, inserida na rede mundial de computadores por órgãos públicos ou entidades que prestam serviços públicos;

II - *sítio oficial*: todo sítio de internet vinculado a órgãos da administração pública direta ou indireta, hospedado sob as extensões “gov.br”; “leg.br”; “jus.br”, para entidades governamentais civis, e “mil.br”, para entidades militares.

Quanto às atribuições:

Art. 3º Ficam os *Chefes dos Poderes Públicos* incumbidos da preservação e manutenção do conteúdo digital institucional em seu formato original disponível na rede mundial de computadores, *garantindo o acesso público e facilitado aos usuários*.

De acordo com Santos (2020, p. 4), outros projetos têm implicações diversas, sendo a maioria relativa ao apagamento de informações que possam comprometer a honra, a imagem das pessoas e suas heranças digitais (grifos nossos):

- PL 3050/2020 altera o Código Civil: “Serão *transmitidos aos herdeiros* todos os conteúdos de qualidade patrimonial de contas ou arquivos digitais de titularidade do autor da herança”.
- PL 3051/2020 “...os provedores de aplicações de internet devem excluir as respectivas contas de usuários brasileiros mortos, imediatamente, *se for requerido por familiares* após a comprovação do óbito”.

- PL 3395/2020. “É vedado às plataformas digitais *remover conteúdos publicados* por seus usuários, *salvo* por força de ordem judicial, à exceção da hipótese prevista no art. 21, em que o provedor procederá à indisponibilização do conteúdo independentemente de ordem judicial.”
- PL 3573/2020, altera o Marco Civil da Internet em: “Com o intuito de assegurar a liberdade de expressão e impedir a censura, o provedor de aplicações de internet *não poderá retirar conteúdo gerado por terceiro, exceto* por ordem judicial ou com a indicação expressa do crime que se está cometendo mediante a divulgação do conteúdo retirado”.

Ainda segundo Santos (2020, p. 4, 5) “pode-se afirmar que não há uma norma que considere os sítios como publicações ou publicações oficiais no contexto do depósito legal de materiais bibliográficos”. A Lei 10.994/2004 (BRASIL, 2004) explicita:

Art. 1º “...regulamenta o depósito legal de publicações, na Biblioteca Nacional, objetivando *assegurar o registro e a guarda da produção intelectual nacional...*”;

Art. 2º Considera:

I - Depósito legal: a exigência estabelecida em lei para *depositar*, em instituições específicas, *um ou mais exemplares, de todas as publicações*, produzidas por qualquer meio ou processo, para distribuição gratuita ou venda;

Art. 3º *Abrange as publicações oficiais* da administração federal, estadual e municipal, dos órgãos e entidades de administração direta e indireta, das fundações criadas, mantidas ou subvencionadas pelo poder público;

Art. 5º O depósito legal *será efetuado pelos impressores, devendo ser efetivado até 30 (trinta) dias* após a publicação da obra, cabendo ao seu editor e ao autor verificar a efetivação desta medida. Não foi identificado o termo, artigo ou inciso para incluir os sítios entre as publicações submetidas ao depósito legal, também fica fora os Diários Oficiais, Diários de Justiça, Diário do Congresso Nacional, Boletins administrativos e similares.

O mesmo autor ainda ressalta que existe legislação que dispõe sobre a publicação de informações de interesse público em sítio oficial das instituições, e isso pode ser interpretado como uma “publicação oficial”. A legislação contemporânea formaliza o uso de sítios *Web* como ferramenta oficial de difusão das informações institucionais, e ele ilustra esse entendimento com a Lei nº 12.527/ 2011, conhecida como a Lei de Acesso à Informação (LAI). Em seu artigo 8º ela mostra que:

É dever dos órgãos e entidades públicas promover, independentemente de requerimentos, a divulgação em local de fácil acesso, no âmbito de suas competências, de informações de interesse coletivo ou geral por eles produzidas ou custodiadas (grifos nossos).

O Decreto 7.724/2012 (BRASIL, 2012 apud Santos, 2020, p. 4, 5), em seu Art. 7º, amplia essa questão (grifos nossos):

§ 3º Deverão ser divulgadas [...] informações sobre:

I -estrutura organizacional, competências, legislação aplicável, principais cargos e seus ocupantes, endereço e telefones das unidades, horários de atendimento ao público;

II - programas, projetos, ações, obras e atividades, com indicação da unidade responsável, principais metas e resultados e, quando existentes, indicadores de resultado e impacto;

III - repasses ou transferências de recursos financeiros;

IV - execução orçamentária e financeira detalhada;

V - licitações realizadas e em andamento, com editais, anexos e resultados, além dos contratos firmados e notas de empenho emitidas;

VI - remuneração e subsídio recebidos por ocupante de cargo, posto, graduação, função e emprego público, incluídos os auxílios, as ajudas de custo, os jetons e outras vantagens pecuniárias, além dos proventos de aposentadoria e das pensões daqueles servidores e empregados públicos que estiverem na ativa, de maneira individualizada, conforme estabelecido em ato do Ministro de Estado da Economia;

VIII - contato da autoridade de monitoramento, designada nos termos do art. 40 da Lei nº 12.527, de 2011, e telefone e correio eletrônico do Serviço de Informações ao Cidadão -SIC [...];

IX - programas financiados pelo Fundo de Amparo ao Trabalhador -FAT.

Art. 64-A. As entidades com *personalidade jurídica de direito privado* constituídas sob a forma de serviço social autônomo, destinatárias de contribuições, divulgação, *independentemente de requerimento*, as informações de interesse coletivo ou geral por elas produzidas ou custodiadas [...] em *local de fácil visualização em sítios oficiais na internet*.

O marco legal da internet, Lei 12.965/2014, no Art. 7º, mostra (BRASIL, 2014) (grifos nossos):

O acesso à internet é essencial ao exercício da cidadania, e ao usuário são assegurados [...] IX -*consentimento expresso* sobre coleta, uso, armazenamento e tratamento de *dados pessoais*, que deverá ocorrer de forma destacada das demais cláusulas contratuais;

X -*exclusão definitiva dos dados pessoais* que tiver fornecido a determinada aplicação de internet, a seu requerimento, *ao término da relação entre as partes*, *ressalvadas* as hipóteses de guarda obrigatória de registros previstas nesta Lei e na que dispõe sobre a proteção de dados pessoais.

Santos (2020, p. 8) conclui dizendo que:

Embora não exista uma legislação específica sobre a proteção, captura e preservação de sítios web de instituições públicas ou privadas, entende-se que é possível [...] defender a necessidade de que esse serviço seja prestado por instituições de memória.

Quanto às responsabilidades, o *Arquivo.pt* (Portugal) atribui aos arquivos o encargo pela preservação do patrimônio digital do país, já a Biblioteca Digital Hispânica (Biblioteca Nacional de Espanha), vê a responsabilidade como sendo da biblioteca, como derivação do conceito do depósito legal.

A lei brasileira traz alguns problemas legais quanto aos direitos do autor. A Lei 10.994/2004, que dispõe sobre o Depósito Legal Brasileira (BRASIL, 2004) exige que os autores depositem na Biblioteca Nacional um ou mais exemplares de todas as suas publicações, produzidas por qualquer meio ou processo (Art. 2º, inciso I), acarretando multa para quem não a cumprir (§ 1º, inciso I). Diz a lei que as despesas do depósito legal são de responsabilidade exclusiva dos (autores) depositantes (Art. 6º). Porém, para Gomes (2010, p. 3), essa abordagem é difícil de operacionalizar porque impõe custos para os autores; as imposições legais têm fronteiras que não existem na Internet; e faltam normas e ferramentas que permitam aos autores depositar as suas publicações de uma forma eficaz e eficiente. Vê-se que o depósito imposto ao autor é uma analogia ao modelo de depósito legal existente para as publicações impressas, em que o autor é legalmente obrigado a entregar cópias das suas publicações para arquivo, sendo um fator que dificulta a formação do lastro da Biblioteca Nacional brasileira.

Quanto aos direitos autorais no Brasil, a representação digital de conteúdos informacionais tem questões éticas, legais e sociais, como a da propriedade intelectual, da proteção da confidencialidade e da privacidade. Pela lei brasileira, talvez das mais restritivas do mundo, observa-se o que aparece em seu artigo 29 (BRASIL, 1998), no que se refere às obras digitais: é vedada sua inclusão em base de dados, o armazenamento em computador, a microfilmagem e as demais formas de arquivamento do gênero, em quaisquer outras modalidades de utilização existentes ou que venham a ser inventadas (BRASIL, 2015). Nesse caso, tem-se uma questão séria e, ainda, não resolvida em âmbito nacional: como operacionalizar o arquivamento da *Web* se a lei brasileira não permite que as informações digitais sejam disponibilizadas em ferramentas de TI que possam permitir-lhes tal resguardo futuro? (BOERES, 2017, p. 114).

Segundo Rockembach (2018, p. 15), o uso e reuso dessas informações é duvidoso no que diz respeito às permissões, que nem sempre estão explícitas nas páginas *Web*, como é o caso da aplicação das licenças *Creative Commons* e seus formatos de concessões. Isso leva a problemas legais e éticos, pois envolve o uso de informações sob direitos autorais e deve levar à reflexão sobre a ética informacional ligada à proteção de dados e privacidade na internet. Uma das alternativas legais e éticas para essa coleta e uso do arquivamento da *Web* vem por meio do *fair use*, ou uso justo, termo usado na legislação americana ligado ao uso de conteúdo sob direitos autorais, como o uso pedagógico e educacional, como para gerar notícias ou pesquisas. Minow (2003) defende que pelo fato de a *Web* ser um recurso importante na produção e transmissão da informação, há o interesse em preservar partes do seu conteúdo. E já que a maior parte dos sites são protegidos por direitos autorais, como proceder para coletar esses dados? Isso poderia ser resolvido com o uso do *fair use*, mostrando a intencionalidade de não-coleta pelos motores de busca ou o controle de permissões de acesso com o uso de um arquivo *robots.txt*, e com a possibilidade da manifesta solicitação de retirar algum conteúdo capturado. No entanto, essa legislação e o direito ao uso justo não se aplicam a todos os países, por isso é preciso ser avaliado individualmente. Percebe-se como o Brasil ainda carece de normas.

De acordo com Rockembach (2018, p. 3) a dificuldade em encontrar estudos no contexto brasileiro ocorre por ser um tema ainda recente de pesquisa, embora importante, pois envolve a memória da *Web* brasileira e o acesso às informações no presente e no futuro. Apenas como exemplo, já que não há um modelo oficial para o arquivamento da *Web*, a escassez de metodologias voltadas ao acompanhamento do processo evolutivo dos dados na internet fez com que Silva (2019, p. 20) propusesse um Modelo de Ciclo de Vida de Dados na *Web* (Data on the

Web Lifecycle Model - DWLM). Ele tem como principais objetivos prover um entendimento comum do processo de publicação e consumo de dados na *Web* e especificar os papéis, fases e atividades envolvidos nesse processo.

3.3.2.3 Um pouco sobre repositórios e política de arquivamento

Os repositórios digitais são uma forma de integração de diferentes usuários, para o avanço científico e a preservação dos documentos. As primeiras iniciativas de repositórios digitais datam da década de 1990, e o primeiro repositório digital implementado foi o *Los Alamos National Laboratory (LANL) Physics Archive*, atualmente denominado *ArXiv*, que foi fundado em 1991, no Novo México (COSTA, 2015, p. 23).

Para os repositórios, o arquivamento da produção intelectual aumentou a visibilidade da instituição e do reconhecimento dos pesquisadores por intermédio de citações feitas por outros autores. A autora mostra que grande parte dos repositórios institucionais adota o auto arquivamento como maneira de depósito. A partir disso há autores que defendem a necessidade do estabelecimento de uma política em que se apresente uma licença do tipo *Non exclusive licence*, na qual devem ser apresentadas permissões de depósito, distribuição e preservação dos trabalhos (COSTA, 2015, p. 16, 32, 33).

Para a autora, a criação de políticas de arquivamento deve considerar alguns requisitos importantes, como os direitos autorais. Parte da criação de uma política de acesso às informações disponíveis na internet deveria incluir a escolha da tecnologia/ferramenta específica para fornecer acesso às páginas *Web* arquivadas. A equipe do Archive-It considera as decisões de nível superior as ligadas à política de acesso, uma vez que em uma pesquisa feita por Bragg e Hanna (2013, p. 12), os entrevistados acessavam arquivos da *Web* usando o *software Wayback*, ferramenta de visualização de código aberto que permite ao público navegar em páginas *Web* arquivadas, da mesma forma que experimentariam uma página da *Web* ao vivo, buscando acesso ao conteúdo.

Já a política instituída para os repositórios deve abarcar a que versa sobre os conteúdos, a preservação digital, a de metadados, a de submissão/arquivamento e a política de acesso. Dessa forma, entende-se que se requer tempo para estudar e planejar o escopo do repositório para sua implantação em uma instituição (COSTA, 2015, p. 30). Analisando as políticas no âmbito do arquivamento, verificou-se que de acordo com o critério *Size* (tamanho), o número de *links* dos artigos presentes na *Web* é maior, provendo mais chances de os autores serem citados (IDEM, p. 63).

Com relação ao tempo para efetivar o depósito legal na Biblioteca Nacional brasileira, grande parte das editoras estabelecem um período de embargo (tempo de tramitação para o depósito legal), pois os autores não podem auto arquivar seus trabalhos em repositórios assim que são aceitos em periódicos científicos. Os repositórios também estabelecem políticas de arquivamento tendo o acesso restrito ou a permissão do auto arquivamento para depósito sem acesso, ou com acesso livre (COSTA, 2015, p. 65).

Ressalta-se que o investimento em capacitação de pessoal especializado em preservação digital é primordial. Esses irão verificar se a instituição possui política de preservação da informação em formato digital; se estão sendo utilizados padrões de formatos de arquivos para a preservação; verificar como a instituição irá digitalizar

documentos para inclusão no seu acervo digital, entre outros tantos questionamentos importantes a serem feitos para a implantação de ações de preservação de informação digital (COSTA, 2015, p. 29; BOERES, 2017).

Em suma, quanto à legislação, o arquivamento e a preservação dos dados digitais estão pautados nas seguintes legislações:

- Constituição Federal Brasileira (art. 5º, XXXIII; art. 23, III; art. 216), que assegura ao cidadão o direito de receber e ter acesso à informação; à proteção dos documentos, às obras e outros bens de valor histórico, artístico e cultural que compõem o patrimônio cultural brasileiro material e imaterial;
- Lei nº 9.610/1998, que dispõe sobre direitos autorais;
- Lei de Acesso à Informação nº 12.527/2011, que regula o acesso às informações, previstas na CF/1988. Estão subordinadas ao regime dessa lei as entidades da administração direta e indireta, de todas as esferas e poderes;
- Decreto nº 8.539/2015, que dispõe sobre o uso do meio eletrônico para a realização do processo administrativo no âmbito dos órgãos e das entidades da administração pública federal direta, autárquica e fundacional. Parágrafo único. O estabelecido no *caput* deverá prever, no mínimo: I - proteção contra a deterioração e a obsolescência de equipamentos e programas; e II - mecanismos para garantir a autenticidade, a integridade e a legibilidade dos documentos eletrônicos ou digitais;
- Lei nº 12.682/2012, que dispõe sobre a elaboração e o arquivamento de documentos em meios eletromagnéticos;
- Lei nº 13.709/2018, dispõe sobre a proteção de dados pessoais.

3.3.2.4 Tecnologia para arquivamento Web

O arquivamento da *Web* envolve coletar dados da internet, armazená-los e preservá-los em algum nível, tornando-os acessíveis aos usuários. Julien Masanès (*apud* DAVIS, 2014), da *Internet Memory Foundation*, identificou três abordagens técnicas para o arquivamento da *Web*: a transacional, a do lado do servidor e a do lado do cliente.

O arquivamento transacional registra o conteúdo apresentado a um usuário em uma data e hora, criando um registro de todas as solicitações e respostas HTTP. Isso requer configuração do servidor, com a cooperação de um *host*. O *SiteStory*, desenvolvido no Laboratório Nacional de Los Alamos, é uma ferramenta transacional de arquivamento da *Web*, ele habilita um servidor quando recebe uma solicitação de um navegador e a envia para um arquivo da *Web* associado. Como resultado, todas as versões de todos os recursos que estão sendo solicitados pelos navegadores também estão disponíveis no arquivo. O arquivo resultante é efetivamente representativo de todo o histórico de um servidor, embora as versões de recursos que nunca são solicitadas por um navegador também nunca sejam arquivadas. Adicionar recursos de arquivamento do *SiteStory* a um *Apache Web Content Server* não afeta seu desempenho de maneira significativa (DAVIS, 2014).

O arquivamento do lado do servidor também requer suporte de *hosts* de conteúdo porque tem relação com a cópia de arquivos diretamente de um servidor. Esse tipo de arquivamento pode ser empregado quando o conteúdo não pode ser facilmente capturado por meio de solicitações HTTP. O *DeepArc* é um exemplo dessa ferramenta, foi desenvolvido pela Biblioteca Nacional da França, e transforma o conteúdo do banco de dados relacional em XML para arquivamento.

Das três opções, a forma mais comum de arquivamento da *Web* é o lado do cliente, ou coleta remota. Aqui, os rastreadores da *Web* usam o protocolo HTTP para coletar conteúdo diretamente de um servidor, buscando todos os links associados a uma URL de “semente” específica. O comportamento do rastreamento é determinado fornecendo ao buscador/rastreador instruções sobre os parâmetros de busca/rastreamento, como sua profundidade. A tecnologia de arquivamento do lado do cliente mais conhecida foi desenvolvida no *Internet Archive* (IA), organização fundada em 1996 por Brewster Kahle para construir uma biblioteca na Internet. Depois desse trabalho as bibliotecas e arquivos nacionais no mundo começaram a ver a importância de preservar esse recurso global (*Web*). Muitas bibliotecas nacionais passaram a estabelecer iniciativas de arquivamento da internet. O IA desenvolveu o *Wayback Machine* (IDEM, 2014).

O IA também desenvolveu a ferramenta *Heritrix*, um aplicativo *Java* que rastreia a *Web* e armazena conteúdo especificamente para fins de arquivamento da *Web* (BRAGG; HANNA, 2013). Esse conteúdo é armazenado nos formatos ARC ou WARC, que especificam um método para combinar vários recursos digitais em um arquivo agregado com informações relacionadas. Isso significa que além do texto HTML, imagens, *JavaScript*, arquivos de mídia etc., outros são combinados em um único arquivo W/ARC, que contém metadados específicos de rastreamento para indicar o momento da coleta, o endereço IP da máquina (coleta), o tipo de mídia da Internet (tipo MIME), o código de resposta para a transação (coleta) e o objetivo da coleta. Os arquivos W/ARC podem ser reproduzidos inserindo o URL da página original em uma instância do *Wayback Machine*, um aplicativo *Java Servlet* que recupera, exibe e indexa o conteúdo arquivado da Internet. O rastreador da Web *Heritrix* e o *Wayback Machine* são de código aberto, o WARC é um padrão ISO (28500:2009), e os arquivos W/ARC podem ser indexados para pesquisa de texto completa, usando ferramentas de código aberto como *NutchWAX* (desenvolvido pelo *Internet Archive* e *Nordic Web Archive*), *Solr* ou *Elasticsearch*, todos baseados em *Lucene* (DAVIS, 2014).

A *Web* cresceu em tamanho e em complexidade, e o trabalho desse autor mostra os tipos de arquivo do *Archive-it*, apontando quantas URLs foram capturadas para cada tipo de arquivo especificado. Os sites estão se tornando cada vez mais dependentes de tecnologias como *JavaScript*, XML, JSON e AJAX, o que pode ser problemático para um rastreador como o *Heritrix*. Antes do AJAX e funcionalidades semelhantes não era possível gerar conteúdo XML dinamicamente para exibição em uma única página, e os usuários eram levados para outra página por uma tag âncora (< a>), contendo uma URL com parâmetros de consulta de banco de dados específicos. Essa âncora permitiria que o rastreador *Heritrix* executasse a consulta de banco de dados específica via HTTP, obtendo o documento associado e o capturando. Também há que se considerar a preservação e o acesso digital. Os arquivos WARC capturados pelo *Archive-it* são armazenados nos Estados Unidos, e para outros países isso é problemático porque as disposições de notificação e remoção da Lei de Direitos Autorais do Milênio Digital Americano (DCMA) diferem de outras leis. Isso pode levar à remoção de conteúdo, pelo *Internet Archive*, o que poderia não ter acontecido se uma solicitação de remoção semelhante tivesse sido emitida pelo outro país. E mais, ferramentas como o *Archive-it* são usadas para capturar sites nos quais os proprietários podem ver que o

conteúdo foi removido, e o argumento convincente pode ser feito para manter o acesso público disponível ao conteúdo capturado, e ter esse conteúdo localizado nos EUA pode ser problemático.

A preservação digital é outra questão importante, a captura de sites no formato WARC para reprodução e pesquisa de texto completo é apenas uma parte do que é necessário para tal preservação. Segundo Davis (2014), os arquivos WARC com backup do *Internet Archive* são suscetíveis à corrupção, assim, o IA pode ser visto como uma amostra estatística de toda a *Web*. Alguma perda de dados pode ser esperada, e ela introduz ruído em uma amostra grande, o que pode ser problemático em uma instituição. O *Archivematica* é um sistema de preservação digital de código aberto projetado para manter o acesso de longo prazo, baseado em padrões de coleções de objetos digitais. O sistema pode ingerir arquivos WARC, que são executados por processos baseados no *Open Archival Information System Reference Model* (ISO-OAIS), que permite o planejamento de preservação futura com base na identificação do formato e outras informações. O *Archivematica* cria Pacotes de Informações de Arquivo (AIPs) de WARCs usando *Bagit*, daí os arquivos ‘empacotados’ podem ser enviados para o COPPUL-PLN, onde uma única cópia, se danificada, pode ser reparada por outras cópias na rede. Esse trabalho potencial de desenvolvimento de integração *Archivematica/LOCKSS* não é trivial, especialmente a criação de um manifesto para LOCKSS para coletar AIPs do *Archivematica*, mas é importante ver que questões jurisdicionais e considerações de preservação digital podem exigir que o gerenciamento local de arquivos WARC garanta o acesso contínuo e a preservação a longo prazo.

3.3.2.5 Modelos de ciclo de vida de dados na *Web*

Dentro da ideia de tecnologias para arquivamento de conteúdo na internet, a literatura aponta para os modelos de ciclo de vida, que geralmente são propostos relacionados a domínios mais específicos. Os poucos modelos conhecidos, que versam sobre domínios genéricos, praticamente não contemplam as fases no domínio de Dados na *Web*, ou não tratam das fases utilizadas nesse domínio. O chamado “Modelo de Ciclo de Vida de Arquivamento da *Web*” visa criar um conjunto de práticas recomendadas para estabelecer e manter um programa de arquivamento da *Web*. Tal modelo é uma tentativa de unir as vertentes tecnológicas e programáticas do arquivamento da *Web* em uma estrutura relevante para organizações que pretendem arquivar a internet, independentemente do tamanho da organização, orçamento ou métodos técnicos de arquivamento da *Web* (BRAGG; HANNA, 2013, p. 27 e 28).

Silva (2019, p. 48) propõe um Modelo de Ciclo de Vida de Dados na *Web* (*Data on the Web Lifecycle Model* - DWLM) que descreve essas fases desde a concepção até a retirada de acesso ao grupo dos dados e aos papéis, atividades, entradas e saídas de cada fase. O DWLM tem o objetivo de informar sobre as etapas por que um conjunto de dados passa ao longo de sua vida na *Web* e visa a garantir que os grupos de dados publicados atendam a alguns requisitos que permitam seu processamento, por humanos e máquinas. Para isso, o modelo incorpora as Boas Práticas para Dados na *Web* (DWBP) propostas pelo W3C.

A partir de informações dos parceiros a equipe do *Archive-It*, concluiu-se que o chamado “círculo externo do modelo” (Modelo de Ciclo de Vida de Arquivamento da *Web*) se relaciona com as questões mais amplas sobre a criação e definição de um programa institucional de arquivamento da *Web*. Essa é a parte que envolve os metadados e a descrição desse ciclo de arquivamento, como a política, e se sobrepõe a outras etapas. Foi tomada

a decisão de apresentar metadados e a descrição como uma faixa extensiva do modelo, em vez de sua própria parte escondida do processo (BRAGG; HANNA, 2013, p. 20).

3.3.3 Modelo de preservação de sites governamentais

O objetivo maior da preservação digital é dar acesso a conteúdos digitais que guardem suas características de autenticidade, confiabilidade e integridade ao longo do tempo. Falar de iniciativas de preservação e arquivamento envolve se remeter ao modelo padronizado por norma internacional (ISO 14721:2012), ou norma ABNT 15472:2021 (Sistema Aberto de Arquivamento de Informações - SAAI), ou *Open Archival Information System* (OAIS), mais conhecido como modelo de referência *Open Archival Information System* (OAIS). Ele foi criado pela *Consultative Committee for Space Data Systems* (CCSDS, 2012) e especifica os principais critérios para orientar as iniciativas de preservação digital. É um esquema conceitual que define e modela o que se precisa para armazenar, preservar e dar acesso a informações digitais. Conta com um vocabulário padrão (metadados), define um modelo de referência que orienta e organiza os sistemas para a preservação de objetos digitais, em longo prazo (ROCKENBACH, 2018, p. 11).

O OAIS é uma das iniciativas que orienta a manutenção do acesso aos objetos digitais para sua preservação. Promovido pela International Organization for Standardization (ISO) e publicado pelo *Consultive Committee for Space Data Systems* (CCSDS), é uma iniciativa que utiliza um vocabulário padrão (metadados). Ele define um modelo de referência que orienta e organiza os sistemas para a preservação de objetos digitais e é adotado quando é necessária uma preservação em longo prazo (CAMPOS, 2007). O modelo OAIS se expressa na Figura 1:

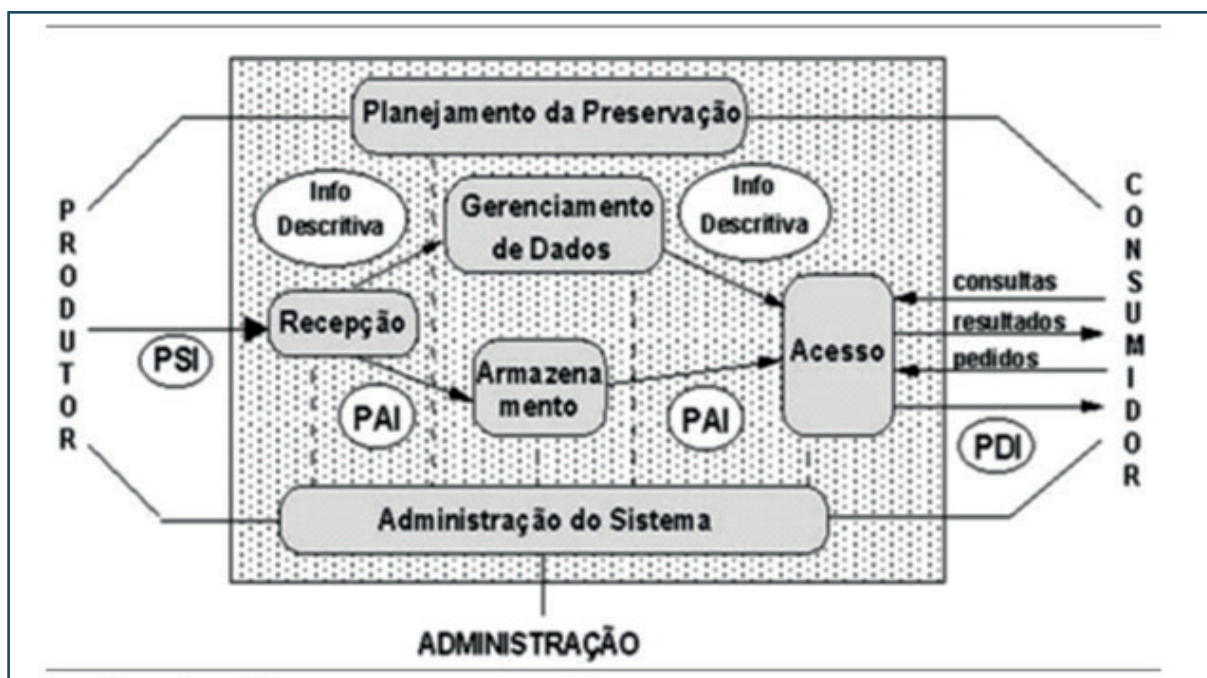


Figura 1 - Modelo Funcional OAIS

Fonte: SOUZA; OLIVEIRA; D'AVILA; CHAVES, 2012, p. 68.

A preservação digital reúne mecanismos de armazenamento, gerenciamento de objetos digitais, estratégias metodológicas e parâmetros de arquivamento. O modelo OAIS objetiva estabelecer um sistema de arquivamento de

informações por meio de pessoas de dada organização que se responsabilizam por preservar as informações e disponibilizá-las para sua comunidade. Nele, cada entidade é responsável por algum dos processos no ambiente OAIS, a entidade recepção recebe os Pacotes de Submissão de Informação (PSIs - para criar ou atualizar PAIs e/ou suas informações descritivas associadas) que chegam dos produtores e no tratamento do conteúdo que será armazenado. Assim, ele gera Pacotes de Arquivamento de Informação (PAIs - informações do conteúdo a ser preservado e as descritivas de preservação), e transfere os novos PAIs, indo as informações descritivas associadas para a entidade armazenamento e para a entidade gerenciamento de dados, respectivamente (SOUZA; OLIVEIRA; D'AVILA; CHAVES, 2012, p. 66-68).

Uma tradução para o português do trabalho de Zierau; Schultz (2013) mostra um modelo “estendido” do OAIS, é o Modelo Outer OAIS-Inner OAIS, ou Modelo OO-IO, traduzido como OAIS Externo – OAIS Interno, para preservação digital distribuída. Desenvolvido por um projeto de repositório de bits dinamarquês para resolver uma necessidade do OAIS, que evoluiu para o modelo OO-IO, pelo projeto Creating a Framework for Applying OAIS to Distributed Digital Preservation. O modelo OO-IO descreve sistemas de preservação digital distribuída de forma a estar em conformidade com o Modelo de Referência OAIS. Ele (OO-IO) suporta a especificação e auditoria de interações colaborativas entre múltiplas implementações de repositórios OAIS. O modelo OO-IO baseia-se na seção 6 do Modelo de Referência OAIS e especifica uma abordagem para usar o modelo OAIS para a interoperabilidade dos arquivos, na forma de preservação digital distribuída, a qual a seção 6 não fornece. Segundo Zierau; Schultz (2013, p. 6):

A ideia por trás do modelo OO-IO é que cada entidade funcional de um repositório OAIS possa ser descrita como um repositório OAIS completo (um OAIS Interno), que ajuda uma organização (um OAIS Externo) a abordar essa entidade funcional. Cada OAIS Interno é tratado como um repositório OAIS completo, embora seja dedicado a administrar uma única entidade funcional no OAIS Externo. Usar o modelo OO-IO pode envolver apenas uma das entidades funcionais do OAIS Externo, o Arquivamento por exemplo, como um repositório de *bits* separado, e gerenciado por uma organização externa.

O modelo OO-IO pode ser representado na Figura 2:

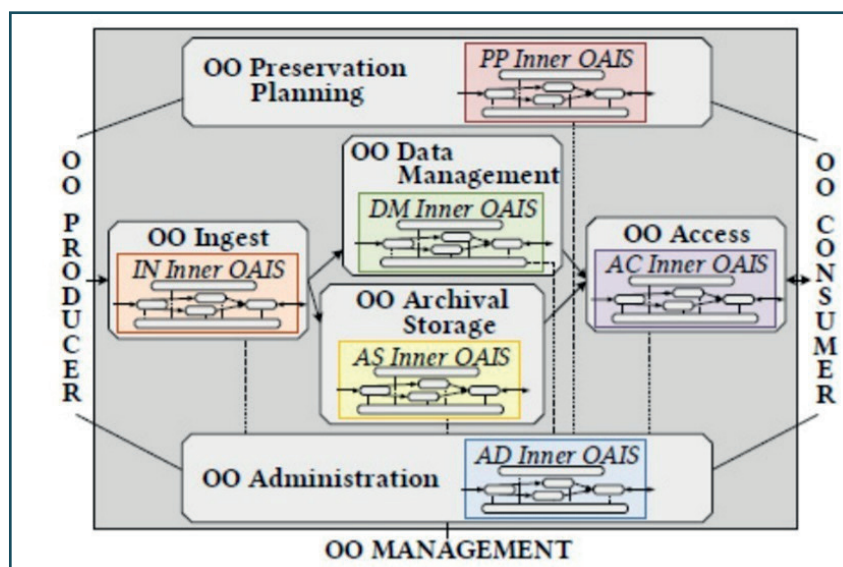


Figura 2 - Modelo OO-IO

Fonte: Zierau; Schultz (2013, p. 6).

3.3.3.1 Arquivamento da *Web* e o modelo OAIS

O trabalho de Rockenbach (2018, p. 11) mostra as etapas para um arquivamento da *Web* que envolvem recolher a informação, coletando o arquivo, armazenando, extraindo os endereços a partir dos *hiperlinks*, e inserindo os novos endereços para a recolha. Os outros dois passos são indexar e disponibilizar serviços. Ele ressalta a necessidade de preservação do conteúdo digital como essencial, dada a dinamicidade e não acessibilidade da *Web* pregressa.

A coleta dos dados da *Web* pode ser feita pelo proprietário da informação, quando ele a envia para o arquivamento, ou quando a coleta é ativa, feita pela instituição responsável pelo arquivamento. No caso dos repositórios, quanto aos fluxos de submissão e depósito, estes podem variar em desabono, tanto da política estabelecida, quanto do software escolhido. A submissão de um documento no repositório requer o preenchimento de metadados e o carregamento do arquivo. Alguns dos componentes essenciais para a submissão são (LEITE, 2009, p. 69):

- Captura de metadados que descrevem a conteúdo dos documentos;
- Carregamento de arquivos eletrônicos;
- Atribuição de licença obrigatória por parte do autor quanto à concordância com os termos de licença do conteúdo que se submete;
- Política de auto arquivamento que determina por quem os arquivos serão arquivados (pelo próprio autor ou por um mediador);
- Mandato de auto arquivamento ou depósito obrigatório.

Nesse contexto, as instituições devem disponibilizar formação e treinamento aos pesquisadores para que eles estejam familiarizados com o sistema (COSTA, 2015, p. 34). Para Rockenbach (2021, p. 7), o modelo de arquivamento da *Web* (Figura 3) pode ser assim representado:

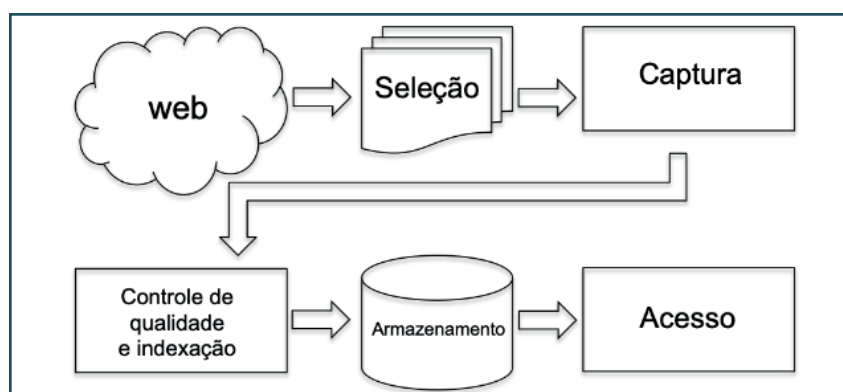


Figura 3 - Modelo de Arquivamento da *Web*

Fonte: Rockenbach (2021, p. 7).

Já para Flores (2019, *slide 13*), a partir do modelo OAIS, traduzido no Brasil como a NBR 15.472 e contemplado na Resolução nº 43 do Conarq, o arquivamento de documentos de arquivo (Figura 4) ficaria assim representado:

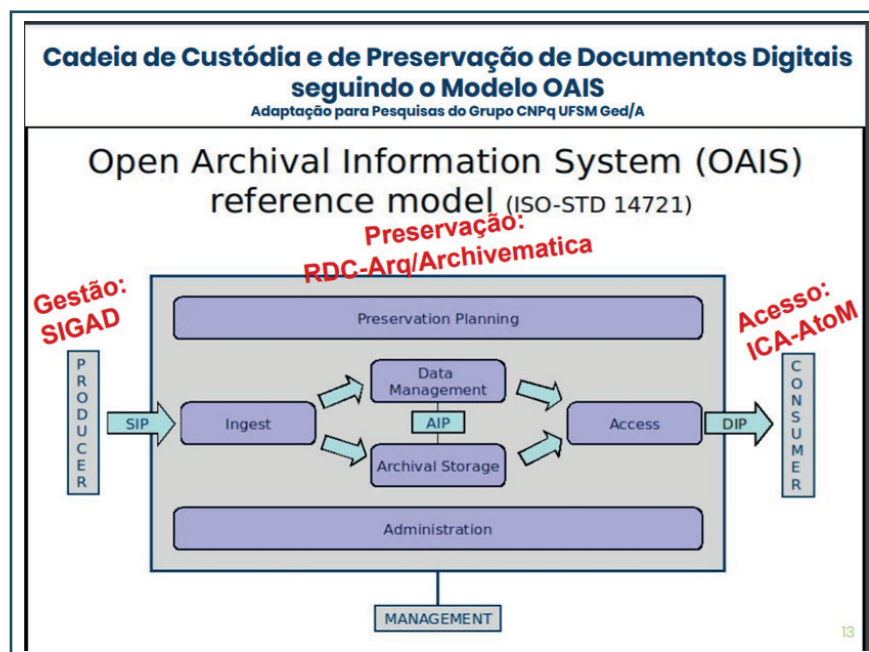


Figura 4 - Arquivamento de Documentos: cadeia de custódia e de preservação de documentos digitais

Fonte: Flores (2019, *slide 13*).

E, no caso de um repositório, a representação ficaria conforme a Figura 5 (FLORES, 2019, *slide 28*):

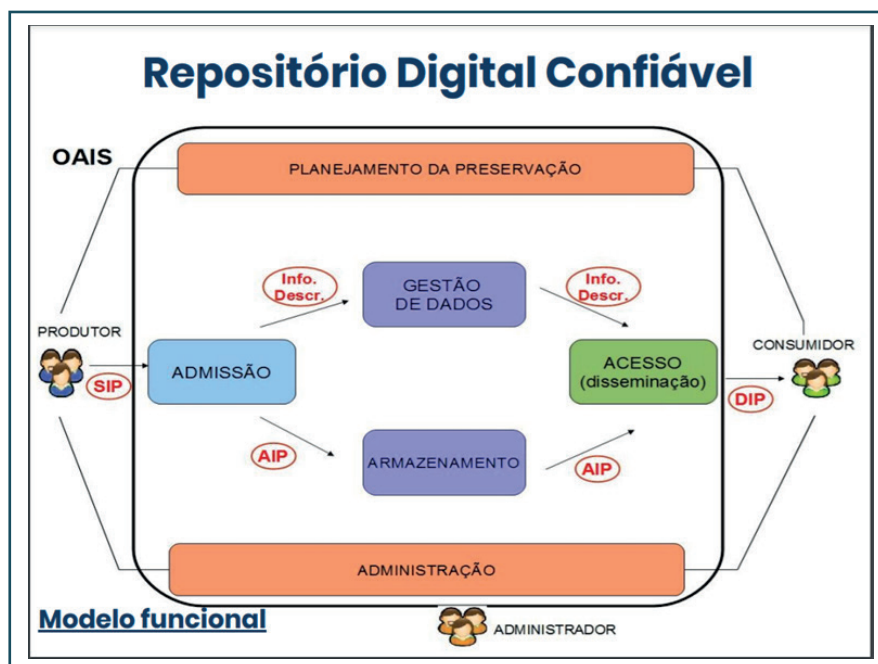


Figura 5 - Arquivamento de Documentos: Repositório Digital Confiável

Fonte: Flores (2019, *slide 28*).

O *software Archivematica*, muito utilizado para efetuar a preservação digital, utiliza o modelo OAIS, é aberto, compatível com vários formatos, de customização flexível e pode ingerir arquivos WARC, que são executados por uma série de processos baseados no Open Archival Information System Reference Model (ISO-OAIS), que permite um planejamento futuro de preservação com base na identificação de formato e outras informações (DAVIS, 2014).

3.3.3.2 Descarte de conteúdos Web

Segundo (Werf; Werf 2014, p. 17), da mesma forma que alguns autores apoiam o “ter o texto seguro”, esses autores apoiam o “ter o texto limpo”, e nenhuma das duas situações vai acontecer rapidamente, é necessário trabalhar para isso. As trias hereditárias, uma estrutura de colaboração entre a indústria da informação, as autoridades públicas e as instituições de patrimônio cultural, podem ajudar a fazer isso acontecer. O descarte pode contribuir para tornar o ciberespaço limpo, lembrando que até as nuvens de armazenamento têm espaço limitado, e o comportamento do consumidor pode ser influenciado com campanhas, inclusive com as que incentivem limpar o lixo digital, contanto que eles sejam incentivados corretamente.

Assim como os bens comuns digitais, as comunidades de patrimônio cultural ou acadêmico e as instituições de memória podem atuar junto com as autoridades públicas, em nível internacional, pois estas últimas têm um papel a desempenhar por meio da colaboração entre a indústria da informação, as autoridades públicas e as instituições de patrimônio cultural (WERF; WERF 2014, p. 17).

Para Heikkinen, *et al.* (2021, p. 2, 6, 7), da Biblioteca Nacional da Finlândia, uma vez que a publicação digital muda constantemente e se torna cada vez mais internacional, no que tange ao desenvolvimento do acervo (seleção e descarte) a cooperação é necessária, tanto na seleção dos materiais a serem preservados, quanto no desenvolvimento e implementação de tecnologias de coleta e preservação. A cooperação é de grande importância, quer seja a internacional, com organismos como o IIPC (International Internet Preservation Consortium) e outras organizações que arquivam materiais on-line, que continuarão a ganhar importância, como a cooperação com a indústria editorial e com outras organizações de arquivamento da *Web*, nacionais e internacionais. É uma necessidade absoluta para coletar e preservar publicações online, além do desenvolvimento constante das competências da equipe (BOERES, 2017). E, por fim, é necessário mostrar que a cooperação com pesquisadores e outros especialistas está se tornando cada vez mais importante na coleta de materiais on-line, tanto em termos de levantamento do material a ser coletado, quanto na promoção do uso de pesquisa desses materiais.

Nas pesquisas em bases de dados especializadas é mais fácil encontrar conteúdo que versa sobre a seleção de material para formar os acervos, do que sobre seu descarte. Talvez porque esta última ação dependa do tipo de centro de informação, se, biblioteca, arquivo, museu ou outro, que têm objetivos e finalidades diferentes, como também pelo motivo que o descarte pode envolver insatisfação dos usuários e discordância, principalmente por parte da direção das instituições, quanto aos métodos utilizados. É importante lembrar que em algum momento o descarte deverá acontecer, a partir dos critérios estabelecidos internamente e da política institucional, que apontará as condições para isso.

3.3.3.3 Iniciativas e estratégias de arquivamento e preservação digital

De acordo com a Rede Cariniana, preservação significa a conservação a longo prazo do conteúdo intelectual e do aspecto do recurso, mas também a garantia do cumprimento das regras de acesso estabelecidas para esse documento. Um programa de preservação deve garantir o acesso ao recurso para sempre, porém, se o conteúdo estiver protegido por direitos de autor e restrições de outra natureza, isso deve ser assegurado. A visão australiana de tomar os recursos como contínuos mostra a importância de preservá-los desde a criação. Dessa maneira, a preservação seria algo intrínseco e presente em todo o processo de gestão desse recurso. No campo das revistas científicas, duas entidades internacionais avançaram com propostas de preservação digital, o Portico e o Lockss, que são serviços pagos (INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA 2022?).

O *Archive-It*, fundação americana sem fins lucrativos, é um serviço de preservação digital, arquivamento da *Web*, e um dos membros fundadores do Consórcio Internacional de Preservação da Internet (IIPC). É ligado à iniciativa *Internet Archive* que está arquivando a *Web* desde 1996. Em 2002, o *Internet Archive* lançou o *Heritrix*, um rastreador de código aberto, ferramenta de software que captura o conteúdo da *World Wide Web* (WWW), que se tornou a tecnologia de rastreamento de arquivos da *Web* mais popular e difundida atualmente (MELO, 2020, p. 37). Os resultados dos acompanhamentos são armazenados em um arquivo *WebArchive* (WARC), um formato desenvolvido pelo IIPC, que em 2009 foi adotado como extensão padrão para arquivos *Web*, definido na ISO 2850017. No início de 2006, o *Internet Archive* lançou o serviço de arquivamento por assinatura para ajudar organizações parceiras a coletar, construir e gerenciar coleções nascidas digitais. Em 2009, o arquivo *WebArchive* (WARC), saída de arquivo do rastreador *Heritrix*, foi adotado como um padrão ISO para arquivamento da *Web*, mostrando o domínio de programas de arquivamento ativos e a importância do próprio rastreador da *Web*. (BRAGG; HANNA, 2013, p. 1). Em 2013, o *Internet Archive* desenvolveu um *White Paper* onde mostraram-se as diversas dimensões envolvidas (ROCKEMBACH, 2018, p. 13).

Em uma breve síntese dos fatos, o *Internet Archive* definiu o formato ARC para salvaguardar os arquivos de websites (em 2009), a *Internet Organization for Standardization* (ISO) estabeleceu o formato WARC como padrão oficial para arquivos *Web*. E veio a ferramenta *Wayback Machine*, que permite acessar determinado sítio como ele era em certa época (*snapshot*). Para isso, é preciso inserir o domínio a ser pesquisado e, com os resultados da pesquisa, selecionar alguma data específica, a partir dos dados capturados. O *Wayback Machine* coleta por domínios de primeiro nível, nacionais e regionais, seletivo, por eventos, temático, coleta em massa, *survey*, sites que serão encerrados e sites específicos. A coleta pode ocorrer de forma ativa, pelo usuário, na página da ferramenta (*Wayback Machine*), é possível inserir o endereço eletrônico que se pretende capturar e salvar a página no banco de dados da *Internet Archive*, podendo fazer uma citação confiável de tal referência futuramente. Porém, essa ferramenta só está acessível para sítios que permitam o uso de rastreadores *Web* (*crawlers*) para a indexação desses sites, que é a mesma tecnologia utilizada por motores de busca. São arquivados pela *Internet Archive* não apenas sítios da *Web*, mas livros, textos, filmes, softwares, música e imagens, assim como coleções cujo total de itens disponíveis aparece em cada coleção (ROCKEMBACH, 2018, p. 17).

Em 2004, foi criado o projeto *Archive-Access*, liderado pelo *Internet Archive*, que reuniu e desenvolveu ferramentas gratuitas de software para arquivar a *Web* (INTERNET ARCHIVE, 2007). O Arquivo da *Web* Portuguesa (AWP) adotou três sistemas pertencentes ao *Archive-Access*: o sistema de recolha *Heritrix*, o sistema de pesquisa *NutchWAX* e o

sistema de acesso a conteúdo arquivado *Wayback*. O sistema de recolha foi adaptado sem alterações significativas, já os sistemas de pesquisa e acesso sofreram modificações para que os objetivos do AWP fossem alcançados, satisfazendo as necessidades da comunidade. (GOMES, 2010, p.4).

O modelo de arquivamento da *Web* começa na seleção do que servirá para as configurações de rastreadores, captura, controle de qualidade, indexação dos conteúdos capturados, armazenamento em padrão ISO 28500 e acesso por meio da reprodução dos arquivos preservados. Inconsistências na qualidade dos conteúdos podem indicar diminuição no potencial de arquivar, dos websites, a partir da acessibilidade, conformidade dos padrões, coesão e metadados, com a possibilidade do uso da ferramenta *Archive Ready* (MELO, ROCKEMBACH, 2020; ROCKEMBACH, 2021). Segundo os autores, a ferramenta mede a “arquivabilidade” e “preservabilidade” dos sítios *Web*, fazendo uma análise na estrutura desses sítios avaliados e seus conteúdos, levando em consideração a acessibilidade, os padrões em conformidade, a coesão e os metadados. Ao avaliar o website o *Archive Ready* gera uma pontuação que demonstra o quão passível de preservação ele é, e gera um relatório que permite visualizar as conformidades (ou não) e as sugestões para aumentar essa “preservabilidade”.

A equipe do *Archive-It* desenvolveu o *Web Archiving Life Cycle Model* (WALCM), modelo baseado nas experiências da equipe e nas lições aprendidas com instituições aliadas, incluindo estudos de caso. O WALCM busca representar fluxos de trabalho comuns e criar um modelo mensurável para as organizações examinarem a fim de criar ou melhorar seus programas de arquivamento na *Web*. (BRAGG; HANNA, 2013, p. 1 e 2). Ele usa *feedbacks* e lições aprendidas em suas parcerias com organizações que arquivam a *Web*, esse retorno vem do uso do serviço. O modelo é uma tentativa de instigar as diferentes fases que uma instituição experimenta à medida que desenvolve e gerencia um programa de arquivamento na *Web*. Embora o modelo seja dividido em etapas individuais, estas estão relacionadas, com uma quantidade significativa de sobreposição. A forma do modelo é circular para sugerir a natureza repetitiva das etapas do ciclo de vida (BRAGG; HANNA, 2013, p. 3). Conforme os usuários passam por cada etapa, eles eventualmente se encontram de volta no início, ou repetindo etapas, dependendo das tarefas. O modelo inclui círculos dentro de círculos para sugerir ciclos repetitivos em um processo maior. O modelo representa os metadados e a descrição, e o *Archive-It* escolheu incorporar metadados como uma banda em vez de um segmento para enfatizar que criar, importar e exportar metadados são processos contínuos, que ocorrem junto com outras atividades no ciclo de vida (BRAGG; HANNA, 2013, p. 3).

O acesso aos conteúdos, segundo Melo (2020, p. 52), poderá ser fornecido pelos próprios sites das iniciativas; já o acesso de longo prazo dependerá de iniciativas capazes de preservar o conteúdo que foi coletado, trazendo-os para o domínio da preservação digital. Os dados reunidos para a preparação do modelo sugerem que a preservação ainda é uma questão em evolução para as instituições que arquivam a *Web*, pois acompanha a evolução da natureza da preservação digital e o desenvolvimento de repositórios digitais como um todo (BRAGG; HANNA, 2013, p. 17). Para a equipe do *Archive-It*, seus parceiros tendem a adotar diferentes estratégias de preservação: algumas utilizam o *Internet Archive* para armazenamento e preservação de seus arquivos WARC e metadados associados; outros recebem uma cópia ou baixam seus arquivos WARC diretamente dos servidores do *Internet Archive* para o disco rígido; algumas instituições aliadas estão trabalhando para incorporar arquivos WARC em nos repositórios digitais locais, embora esses projetos ainda estejam em fase inicial (BRAGG; HANNA, 2013, p. 17).

Uma última fase no fluxo de arquivamento dos sites é permitir que o conteúdo esteja disponível para os interessados. Por meio de iniciativas como a do software *Open Wayback* é possível reproduzir os arquivos WARC armazenados, onde os responsáveis pela curadoria conseguirão validar se a estrutura e o conteúdo do *website* foram mantidos. Além da ferramenta WARC, ALAM, *et al.* (2019, p. 1, tradução nossa) sugerem o Protocolo Memento, financiado pelo Programa Nacional de Infraestrutura e Preservação de Informação Digital (NDIIPP), para negociação de conteúdo *datetime* na reprodução. Ele especifica a negociação de conteúdo baseada em tempo (*time-based content negotiation*, no original) para que um agente de usuário possa solicitar a representação de um recurso original URI (ou URI-R), em ou próximo a um determinado momento, por meio da intermediação de um recurso *TimeGate* (ou URI-G).

3.3.4 Modelos de formatos adequados para preservação de dados e documentos digitais governamentais

Os *websites* governamentais têm um papel importante na disseminação de informações, mas seu conteúdo pode mudar diariamente, ou ser excluído permanentemente. Se não arquivadas ou preservadas, podem levar a perda de informações valiosas para pesquisas e até prestação de contas das ações do próprio governo.

Na última campanha eleitoral brasileira houve grande utilização do *Facebook* para veicular propagandas, utilizando-o com tal intensidade que chegou até mesmo a ultrapassar o tempo previsto para os candidatos na televisão. A partir disso Ferreira (2019) estudou a preservação digital dos vídeos produzidos pelos candidatos, e observou que no Brasil não existem políticas de preservação digital que possibilitem buscar e recuperar os conteúdos publicados apenas on-line. O ideal, para ela, seria uma iniciativa independente, que recebesse um auxílio governamental, e uma legislação que desse respaldo para as universidades conseguirem financiamento para esses projetos. A realização do arquivamento da *Web* depende de recursos para custear as máquinas, os softwares e os funcionários envolvidos no trabalho, lembrando que os arquivos coletados são públicos e disponibilizados gratuitamente (o governo possui a base de dados do projeto Arquivamento das Eleições Brasileiras - Aweb).

Analisando um outro contexto, em dissertação, Melo (2020) avaliou 22 *websites* de ministérios brasileiros e um (1) do governo central, o Portal Único, que pretende reunir todos os sites do Poder Executivo do Governo Federal. Para isso ele utilizou algumas das tecnologias explanadas anteriormente neste trabalho, como o rastreador automatizado de páginas *Heritrix*, para reconstruir os sites da internet arquivados, com o uso do *software* WABAC, e comparou os recursos disponibilizados nas versões ao vivo e arquivadas dos *websites* selecionados. A pesquisa foi amparada nas abordagens de Bragg; Hanna (2013) em seu modelo *The Web Archiving Life Cycle Model*:

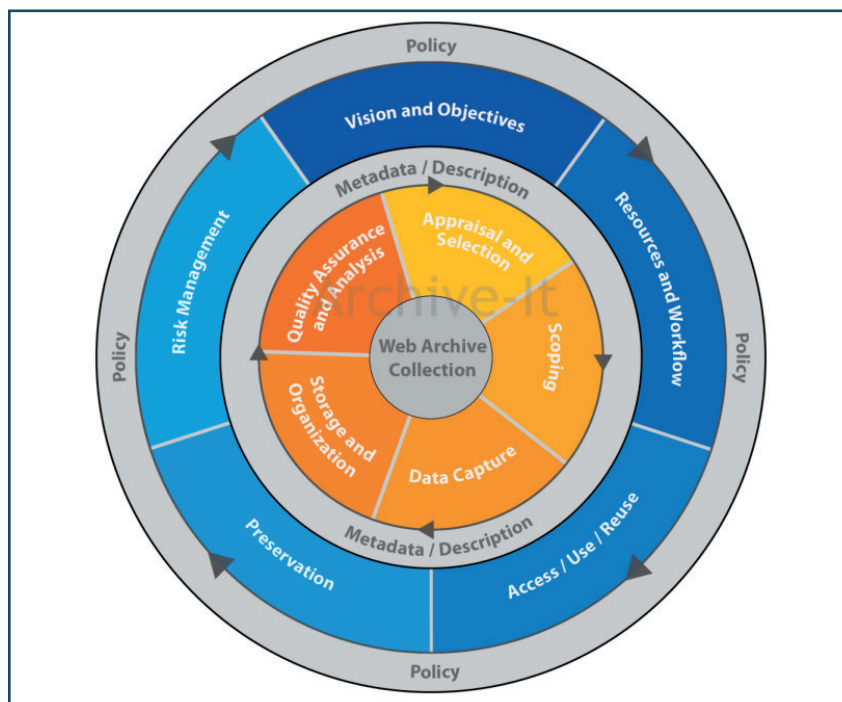


Figura 6 - Web archiving life cycle model

Fonte: Bragg; Hanna (2013, p. 3).

Também se baseou em Khan e Rahman (2019), no trabalho *A Systematic Approach Towards Web Preservation*. Melo (2020, p. 42) listou exemplos de iniciativas internacionais de arquivamento de informações governamentais:

- nos Estados Unidos, a partir de 2008, foi criado um arquivo *web* colaborativo, o *End of Term Web Archive: U.S. Government Websites* (EOT), com coleções de *websites* do governo federal (.gov, .mil) nas áreas do legislativo, executivo e judicial daquele governo;
- naquele mesmo país outras iniciativas governamentais são sobre mudanças climáticas, prevendo o risco de dados do “.gov”, e são a EPA (*Environmental Protection Agency*) e a NOAA (*National Oceanic and Atmospheric Administration*);
- os objetivos do projeto de Sarawak, na Malásia, são: preservar evidências dos conteúdos da *web* publicados pelos departamentos e agências da administração pública; contribuir para facilitar o acesso e a disponibilização de informações à pesquisa; e cumprir com a legislação da Biblioteca Estadual de Sarawak, referentes ao *depósito legal*.

No Brasil, o autor mostra que no censo publicado em 2018, do Instituto Brasileiro de Geografia e Estatística (IBGE), no ano de 2016 a grande produção de informações digitais mostrou que 116 milhões de pessoas estavam conectadas à Internet, equivalente a 64,7% da população do país acima de 10 anos de idade. O Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br) divulgou que em 2018, 67% dos domicílios brasileiros possuíam Internet, e a porcentagem de usuários da rede havia aumentado para 76% da população brasileira, e que 100% dos órgãos públicos Federais e Estaduais utilizam Internet, sendo que 90% destes possuem *websites* (IBGE, 2018).

Especificamente com relação ao arquivamento das páginas do governo do Portal Único, pesquisadas por Melo (2020, p. 114 - 116), mais da metade dos *websites* tiveram seus recursos arquivados, ainda que quantidade relevante não tenha recuperado alguns dos disponibilizados quando o formato do arquivo não era textual e de imagem estática. No caso de áudios e vídeos se percebeu que o tempo de coleta ou a ferramenta WABAC, utilizada para reconstrução dos arquivos WARC capturados, não foram suficientes para recuperar os arquivos. Recursos hospedados em páginas diferentes das arquivadas, para medição de câmbios monetários, previsão do tempo e para interoperabilidade com as redes sociais, tenderam a não recuperação. Essas falhas podem ter ocorrido por diversos fatores, como esses recursos estarem hospedados em servidores diferentes do que está o *website* originalmente. O autor sugeriu o uso de alguma ferramenta auxiliar para a recuperação desses documentos não textuais, como áudios, vídeos e estudos empíricos para entender às necessidades e melhores ferramentas para seu restabelecimento. Em relação à recuperação do conteúdo textual, de imagens ilustrativas, sejam notícias, ícones, e a permanência do *feed* de notícias, percebeu-se que a recuperação foi satisfatória (a maioria dos *websites* arquivados apresentam esses conteúdos de forma integral). O autor aponta um desafio: o arquivamento de alguns conteúdos que não foram recuperados, principalmente os arquivos de áudio e vídeo, e os recursos hospedados em servidores externos.

Em relação à recuperação do conteúdo textual, de imagens ilustrativas, de notícias ou dos ícones, e da permanência do *feed* de notícias, a maioria dos *websites* arquivados apresentam esses conteúdos de forma integral, ainda que alguns não estejam formatados visualmente, tal como o *website* ao vivo. Apenas o *website* do Banco Central do Brasil não seria disponibilizado em uma plataforma de acesso por não ter recuperado seu conteúdo, os demais tiveram seu conteúdo informacional coletado (ainda que com falhas).

Para finalizar essa temática, tem-se a tese de Andrade da Luz (2021, p. 115-116), que mostrou que os produtos da comunicação governamental registram a política, a administração pública, a cultura e a sociabilidade de um determinado momento histórico. Seu trabalho apresentou modelos de preservação de sites oficiais desenvolvidos em seis países (Estados Unidos, México, França, Portugal, Chile e Brasil), para analisar a preservação e o apagamento da comunicação dos mandatos da presidenta Dilma Rousseff no site oficial da Presidência da República, e avaliou o acervo disponível. Segundo a autora, diferentemente dos países listados, o Brasil não possui um programa de arquivamento da *Web* dos seus sites oficiais, embora em 2010, o Arquivo Nacional brasileiro tenha instituído o Programa Permanente de Preservação e Acesso a Documentos Arquivísticos Digitais (AN Digital), no qual se comprometia preservar diversos tipos de documentos nascidos digitais. Outra iniciativa que não foi adiante foi a do Conarq, que em agosto de 2021 aprovou a atualização do “Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos” (e-ARQ Brasil), mas também adiou a discussão sobre os requisitos para a gestão arquivística de *websites*. Para ela, no Brasil, não há política de preservação dos arquivos da comunicação governamental publicada em sites oficiais, o que faz o país sofrer sem legislação específica para impedir a perda desses documentos. Ela aponta um caso único de preservação de *website* oficial identificado no governo federal brasileiro que é o site do governo do presidente Michel Temer (2016-2018), o que mostra que a Presidência da República tem conhecimento, tecnologia e recursos disponíveis para garantir o acesso presente e futuro aos acervos da comunicação governamental. A autora concluiu que o site oficial da Presidência da República brasileira não atende plenamente aos direitos à informação e à memória, reforçando a importância da adoção de uma política de Memória Comunicacional.

3.3.4.1 Políticas de arquivamento de páginas governamentais

De acordo com Brayner (2019, p. 5), dentre as políticas de arquivamento de páginas governamentais, a coleta do domínio do governo brasileiro é feita de maneira automática por um *crawler*, e esse tipo de arquivamento é problemático em virtude do grande volume de dados. O obstáculo não está no armazenamento, mas no tratamento da quantidade de informação coletada. Outra política adotada por muitas instituições é a coleta seletiva, onde se estipulam quais *sites* serão arquivados e a periodicidade das coletas, sendo a captura das páginas realizada com interação humana: a quantidade capturada é substancialmente menor, e a qualidade do arquivamento seletivo é superior ao automático. Tais políticas de arquivamento não são excludentes, e a necessidade de seleção e arquivamento de páginas *Web* é imperativa.

Quanto aos aspectos positivos e sensíveis das políticas de arquivamento, o primeiro a ser elencado é o controle de acesso. O nível de arquivamento se refere à coleta do site em sua integridade ou não, a quantas páginas serão coletadas, se não todas, e quais conteúdos nelas contidos serão priorizados, sejam textos, imagens e/ou vídeos. Já a periodicidade de coleta precisa ser estabelecida de acordo com a dinâmica de alteração/atualização dos sites selecionados. Um dos aspectos mais sensíveis diz respeito à política de reuso dos conteúdos arquivados, determinada por uma série de legislações que protegem a privacidade, os dados pessoais e os direitos autorais.

No que concerne à *Web* brasileira, Brayner (2019) informa que há mais de 4 milhões de endereços registrados, se for considerada a quantidade de páginas, esse número aumenta exponencialmente. Como as páginas *Web* são efêmeras: 80% delas desaparecem entre 12 meses e dois anos, perde-se muita informação, apesar do Brasil contar com a Lei de Depósito Legal (Lei 10.994/2004), no qual o artigo 2 dessa lei ainda não se cumpriu no sentido de promover o arquivamento da *Web* brasileira.

Um cunho muito prático é trazido por Gomes (2019) com sua experiência no *Arquivo.pt* (Portugal), para ele, os principais obstáculos e avanços no percurso do *Arquivo.pt* foram recrutar e treinar recursos humanos especializados, além da imaturidade da tecnologia, à época, para o arquivamento web. Para convencer os órgãos governamentais sobre a importância de preservar conteúdo *Web*, deve-se começar mostrando que informação é poder, e que ela, no século XXI, é conteúdo *Web*, portanto, preservar tal conteúdo também é poder. Sobre o que alguns alegam como empecilho para o arquivamento da *Web*, o alto custo do projeto, ele diz que perder a informação é ainda mais caro.

Uma importante iniciativa do governo brasileiro é o Portal do Livro Aberto, um serviço do Ibict, que tem por objetivo reunir, divulgar e preservar as publicações oficiais em diversas áreas, entre elas a ciência, tecnologia e inovação, as tecnologias da Informação e comunicação, e a Ciência da Informação. Ele objetiva contribuir na visibilidade à publicação oficial brasileira e, como repositório das publicações oficiais, é um projeto que está diretamente ligado à sua missão: fomentar e articular infraestrutura de organização, armazenamento, disseminação, acesso, preservação e análise da informação, por meio dos fundamentos da Ciência da Informação (INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2021?).

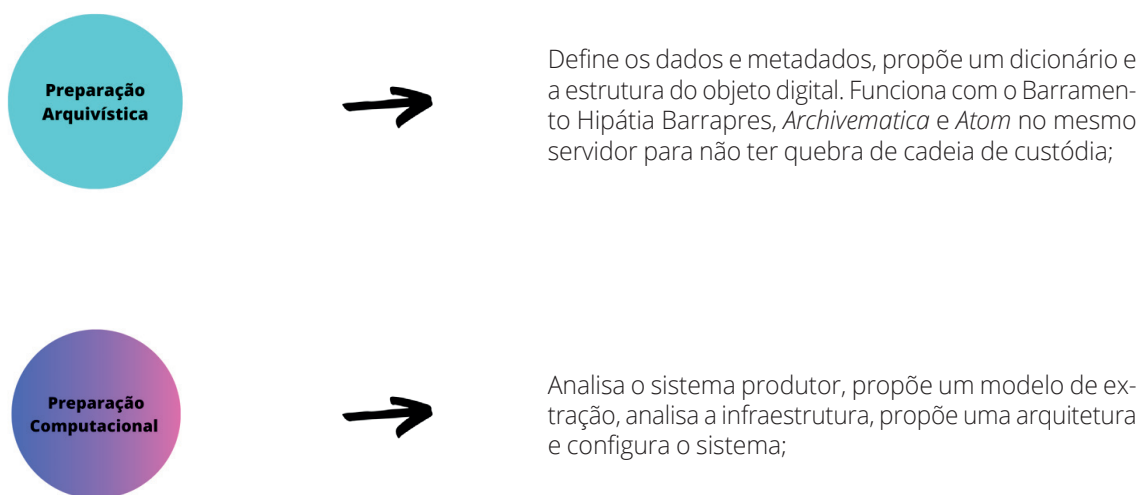
Na maioria das universidades, as funções administrativas de aquisição para as bibliotecas e para os arquivos são distintas, os artigos, relatórios, boletins informativos e outros itens publicados são comuns em sites, mas a relação

entre os documentos também é importante (DAVIS, 2014). O uso do termo “arquivo” em relação ao arquivamento da *Web* também é problemático porque a palavra está mais ligada ao uso computacional do arquivo, como uma cópia de segurança da informação, e depois vem a perspectiva disciplinar dos arquivos (OWENS, 2014). Para enfrentar esse desafio, uma iniciativa na Universidade de Victoria, nos Estados Unidos, foi criada e supervisionada por um grupo de trabalho do *Archive-it*, composto por integrantes com conhecimentos técnicos e de coleções: um Bibliotecário da área de Recursos de Aprendizagem e Pesquisa, o chefe dos serviços de gestão de coleções da biblioteca, o diretor de Coleções Especiais e um Arquivista, o que mostra que nem sempre a equipe precisa ser grande, e também revela a necessidade e a riqueza de uma equipe multidisciplinar (DAVIS, 2014).

Serviços como o Portico, o LOCKSS e o Portal *Scholar* fazem a preservação a longo prazo do conteúdo do *Open Journal System* (OJS), porém, a literatura de acesso aberto, especialmente de editoras menores e dos sites, tende a ser mais vulnerável a perdas do que o conteúdo baseado em assinatura, porque as bibliotecas não possuem o incentivo de investir em preservação a longo prazo, de acordo com o *Public Knowledge Project* (PKP) (PUBLIC KNOWLEDGE PROJECT, 2022). Ao contrário das organizações governamentais americanas, que têm um mandato para coletar certos tipos de *sites*, as bibliotecas e os arquivos universitários raramente têm essas obrigações, e seus focos, geralmente, são empregar o orçamento com livros, periódicos, bancos de dados de pesquisa e outros recursos pagos. Ainda há pouco incentivo para construir coleções de sites arquivados, o que envolve esforço e deixar de lado outras atividades de construção de coleções, relacionadas às aquisições. Assim, qualquer iniciativa de arquivamento da *Web* terá que considerar como envolver os bibliotecários no processo de construção de coleções digitais (DAVIS, 2014).

3.3.5 Estudos para o aprimoramento de modelos de preservação digital arquivística

Um modelo brasileiro para preservação digital que foi desenvolvido pelo Ibict é o Hipátia. O modelo implementa os repositórios RDC-Arq, repositórios digitais arquivísticos confiáveis. Ele atende a cinco 5 etapas (MODELO DE PRESERVAÇÃO DIGITAL HIPÁTIA, 2021):



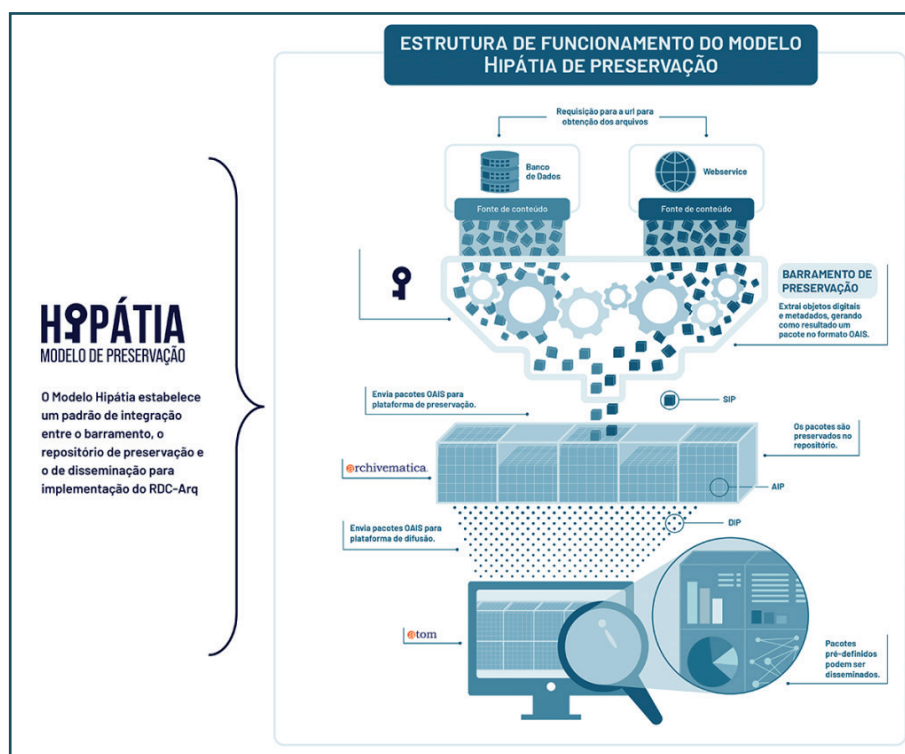
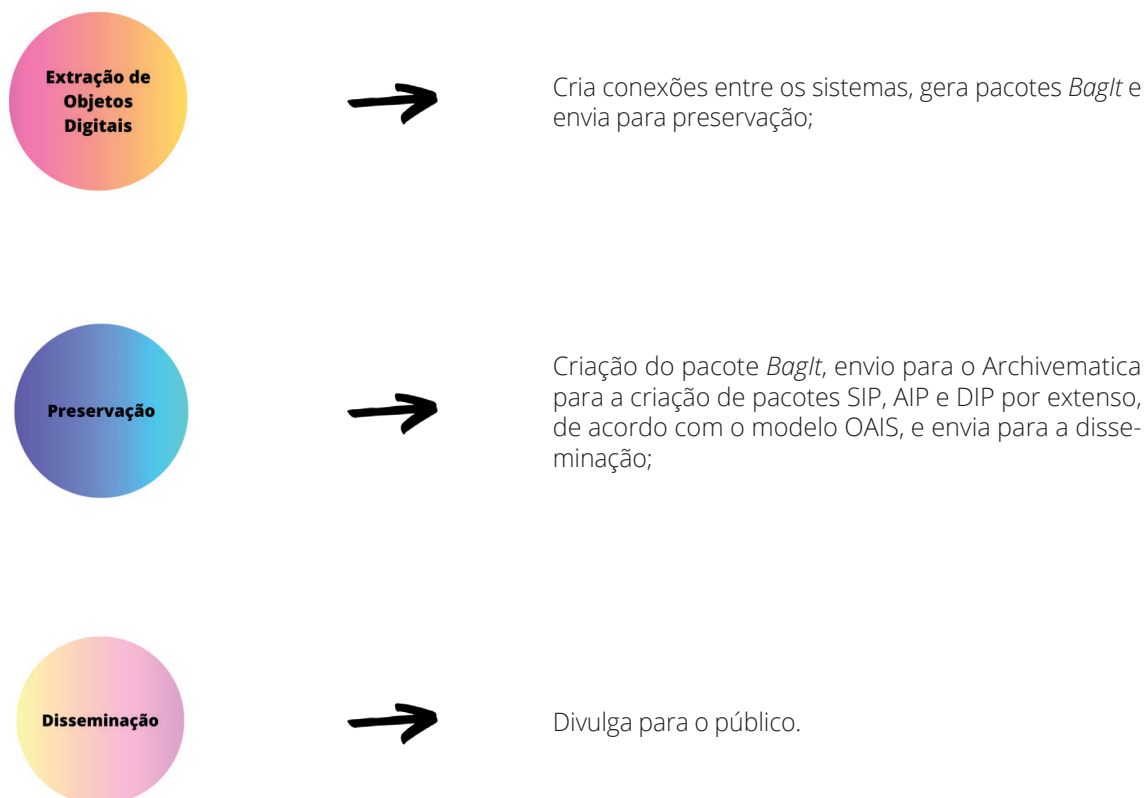


Figura 7 - Estrutura de Funcionamento do Modelo de Preservação Hipátia

Fonte: Modelo de Preservação Digital Hipátia (2021).

Dentre as pesquisas aplicadas a partir do Hipátia, tem-se os casos de sucesso, exemplificados pelo Tribunal de Justiça do Distrito Federal e Territórios (TJDFT), sendo realizada a integração do Diário de Justiça Eletrônico (DJe) e o RDC- Arq e, após o êxito dessa implantação, o TJDFT manteve a parceria de pesquisa para a implementação da solução tecnológica com os demais sistemas da instituição, em especial com o Processo Judicial Eletrônico (PJe). Além do TJDFT, cita-se o Arquivo Nacional, que buscou parceria com o Ibict para realizar a integração com o Sistema Eletrônico de Informações (SEI). Foi um projeto de pesquisa que deu ao Hipátia uma grande flexibilidade de integração com vários sistemas produtores de informações arquivísticas, oferecendo robustez necessária para validar realmente como um barramento de dados, ao garantir integridade e autenticidade e essa automação proporcionada pelo barramento, facilita e agiliza o processo de preservação digital nas instituições.

Elenca-se ainda os projetos que estão em andamento, exemplificados pelo Tribunal de Justiça de Minas Gerais, no qual a pesquisa inicialmente está definindo as demandas de gestão e preservação dos documentos digitais do Tribunal. Em paralelo, a equipe está avaliando o escopo tecnológico e documental da instituição, para então estabelecer as estratégias de trabalho em parceria com a equipe técnica do órgão. Como desdobramento desse trabalho, serão definidos os modelos de preservação para repositórios arquivísticos digitais aplicáveis ao sistema PROJUDI com extração de metadados do SIJUD. Apresenta-se também o TRT4 da 4ª Região que, impulsionado por razões que visam a cumprir recomendações de preservação digital apontadas por instituições como Conarq e CNJ, detectou a necessidade da criação e implementação de um Repositório Arquivístico Digital Confiável como uma forma segura de gerir e preservar informações digitais por meio de pacotes que devem seguir em cadeia de custódia plena e ininterrupta entre os sistemas de negócio, plataforma de preservação e de acesso/disponibilização. A competência desenvolvida pelo Ibict em sua área de atuação está alinhada à demanda apresentada pelo TRT 4ª Região, além disso, o apoio do Ibict às atividades de preservação digital também está expressa no Plano Diretor: 2018-2022 do Instituto.

Resumindo, o Hipátia é um modelo de preservação digital para implantar os Repositórios Arquivísticos Digitais Confiáveis, garantindo a cadeia de custódia dos objetos digitais arquivísticos, e integrando os sistemas de barramento Hipátia, Archivematica e AtoM. Com isso, o Hipátia pode ser considerado como uma plataforma de integração de sistemas voltada à preservação digital, no qual possibilita a integração de diversos sistemas produtores de documentos.

3.4 Considerações finais sobre a Meta 2

O conteúdo que circula na *Web* é dinâmico e temporário, especialmente o nato digital, e isso evidencia uma contradição da Era Digital: a abundância de informações, de um lado, e escassez de memória, de outro, o que torna seu arquivamento digital necessário. Alguns dos desafios para a preservação da rede são o fato de seus documentos terem uma natureza descentralizada, permeada de *hiperlinks*, e essencialmente rápida na obsolescência tecnológica. O ponto de partida para enfrentá-los é o planejamento, no qual são definidos o objetivo do arquivamento, a metodologia e as plataformas para coletar, armazenar e disponibilizar o conteúdo digital. Esses passos (coletar, armazenar e disponibilizar) correspondem às etapas do processo de arquivamento *Web*, ou de seu ciclo de vida.

Estudar sobre arquivamento leva ao sentimento de que, ao mesmo tempo em que se fica comovido por causa de informações do passado que não chegaram aos dias atuais, vê-se quão negligentes muitos são em relação ao arquivamento, em especial, de páginas *Web*. Possivelmente, no futuro, essa geração será conhecida como a “Era do Erro 404”, a sociedade sofrerá o impacto desse vácuo informacional, e esse quadro não se alterará se a guarda do conteúdo digital ficar sem diretrizes e políticas estabelecidas e aplicadas. Todos são responsáveis pelo legado histórico-cultural a ser deixado para as próximas gerações.

A motivação para este trabalho veio de uma necessidade do Ibict ligada ao arquivamento de conteúdo digital do governo, e pode ser realizado com a contratação de especialistas e com o agregar de saberes dos grupos de estudos. É um tema de grande importância, pela contemporaneidade dos objetos arrolados e pela reunião da literatura atualizada, acrescida das discussões que a seguem, o que o justifica. Por mais que o contexto de arquivamento de conteúdo digital seja discutido dentro e fora do Brasil há anos, internamente ainda é mais teórico que prático, haja vista a pouca quantidade de iniciativas nacionais comparadas às internacionais.

No que se refere aos *websites* governamentais, estes têm um papel importante na disseminação de informações, mas seu conteúdo pode mudar diariamente, ou ser excluído permanente e inesperadamente. Se as páginas do governo não forem digitalmente arquivadas, isso pode levar à perda de informações valiosas para pesquisas e até para a prestação de contas das ações do próprio governo.

Alguns desafios se impõem: serviços como o Pórtico, o LOCKSS e o Portal *Scholar* já fazem a preservação a longo prazo do conteúdo do *Open Journal System* (OJS), porém, baseada em assinatura (paga) e assim, a literatura de acesso aberto, especialmente de editoras menores e dos sites, tende a ser mais vulnerável a perdas do que o conteúdo baseado em assinatura. As bibliotecas e os arquivos universitários raramente têm a obrigação de investir em preservação em longo prazo, e seus focos muitas vezes ainda são empregar o orçamento em livros, periódicos, bancos de dados de pesquisa e outros recursos pagos. Ainda há pouco incentivo para construir coleções de sites arquivados, o que envolve esforço e deixar de lado outras atividades de construção de coleções, relacionadas às aquisições. Assim, qualquer ação de arquivamento da *Web* terá que considerar como envolver, e treinar, os bibliotecários no processo de construção de coleções digitais.

Outro ponto importante é o profissional da informação, é de grande relevância nesse contexto. Há que se destacar a necessidade de uma mudança na forma de agir em relação ao planejamento do processo, à quantidade e à velocidade no arquivamento, no que tange o gerenciamento, armazenamento e recuperação de conteúdo *Web*. Não se trata apenas de algo a mais na formação do profissional da informação, mas em uma mudança exponencial na forma como eles lidam com dados, informações, documentos e tecnologia. É importante que esses profissionais aprimorem suas competências tecnológicas para melhorar os sistemas de produção digital e o estabelecimento de uma relação entre o arquivamento da *Web*, com a produção governamental no ambiente *Web*. Seria interessante que essas melhorias considerassem os padrões estipulados pelo consórcio W3C, é a partir desses padrões que serão desenvolvidas ferramentas e metodologias para o arquivamento da *Web* com melhores resultados. A Universidade de Victoria, nos Estados Unidos, buscou uma alternativa para incrementar o potencial de seu pessoal com a implementação de uma equipe multidisciplinar, o que já se sabe ser o ideal. Ela foi criada e supervisionada por um grupo de trabalho do *Archive-it* composto por integrantes com conhecimentos técnicos, de coleções e de tecnologia, o que mostra a necessidade e riqueza de um grupo multidisciplinar.

Embora o arquivamento esteja acontecendo, uma limitação verificada é que alguns recursos presentes nos *websites* ainda não são recuperados, especialmente, quando o formato do arquivo é diferente do textual e da imagem estática. Áudios, vídeos e recursos hospedados em servidores externos tendem a não recuperação, assim entende-se que se faz necessário o uso de ferramenta auxiliar para recuperação dos documentos não textuais. Por outro lado, nos lugares testados e identificados no trabalho, a recuperação foi considerada satisfatória quando os resultados mostraram que a maioria dos *websites* arquivados apresentavam seus conteúdos de forma integral.

Os *websites* do Governo Federal Brasileiro são arquiváveis sem perda de informações relevantes, embora não estejam sendo arquivados, e o país carece de uma política pública para sistematizar tal processo ligado aos *websites* governamentais. A implantação de uma política para preservação de páginas *Web* governamentais no Brasil é um grande desafio, inclusive no convencimento sobre a necessidade e quanto a questões técnicas associadas. Outra dificuldade é encontrar estudos sobre esses temas na literatura brasileira, talvez por ser um tema ainda recente de pesquisa interna (no Brasil), uma vez que já está se consolidando no exterior há alguns anos. Deve-se reforçar sua importância, pois envolve a memória da *Web* brasileira e o acesso a essas informações no presente e no futuro.

4. RESULTADOS META 3

Na sociedade contemporânea, a acessibilidade digital e o acesso ao conteúdo informacional disponibilizado nos *websites* governamentais são cruciais para o exercício da cidadania. De acordo com a legislação brasileira, todas as pessoas, independentemente de suas habilidades ou deficiências, devem poder acessar os conteúdos informacionais nos *websites* governamentais e conseguir realizar tarefas ou terem suas necessidades satisfeitas (BRASIL, 2015; SOUSA; NAKANO, SHINTAKU, VIDOTTI, 2022)

Infelizmente, até hoje, mesmo com as previsões legais associadas à acessibilidade digital e informacional, muitos *websites* e documentos governamentais ainda não são projetados tendo em mente a acessibilidade, apresentando barreiras que dificultam o acesso e a compreensão das informações de que as pessoas com deficiência precisam (NAKANO, 2022). É nesse cenário que a acessibilidade se tornou uma questão crítica para organizações e instituições que se esforçam para garantir a igualdade de oportunidades para todos os indivíduos.

Este relatório apresenta os resultados parciais da meta 3: *Desenvolver modelos e orientações para criar ou converter documentos digitais, de forma a torná-los acessíveis, contribuindo para a democratização da informação*, realizada no contexto de apresentação da informação acessível em *websites* governamentais, documentos acessíveis e linguagem acessível para a democratização da informação.

A pesquisa empregada para alcançar os objetivos da meta 3 envolveu uma abordagem de métodos mistos, combinando métodos de coleta de dados qualitativos, pesquisa exploratória e pesquisa aplicada. O estudo incluiu pesquisa bibliográfica e documental, técnicas de observação, bem como avaliações da acessibilidade de documentos e *websites* digitais. O estudo também explorou as melhores práticas e os padrões para documentos acessíveis e linguagem acessível, incluindo as Diretrizes de Acessibilidade de Conteúdo da Web (WCAG) 2.1.; o e-MAG do governo federal bem como as orientações da literatura para redação em linguagem clara e concisa.

Além de apresentar os resultados da Meta 3, o relatório fornece *insights* sobre os desafios enfrentados por pessoas com deficiência no acesso a conteúdo digital, por meio das publicações resultantes das pesquisas, e destaca a importância de projetar documentos acessíveis e usar linguagem acessível para promover a democratização da informação.

O relatório também propõe estratégias práticas para que organizações e instituições, aqui incluído o governo federal, se tornem mais inclusivas melhorando a acessibilidade de seu conteúdo digital e garantindo que todos, independentemente das habilidades e deficiências, tenham acesso à informação.

4.1 Objetivo geral da Meta 3

Desenvolver modelos e orientações para criar ou converter documentos digitais, de forma a torná-los acessíveis, contribuindo com a democratização da informação.

4.2 Objetivo específico da Meta 3

- Desenvolver modelo e orientações para criação ou conversão para documentos digitais acessíveis;
- Criar ou converter documentos para o modelo;
- Avaliar e ajustar modelos e orientações.

4.3 Descrição de resultados da Meta 3

As investigações preliminares foram realizadas no sentido de fazer um diagnóstico dos estudos que envolvem não apenas acessibilidade digital, mas também estudos sobre a acessibilidade do conteúdo informacional, ou seja, a intenção era aprofundar os estudos já existentes que focam na eliminação de barreiras na interface do ambiente digital e contemplar até o conteúdo, no caso do nosso projeto, documentos governamentais, e como criá-los de forma acessível. Preliminarmente, também realizamos estudos para levantar a legislação que ampara a acessibilidade digital e informacional no Brasil.

Com relação ao levantamento sobre documentos governamentais acessíveis, percebemos uma lacuna na literatura. É possível encontrar inúmeras publicações sobre acessibilidade digital, como produzir documentos textuais acessíveis, mas não encontramos estudos sobre como criar e-mails acessíveis. Considerando que os e-mails são formas oficiais de comunicação entre servidores públicos e entre servidores e cidadãos, consideramos essa uma lacuna relevante. Assim, justificamos o primeiro resultado da Meta 3, já publicado em revista em 2022, o artigo intitulado: *Elementos fundamentais para a criação de e-mail acessíveis*.

Com relação ao estudo preliminar de levantamento da legislação brasileira sobre acessibilidade, destacamos o estudo intitulado: *Elementos de amparo legal sobre acessibilidade digital para os surdos no Brasil*, artigo publicado em revista em 2022.

Com o projeto em andamento, alguns dos resultados descritos a seguir ainda podem sofrer mudanças, na medida em que outros resultados se mostram relevantes e divergem do que já está estabelecido. Assim, apresentamos os resultados parciais do projeto.

4.3.1 Desenvolver modelo e orientações para criação ou conversão para documentos digitais acessíveis

O desenvolvimento de guias para orientar sobre a criação ou conversão de documentos digitais são uns dos resultados parciais dos estudos. Tais obras se encontram em fase final de editoração ou em fase final de elaboração. Como demonstrado a seguir:

- Guia de boas práticas para a criação de e-mails acessíveis

O guia está na fase final de editoração, faltando a parte de revisão ortográfica e diagramação. A obra tem como objetivo apresentar uma série de orientações, resultados de estudos sobre acessibilidade, voltadas a tornar os e-mails mais acessíveis, principalmente para órgãos de governo, que os utilizam, muitas vezes, como forma de interação interna e externa. Desse modo, colabora com a discussão sobre acessibilidade de documentos governamentais, na medida em que o Brasil é um dos países signatários do Tratado de Marraqueche, que diz respeito à acessibilidade de documentos, sobretudo para os deficientes visuais. Nesse contexto, oferece à sociedade brasileira um compêndio de boas práticas na confecção de e-mails, de forma a torná-los acessíveis às diversas necessidades.

- Guia Prático do Português Simplificado para documentos acessíveis

O guia ainda está em fase de elaboração de conteúdo. Esse trabalho tem como objetivo ofertar modelos para criar ou converter documentos, prioritariamente governamentais, a orientações para acessibilidade, com foco na simplificação da língua portuguesa. Visa a contribuir com a discussão sobre acessibilidade à informação por meio de orientações para simplificação do português a ser utilizado na apresentação de informações em documentos prioritariamente governamentais. Nota-se que esse modelo de simplificação da língua já é utilizado para outros idiomas, como no inglês simplificado (*plain english*), amplamente fomentado no ensino dessa língua como idioma estrangeiro.

- Guia de Boas Práticas para Criação de Documentos Digitais Visuais Acessíveis em Libras

O guia está em fase final de desenvolvimento e se fundamenta nas recomendações da WCAG, o eMAG, no manual de Carpes e Soster (2016), e na norma NBR 16.452:2016 da Associação Brasileira de Normas Técnicas (ABNT), que trata sobre acessibilidade na comunicação. (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2016). Vale notar que a equipe conta com uma pesquisadora CODA (*child of deaf adult*) e um pesquisador surdo para gravação e revisão dos vídeos.

- Guia de Boas Práticas para Criação de Documentos Digitais Acessíveis em Libras

Esse guia está em fase final de desenvolvimento pelos mesmos pesquisadores do guia mencionado acima.

- Documentos acessíveis: da construção às boas práticas

A obra está em fase de elaboração de conteúdo. Essa obra tem como objetivo apresentar os estudos realizados ao longo do projeto. Nesse sentido, o livro abordará sobre a acessibilidade e a documentação, as questões legais relacionadas a essa documentação acessível, o banco lexical desenvolvido no projeto, o modelo de simplificação da língua portuguesa, a acessibilidade da informação em *website*, o tesaurus desenvolvido no projeto, a qualidade da informação do tesaurus e, por fim, a adaptação de textos para linguagem clara.

4.3.2 Criar ou converter documentos para o modelo

Outro resultado do estudo que converge as pesquisas realizadas é a criação de um modelo de *website* de tesauros acessível na ferramenta Tematres. O *website* converge e integra os recursos de acessibilidade preconizados pelo W3C e pelo governo federal.

A publicação referente à criação do ambiente está em desenvolvimento para ser submetida à conferência ISKO 2023, em junho, em Londrina. O artigo é intitulado: *Proposta de modelo para apresentação da informação acessível em tesauros implementado com a ferramenta Tematres*.

A equipe de pesquisa multidisciplinar do projeto permitiu o desenvolvimento dos seguintes trabalhos:

- Criação dos termos do tesauros e sua descrição padrão;
- Criação da descrição do termo em linguagem simples;
- Gravação e edição dos vídeos traduzidos em linguagem brasileira de sinais;
- Criação do *website* e apresentação da informação acessível, e avaliação do *website*.

Apresentados na subseções seguintes.

4.3.2.1 O Tesouro de Documentos Acessíveis

O Tesouro de Documentos Acessíveis tem um sistema de informação que visa a apresentar os termos relacionados à acessibilidade de documentos digitais, preferencialmente governamentais, atendendo a orientações de acessibilidade que abrangem diversas necessidades especiais. O tesouro mapeia os conceitos reunidos em três categorias: 1. Acessibilidade; 2. Documento Governamental Acessível; 3. Pessoa com Barreira de Acessibilidade. Está em consonância com as normas ISO 25964 Part I e Part II (2011; 2013), que cobrem todos os aspectos relacionados ao desenvolvimento de tesauros monolíngues e multilíngues destinados à recuperação de informação, e com a ANSI/NISO Z39.19 (2010), que apresenta as diretrizes para a construção, o formato e a gestão de vocabulários controlados monolíngues.

Relacionado a este item e à equipe do tesauros acessível, destacamos a publicação em desenvolvimento intitulada: *Formação da terminologia sobre acessibilidade em tesauros implementado com o Tematres*. A publicação objetiva apresentar o processo de construção da terminologia sobre acessibilidade na ferramenta Tematres, e deve ser submetida ao evento ISKO 2023. Além da publicação em evento, enfatizamos a publicação da política intitulada: *Política de gestão e governança do tesouro acessível*. Trata-se de documento que tem o propósito de orientar o gerenciamento do Sistema de Tesauros de Documentos Acessíveis, implementado com a ferramenta livre TemaTres.

4.3.2.2 Descrição padrão do termo em linguagem simples

O modelo de tesouros traz a descrição padrão do termo e inclui uma descrição em linguagem simples. A descrição em linguagem simples foi estudada e elaborada pelas linguistas que compõem a equipe multidisciplinar envolvida na construção do tesouros acessível, seguindo a recomendação do e-MAG para linguagem clara, acessível e com léxico que não exija mais do que o ensino fundamental.

Aqui vale apresentar o resultado relacionado em desenvolvimento intitulado: *Formação de banco lexical para simplificação da informação em língua portuguesa*. O banco lexical está sendo desenvolvido com o objetivo de apresentar o léxico apontado pelo e-MAG como sendo fundamental para compor a linguagem clara e acessível. Embora a recomendação do e-MAG seja de que léxico não deve exigir mais do que o ensino fundamental, não foram levantados dados na literatura sobre quais palavras compõem esse léxico. Assim, diante da lacuna percebida, temos como resultado parcial o desenvolvimento do banco lexical.

Ainda relativamente à recomendação de utilização de linguagem clara, concisa e em português simplificado, o artigo *Acessibilidade e disseminação da informação: eliminação de barreiras físico-cognitivas e de linguagem na construção de um tesouro acessível*, tem como objetivo apresentar os resultados de um estudo voltado à criação de um tesouros on-line que ofereça acessibilidade ampla ao seu conteúdo, com funcionalidades que atendam às necessidades de pessoas com diferentes tipos de deficiência, bem como das pessoas que não foram plenamente alfabetizadas. Nessa linha, também destacamos o trabalho enviado para o XI Congresso Argentino de Semiótica intitulado: *Práticas semióticas de acessibilidade: questões semânticas e estratégias enunciativas*.

4.3.2.3 Gravação e edição dos vídeos traduzidos em linguagem brasileira de sinais

Os conteúdos sem acessibilidade representam barreiras “nas comunicações e na informação” ou barreiras “tecnológicas”, que atrapalham o cumprimento do Estatuto da Pessoa com Deficiência (EPcD) e o pleno exercício da cidadania por todas as pessoas. (BRASIL, 2015). Assim, a criação de vídeos para tornar acessíveis conteúdos físicos ou digitais, como esse tesouros, deve contribuir para o acesso de mais pessoas à informação, para que seja cumprido o EPcD.

4.3.2.4 Criação do *website* e apresentação da informação acessível, e avaliação do *website*

Diferentes instituições de ensino e pesquisa, bem como organizações governamentais regulatórias, tratam o assunto acessibilidade – seja digital ou física – relacionado ao Objetivo de Desenvolvimento Sustentável (ODS) número 10: reduzir a desigualdade dentro dos países e entre eles. Nesse projeto de pesquisa, destacamos especialmente o W3C e o e-MAG como diretrizes principais para eliminação de barreiras à compreensão da mensagem contida em endereços na internet e, conseqüentemente, eliminação de barreiras para acessar seu conteúdo informacional. De acordo com o e-MAG, há três passos para a criação de um *website* acessível, primeiramente, seguir os padrões *Web*; em segundo lugar, seguir as recomendações de acessibilidade, e finalmente realizar a avaliação de acessibilidade.

Além das diretrizes internacionais e do e-MAG, a apresentação da informação no *website* do tesaurus acessível foi pautada nos princípios do Design da Informação, no Design Universal e Encontrabilidade da Informação (VECHIATO, VIDOTTI, 2014). Um artigo detalhando esses princípios que nortearam a construção do ambiente do tesaurus está em fase final de desenvolvimento e será submetido para publicação em revista científica com o seguinte título: *Princípios do Design da Informação, do Design Universal e a Encontrabilidade da informação como fundamentos para criação de um website de tesaurus acessível*.

Ainda com relação à ferramenta Tematres utilizada na construção do *website* do tesaurus acessível, temos como resultado da pesquisa o artigo intitulado: *Proposta de modelo para apresentação da informação acessível em tesaurus implementado com a ferramenta Tematres*.

4.3.3 Avaliar e ajustar modelos e orientações

Finalizado o estágio de construção da estrutura e alimentação, o *website* do tesaurus acessível se encontra em estágio de avaliação por software para validação dos padrões *Web* e validação automática de acessibilidade. Para validar os códigos do conteúdo HTML e das folhas de estilo, os validadores automáticos do W3C foram utilizados. Ao passo que, para a validação automática, foram utilizadas ferramentas gratuitas e específicas de acessibilidade de acordo com a lista disponibilizada pela WAI com as diretrizes do WCAG.

A literatura também recomenda uma validação manual dos recursos de acessibilidade. Na etapa de validação manual foram realizados testes com os recursos do leitor de tela e navegação por tab e teclado para garantir as diretrizes específicas do WCAG e do e-MAG. Nessa validação foram identificados pontos importantes para adequação do Tesaurus acessível, que estão em fase final de ajustes. Como limitação da fase de avaliação, relatamos a falta de avaliação com pessoas deficientes, que pode ocorrer, em um momento futuro, com questionários remotos e grupos focais presenciais, por exemplo.

4.4 Considerações finais sobre a Meta 3

Neste relatório parcial, destaca-se os trabalhos finalizados, aqueles em fase final e alguns trabalhos em andamento. O projeto de pesquisa “Ecosistema de Informação Governamental”, desenvolvido pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) – unidade de pesquisa vinculada ao Ministério da Ciência, Tecnologia e Inovação (MCTI) –, por meio da Coordenação de Tecnologias para Informação (Cotec), assume o compromisso de colaborar com o acesso às tecnologias e conteúdo da Internet que são hoje imprescindíveis para o pleno exercício da cidadania de todos, independentemente das suas habilidades e deficiências. Para tanto, temos, diante de nós, a tarefa de responder com estudos que promovam avanços e ampliação nos modelos, aportes teóricos e metodológicos, na integração de conceitos provenientes de outras áreas, e na adoção de uma conduta interdisciplinar apta a estabelecer interfaces que tragam novos suportes e respostas.

5. RESULTADOS META 4

Como consequência do avanço e aprimoramento das ações do governo eletrônico (e-GOV), a informação governamental passou a apresentar características mais acessíveis. Como a Lei de Acesso à Informação (LAI) e o incentivo à transparência por meio dos dados abertos de governo e a Lei Orçamentária Anual (LOA), por exemplo, que possibilitam o acesso a dados e informações governamentais, mas sempre de forma particionada. Todavia, mesmo com os avanços observados, ainda existem, principalmente nos níveis estadual, municipal e distrital, barreiras técnicas e tecnológicas a serem superadas.

Dentre os desafios, faz-se válido destacar a apresentação de informações e dados governamentais em formato geolocalizado. Essa modalidade de visualização de dados oferece facilidades, mas também exige precisão nas informações, dados públicos e de qualidade, bem como ferramentas que viabilizem a criação de apresentações gráficas e dinâmicas, preferencialmente gratuitas e abertas. Em muitos casos, a geolocalização é fundamental para contextualizar ou compor indicadores e informações sobre entidades e temáticas diversas, importantes para as tomadas de decisão dos gestores, e que impactam diretamente a vida e o bem-estar da população.

Diante da compreensão sobre a relevância dos dados e informações geolocalizadas para o contexto de proposição de um ecossistema de informação governamental, a Meta 4 do projeto de pesquisa, utilizando principalmente os aspectos de pesquisa aplicada, busca englobar atividades que tratam do desenvolvimento de novas funcionalidades informatizadas para apresentação de dados e informações geolocalizadas a partir da perspectiva da Ciência da Informação.

5.1 Objetivo geral da Meta 4

Desenvolver novas funcionalidades informatizadas para apresentação de dados e informações geolocalizadas amparada na Ciência da Informação.

5.2 Objetivo específico da Meta 4

- Estudos para verificar as relações entre ciência de serviços, informação estratégica, Business Intelligence, e-GOV, entre outros e informação e dados geolocalizados;
- Realizar estudos para implementação de melhorias no Visão para atendimento às novas funcionalidades;
- Desenvolver modelo de integração de dados de sistemas para o Visão.

5.3 Descrição de resultados da Meta 4

As subseções abaixo sintetizam os resultados obtidos até então para as atividades e objetivos previstos para a Meta 4.

5.3.1 Estudos para verificar as relações entre ciência de serviços, informação estratégica, Business Intelligence, e-GOV, entre outros e informação e dados geolocalizados

Inicialmente, foram identificados estudos e referências importantes para orientar os estudos comparativos entre as temáticas de interesse da Meta 4, assim, são listadas a seguir publicações caracterizadas como literatura de apoio às atividades da meta em questão, são elas:

ALBAGLI, Sarita; PY, Hesley; IWAMA, Allan Yu. Geovisualização de dados e Ciência aberta e cidadã-a experiência da Plataforma LindaGeo. **Digital Humanities Quarterly**, v. 14, n. 2, 2020.

DANKO, David M. Iso/tc211: geographic information–metadata iso 19115. In: **World Spatial Metadata Standards**. Elsevier Science, 2005. p. 535-555.

DE METADADOS GEOESPACIAIS, Comitê de Estruturação. Perfil de metadados geoespaciais do Brasil: perfil MGB. **Rio de Janeiro: CEMG-Concar**, 2009.

DA INFRAESTRUTURA, CINDE-Comitê de Planejamento et al. Plano de Ação para Implantação da Infraestrutura Nacional de Dados Espaciais. 2010.

REED, Carl N. The open geospatial consortium and web services standards. In: **Geospatial Web Services: Advances in Information Interoperability**. IGI Global, 2011. p. 1-16.

PY, Hesley. Visualização por meio de mapas como recurso para a democratização dos dados e para a apropriação social da informação. 2019.

5.3.2 Realizar estudos para implementação de melhorias no visão para atendimento às novas funcionalidades

Ao longo do andamento do projeto de pesquisa, por meio de testes de usabilidade e navegabilidade, e tendo em vista *feedbacks* recebidos pelos usuários do sistema, foi estabelecido um conjunto de melhorias a serem arquitetadas e desenvolvidas para o Visão. O andamento e a descrição de alguns aprimoramentos já realizados no sistema são apresentados abaixo.

5.3.2.1 Objetivos principais e secundários do sistema Visão definidos

O primeiro passo para encaminhar os estudos de melhorias focou na delimitação de objetivos claros para o sistema Visão. Após momentos de *brainstorming* e reuniões técnicas entre a equipe de pesquisadores que atuam no desenvolvimento do sistema, foi delimitado o seguinte escopo principal para o Visão:

Objetivos principais

- Ser uma ferramenta de democratização e popularização do acesso e entendimento ao significado presente nos dados, independente de sua origem, por meio da criação de visualizações, gestão e disponibilização de diferentes tipos de dados;
- Fomentar a construção de uma comunidade de usuários para discussão sobre dados e visualização de dados e informações.

Objetivos secundários

- Promover a integração de dados com outras bases e plataformas para a criação de visualizações;
- Ser um ambiente que popularize o manuseio dos dados, com uma interface intuitiva, acessível e inclusiva.

5.3.2.2 Requisitos da versão 2.0 do sistema Visão definidos

Além da definição de escopo, foram delimitados requisitos informacionais e tecnológicos necessários para a proposição e implementação da versão 2.0 do Visão. São eles:

Potenciais usuários

- Indivíduos e/ou projetos que receberam recursos públicos para pesquisa no território nacional e que precisam de um local para a publicação dos dados e informações geradas;
- Cientistas cidadãos (grupos ou indivíduo - sem recursos públicos) que precisam de uma ferramenta para guarda e publicação dos dados e informações;
- Órgãos públicos que precisam gerar dados e informações visuais;
- Cidadãos.

Principais funcionalidades

As funcionalidades mapeadas e classificadas como principais para serem implementadas no sistema são apresentadas no Quadro 3:

Principais funcionalidades
Elaboração de relatórios acerca dos dados contextualizados no Visão pelo Ibict. Imagina-se que os relatórios possam ser baixados na própria página do mapa;
Disponibilização do conjunto de dados pelo link da “fonte” que é apresentado do <i>pop up</i> da visão;
Simplificar o processo de <i>upload</i> de csv no sistema (reconhecimento do arquivo sem a necessidade de seleção de tipo de delimitador e separador);
Estrutura de metadados melhor delimitada de modo a permitir integração com outros sistemas;
Acessibilidade com esquema de cores;
Acesso simplificado à documentação do sistema, dando mais visibilidade aos documentos publicados;
Código aberto de verdade (Disponibilização no GitHub);
Interface com API, disponibilizar API de consulta, de integração e coleta de dados;
Estruturar um repositório público para a disponibilização de dados públicos utilizados pelo Visão;
Disponibilizar recursos que permitam que o usuário personalize o mapa (minhas visões) de acordo com seu interesse - opção de cor para daltônicos, personalização dos pins das camadas, integração com Voyant Tools, mapa com barras, mapa de rede;
Permitir que os usuários criem grupos dentro do sistema;
Ampliar as opções de compartilhamento das visões;
Filtrar/organizar os mapa/visões públicas, seja por acesso, tema, instituição;
Disponibilizar modelos de visualizações e permitir que os usuários editem segundo seus interesses;
Tradução do sistema no todo;
Permitir que os usuários façam as edições online e dentro do sistema Visão;
Permitir que o usuários criem ou importem conjuntos de dados;
Criar área administrativa única, que permita ao usuário o acesso e edição dos elementos seus e de seu grupo.

Quadro 3 - Principais funcionalidades a serem incorporadas ao Visão.

Fonte: Equipe Visão.

5.3.2.3 Protótipo da versão 2.0 do sistema validado

Durante a fase de execução da meta 4 do projeto, foi construído um protótipo funcional para as funcionalidades que demandavam maior necessidade de entendimento por parte dos desenvolvedores. A Figura 8 abaixo apresenta o protótipo para as funcionalidades de inserção de indicadores, camadas e filtros.

visão Fazer login

Botões de teste

X = Inserir X indicadores Post grup indicator

Validar +PageNumber +PageItems

Numero de indicadores: 2;
Numero de páginas: 0;
Página atual: 0;
Items por página: 100

Exportar JSON Importar JSON

Exportar CSV Importar CSV

Delimitador (células)
Sem aspas ()

Separador (colunas)
Ponto e vírgula (,)

Nome do grupo de indicadores

Tipo de dado

Grande área

Palavras chave

Descrição

Fonte (URL, Nome)

1

Ir para página 1

Linha	Região	Valor	Período	
1	AC	12	04/03/2023	Editar Excluir
2	MG	14	10/03/2023	Editar Excluir
Novo indicador	AC	0	mm/dd/yyyy	Adicionar

1

Ir para página 1

Salvar

Figura 8 - Protótipo de *upload* de indicadores, camadas e filtros

Fonte: Equipe Visão.

5.3.2.4 Definição e implementação de catálogo de metadados do sistema Visão

Para a implantação do catálogo foi utilizado o software Geonetwork, disponível em <https://geonetwork-opensource.org/>. O Geonetwork é um software livre disponibilizado pela OSGeo (The Open Source Geo Spatial Foundation), que provê um catálogo de metadados para a gestão de dados e informações espaciais. O software disponibiliza ferramentas para a catalogação, edição e busca de metadados. É utilizado em inúmeras infraestruturas de dados ao redor do mundo, incluindo a infraestrutura de dados espaciais Brasileira, a INDE.

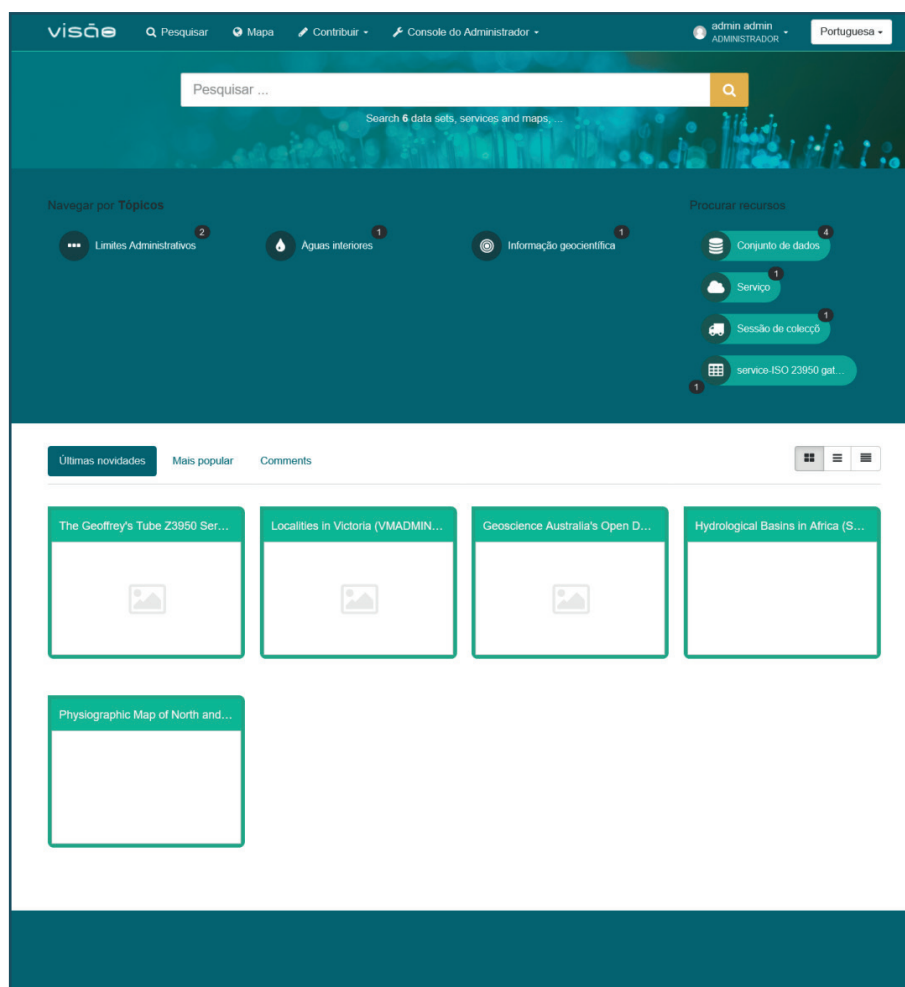


Figura 9 - Catálogo Geonetwork de metadados customizado para o Visão

Fonte: Equipe Visão.

5.3.2.5 Ferramentas para *upload* dos dados para a plataforma

Foram implementadas melhorias para facilitar o processo de carregamento e cadastramento de dados no Visão. O resultado do desenvolvimento pode ser visualizado na Figura 10.

Indicadores

Grupo _____

Nome do grupo de indicadores *

Descrição

Tipo de dado

Grande área

Palavras chave

Fonte (URL, Nome)

Tipo de região

Estado

(*) Campos obrigatórios

Inserir Indicadores _____

Figura 10 - Tela para inclusão / upload de indicadores

Fonte: Equipe Visão.

5.3.2.6 Desenvolvimento de área para gerenciamento de repositório pessoal de dados

A atividade resultou em uma área administrativa denominada “Minha biblioteca” que agrega, em um só lugar, todos os dados e visões de um usuário cadastrado no sistema.

Minha biblioteca

Minhas visões Adicionar visão

Visao 01
Proprietário: Grupo
Visões secundárias:
Metadados:

Visao 02
Proprietário: Grupo
Visões secundárias:
Metadados:

Visao 03
Proprietário: Usuário
Visões secundárias:
Metadados:

Meus dados Adicionar

Legenda: ☒ Indicador ☐ Filtro ☐ Camada

	Descrição	Metadados
☒ Usuário	name_ind 2 desc_ind 2	Ver
☒ Grupo	1 desc_lay	Ver
☒ Grupo	name_geo desc_geo 2	Ver
☒ Grupo	name_ind desc_ind	Criar

Figura 11 - Tela de minha biblioteca

Fonte: Equipe Visão.

5.3.2.7 Grupo de usuários

A versão atual do Visão conta com usuários únicos, autônomos, contudo verificou-se a necessidade da inclusão do conceito de grupos, de forma que um conjunto de usuários pudesse ser agregado em torno de um projeto ou seguimento específico. Testes estão sendo realizados para implantar a validação das funcionalidades para o conceito de usuários.

5.3.2.8 Visualizações secundárias

Inserção de funcionalidades para a criação de visões secundárias na interface de criação de visões. Trata-se da organização das visualizações de forma hierárquica, permitindo que algumas visões deem origem a outras, permanecendo vinculadas.

5.3.2.9 Integração de ferramentas

Integração das ferramentas Geonetwork e do sistema Visão para a geração da versão 2.0 da plataforma Visão. de visualização de dados contendo um cadastro e *login* integrado para o usuário do sistema.

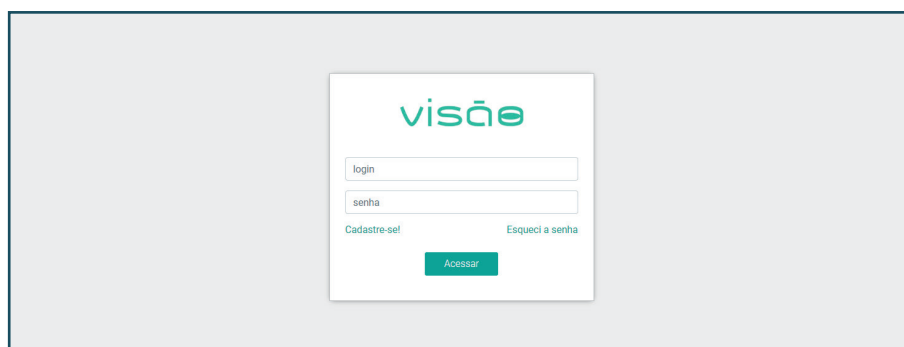


Figura 12 - Tela de login único

Fonte: Equipe Visão.

5.3.3 Desenvolver modelo de integração de dados de sistemas para o Visão

A seguir, são elencadas questões importantes para possibilitar a proposição de um modelo de integração de dados no âmbito do Visão.

5.3.3.1 Perfil de metadados, baseado no ISO 19115/19139 e no Perfil Brasileiro de Metadados Geoespaciais (PMGB)

O perfil de metadados adotado no Visão foi definido tendo por base o padrão ISO 19115/19139 e no Perfil de Metadados Geoespaciais do Brasil (PMGB). Sua definição a partir de um padrão ISO garante a compatibilidade entre os metadados disponíveis na plataforma com as principais plataformas para documentação e disseminação de metadados geoespaciais no mundo.

Segue abaixo um exemplo em XML (eXtensible Markup Language) de um metadado preenchido no perfil de metadados adotado no Visão.

```
<?xml version="1.0" encoding="UTF-8"?>
<gmd:MD_Metadata xmlns:gmd="http://www.isotc211.org/2005/gmd" xmlns:gml="http://www.opengis.net/gml" xmlns:gts="http://www.isotc211.org/2005/gts"
xmlns:gco="http://www.isotc211.org/2005/gco" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.isotc211.org/2005/gmd
http://schemas.opengis.net/csw/2.0.2/profiles/apiso/1.0.0/apiso.xsd">
  <gmd:fileIdentifier>
    <gco:CharacterString>da165110-88fd-11da-a88f-000d939bc5d8</gco:CharacterString>
  </gmd:fileIdentifier>
  <gmd:language>
    <gco:CharacterString>eng</gco:CharacterString>
  </gmd:language>
  <gmd:characterSet>
    <gmd:MD_CharacterSetCode codeListValue="utf8" codeList="./resources/codeList.xml#MD_CharacterSetCode"/>
  </gmd:characterSet>
  <gmd:contact>
    <gmd:CI_ResponsibleParty>
      <gmd:individualName>
        <gco:CharacterString>Jippe Hoogeveen</gco:CharacterString>
      </gmd:individualName>
      <gmd:organisationName>
        <gco:CharacterString>FAO - NRCW</gco:CharacterString>
      </gmd:organisationName>
      <gmd:positionName>
        <gco:CharacterString>Technical Officer</gco:CharacterString>
      </gmd:positionName>
      <gmd:contactInfo>
        <gmd:CI_Contact>
          <gmd:phone>
            <gmd:CI_Telephone>
              <gmd:voice>
                <gco:CharacterString/>
              </gmd:voice>
              <gmd:facsimile>
                <gco:CharacterString/>
              </gmd:facsimile>
            </gmd:CI_Telephone>
          </gmd:phone>
          <gmd:address>
            <gmd:CI_Address>
              <gmd:deliveryPoint>
                <gco:CharacterString>Viale delle Terme di Caracalla</gco:CharacterString>
              </gmd:deliveryPoint>
              <gmd:city>
                <gco:CharacterString>Rome</gco:CharacterString>
              </gmd:city>
              <gmd:administrativeArea>
                <gco:CharacterString/>
              </gmd:administrativeArea>
              <gmd:postalCode>
                <gco:CharacterString>00153</gco:CharacterString>
              </gmd:postalCode>
              <gmd:country>
```

```

        <gco: CharacterString>Italy</gco:CharacterString>
      </gmd: country>
      <gmd: electronicMailAddress>
        <gco:CharacterString>jippe.hoogeveen@fao.org</gco:Character-
String>
      </gmd:electronicMailAddress>
    </gmd:CI_Address>
  </gmd:address>
</gmd:CI_Contact>
</gmd:contactInfo>
<gmd:role>
  <gmd:CI_RoleCode codeList="./resources/codeList.xml#CI_RoleCode" codeL-
istValue="pointOfContact"/>
</gmd:role>
</gmd:CI_ResponsibleParty>
</gmd:contact>
<gmd:dateStamp>
  <gco:DateTime>2007-11-06T12:13:00</gco:DateTime>
</gmd:dateStamp>
<gmd:metadataStandardName>
  <gco:CharacterString>ISO 19115:2003/19139</gco:CharacterString>
</gmd:metadataStandardName>
<gmd:metadataStandardVersion>
  <gco:CharacterString>1.0</gco:CharacterString>
</gmd:metadataStandardVersion>
<gmd:spatialRepresentationInfo/>
<gmd:referenceSystemInfo>
  <gmd:MD_ReferenceSystem>
    <gmd:referenceSystemIdentifier>
      <gmd:RS_Identifier>
        <gmd:code>
          <gco:CharacterString>WGS 1984</gco:CharacterString>
        </gmd:code>
      </gmd:RS_Identifier>
    </gmd:referenceSystemIdentifier>
  </gmd:MD_ReferenceSystem>
</gmd:referenceSystemInfo>
<gmd:identificationInfo>
  <gmd:MD_DataIdentification>
    <gmd:citation>
      <gmd:CI_Citation>
        <gmd:title>
          <gco:CharacterString>Hydrological Basins in Africa (Sample record,
please remove!)</gco:CharacterString>
        </gmd:title>
        <gmd:date>
          <gmd:CI_Date>
            <gmd:date>
              <gco:DateTime>2000-07-19T14:45:00</gco:DateTime>
            </gmd:date>
            <gmd:dateType>
              <gmd:CI_DateTypeCode codeList="./resources/codeList.xml#CI_Da-
teTypeCode" codeListValue="creation"/>
            </gmd:dateType>
          </gmd:CI_Date>
        </gmd:date>
        <gmd:edition>
          <gco:CharacterString>First</gco:CharacterString>
        </gmd:edition>
        <gmd:presentationForm>
          <gmd:CI_PresentationFormCode codeList="./resources/codeList.xml#-

```

```

CI_PresentationFormCode" codeListValue="mapDigital"/>
    </gmd:presentationForm>
    </gmd:CI_Citation>
</gmd:citation>
<gmd:abstract>
    <gco:CharacterString>Major hydrological basins and their sub-basins.
This dataset divides the African continent according to its hydrological char-
acteristics.
The dataset consists of the following information:- numerical code and name
of the major basin (MAJ_BAS and MAJ_NAME); - area of the major basin in square
km (MAJ_AREA); - numerical code and name of the sub-basin (SUB_BAS and SUB_
NAME); - area of the sub-basin in square km (SUB_AREA); - numerical code of the
sub-basin towards which the sub-basin flows (TO_SUBBAS) (the codes -888 and -999
have been assigned respectively to internal sub-basins and to sub-basins drain-
ing into the sea)</gco:CharacterString>
    </gmd:abstract>
    <gmd:purpose>
        <gco:CharacterString>This dataset is developed as part of a GIS-based
information system on water resources for the African continent. It has been
published in the framework of the AQUASTAT - programme of the Land and Water
Division of the Food and Agriculture Organization of the United Nations, as
part of FAO Land and Water Digital Media Series #13: "Atlas of Water Resources
and Irrigation in Africa".
For a wider distribution and to support other projects at FAO this map is also
distributed in a DVD as part of a publication entitled: "Jenness, J.; Dooley,
J.; Aguilar-Manjarrez, J.; Riva, C. African Water Resource Database. GIS-based
tools for inland aquatic resource management. 2. Technical manual and workbook.
CIFA Technical Paper. No. 33, Part 2. Rome, FAO. 2007. 308 p."</gco:Character-
String>
    </gmd:purpose>
    <gmd:status>
        <gmd:MD_ProgressCode codeList="./resources/codeList.xml#MD_Progress-
Code" codeListValue="completed"/>
    </gmd:status>
    <gmd:resourceMaintenance>
        <gmd:MD_MaintenanceInformation>
            <gmd:maintenanceAndUpdateFrequency>
                <gmd:MD_MaintenanceFrequencyCode codeList="./resources/codeList.
xml#MD_MaintenanceFrequencyCode" codeListValue="asNeeded"/>
            </gmd:maintenanceAndUpdateFrequency>
        </gmd:MD_MaintenanceInformation>
    </gmd:resourceMaintenance>
    <gmd:graphicOverview>
        <gmd:MD_BrowseGraphic>
            <gmd:fileName>
                <gco:CharacterString>https://localhost:8080/geonetwork/srv/api/
records/da165110-88fd-11da-a88f-000d939bc5d8/attachments/thumbnail_s.gif</
gco:CharacterString>
            </gmd:fileName>
            <gmd:fileDescription>
                <gco:CharacterString>Hydrological Basins in Africa (small pre-
view)</gco:CharacterString>
            </gmd:fileDescription>
        </gmd:MD_BrowseGraphic>
    </gmd:graphicOverview>
    <gmd:graphicOverview>
        <gmd:MD_BrowseGraphic>
            <gmd:fileName>
                <gco:CharacterString>https://localhost:8080/geonetwork/srv/api/re-
cords/da165110-88fd-11da-a88f-000d939bc5d8/attachments/thumbnail.gif</gco:Char-
acterString>

```

```

        </gmd:fileName>
        <gmd:fileDescription>
            <gco:CharacterString>Hydrological Basins in Africa</gco:Character-
String>
        </gmd:fileDescription>
        </gmd:MD_BrowseGraphic>
    </gmd:graphicOverview>
    <gmd:descriptiveKeywords>
        <gmd:MD_Keywords>
            <gmd:keyword>
                <gco:CharacterString>watersheds</gco:CharacterString>
            </gmd:keyword>
            <gmd:keyword>
                <gco:CharacterString>river basins</gco:CharacterString>
            </gmd:keyword>
            <gmd:keyword>
                <gco:CharacterString>water resources</gco:CharacterString>
            </gmd:keyword>
            <gmd:keyword>
                <gco:CharacterString>hydrology</gco:CharacterString>
            </gmd:keyword>
            <gmd:keyword>
                <gco:CharacterString>AQUASTAT</gco:CharacterString>
            </gmd:keyword>
            <gmd:keyword>
                <gco:CharacterString>AWRD</gco:CharacterString>
            </gmd:keyword>
            <gmd:type>
                <gmd:MD_KeywordTypeCode codeList="./resources/codeList.xml#MD_Key-
wordTypeCode" codeListValue="theme"/>
            </gmd:type>
        </gmd:MD_Keywords>
    </gmd:descriptiveKeywords>
    <gmd:descriptiveKeywords>
        <gmd:MD_Keywords>
            <gmd:keyword>
                <gco:CharacterString>Africa</gco:CharacterString>
            </gmd:keyword>
            <gmd:type>
                <gmd:MD_KeywordTypeCode codeList="./resources/codeList.xml#MD_Key-
wordTypeCode" codeListValue="place"/>
            </gmd:type>
        </gmd:MD_Keywords>
    </gmd:descriptiveKeywords>
    <gmd:resourceConstraints>
        <gmd:MD_Constraints>
            <gmd:useLimitation>
                <gco:CharacterString/>
            </gmd:useLimitation>
        </gmd:MD_Constraints>
    </gmd:resourceConstraints>
    <gmd:spatialRepresentationType>
        <gmd:MD_SpatialRepresentationTypeCode codeList="./resources/codeList.
xml#MD_SpatialRepresentationTypeCode" codeListValue="vector"/>
    </gmd:spatialRepresentationType>
    <gmd:spatialResolution>
        <gmd:MD_Resolution>
            <gmd:equivalentScale>
                <gmd:MD_RepresentativeFraction>
                    <gmd:denominator>
                        <gco:Integer>5000000</gco:Integer>

```

```

        </gmd:denominator>
    </gmd:MD_RepresentativeFraction>
    </gmd:equivalentScale>
</gmd:MD_Resolution>
</gmd:spatialResolution>
<gmd:language>
    <gco:CharacterString>eng</gco:CharacterString>
</gmd:language>
<gmd:characterSet>
    <gmd:MD_CharacterSetCode codeList="./resources/codeList.xml#MD_CharacterSetCode" codeListValue="utf8"/>
</gmd:characterSet>
<gmd:topicCategory>
    <gmd:MD_TopicCategoryCode>inlandWaters</gmd:MD_TopicCategoryCode>
</gmd:topicCategory>
<gmd:extent>
    <gmd:EX_Extent>
        <gmd:temporalElement>
            <gmd:EX_TemporalExtent>
                <gmd:extent>
                    <gml:TimePeriod gml:id="timeperiod1">
                        <gml:beginPosition>2006-01-01T04:29:00</gml:beginPosition>
                        <gml:endPosition>2008-01-08T04:29:00</gml:endPosition>
                    </gml:TimePeriod>
                </gmd:extent>
            </gmd:EX_TemporalExtent>
        </gmd:temporalElement>
    </gmd:EX_Extent>
</gmd:extent>
<gmd:extent>
    <gmd:EX_Extent>
        <gmd:geographicElement>
            <gmd:EX_GeographicBoundingBox>
                <gmd:westBoundLongitude>
                    <gco:Decimal>-17.3</gco:Decimal>
                </gmd:westBoundLongitude>
                <gmd:eastBoundLongitude>
                    <gco:Decimal>51.1</gco:Decimal>
                </gmd:eastBoundLongitude>
                <gmd:southBoundLatitude>
                    <gco:Decimal>-34.6</gco:Decimal>
                </gmd:southBoundLatitude>
                <gmd:northBoundLatitude>
                    <gco:Decimal>38.2</gco:Decimal>
                </gmd:northBoundLatitude>
            </gmd:EX_GeographicBoundingBox>
        </gmd:geographicElement>
    </gmd:EX_Extent>
</gmd:extent>
<gmd:supplementalInformation>
    <gco:CharacterString>You can customize the template to suit your needs.
    You can add and remove fields and fill out default information (e.g. contact de-
    tails). Fields you can not change in the default view may be accessible in the
    more comprehensive (and more complex) advanced view. You can even use the XML
    editor to create custom structures, but they have to be validated by the sys-
    tem, so know what you do :-)</gco:CharacterString>
</gmd:supplementalInformation>
</gmd:MD_DataIdentification>
</gmd:identificationInfo>
<gmd:distributionInfo>
    <gmd:MD_Distribution>

```

```

<gmd:distributionFormat>
  <gmd:MD_Format>
    <gmd:name>
      <gco:CharacterString>ShapeFile</gco:CharacterString>
    </gmd:name>
    <gmd:version>
      <gco:CharacterString>Grass Version 6.1</gco:CharacterString>
    </gmd:version>
  </gmd:MD_Format>
</gmd:distributionFormat>
<gmd:transferOptions>
  <gmd:MD_DigitalTransferOptions>
    <gmd:onLine>
      <gmd:CI_OnlineResource>
        <gmd:linkage>
          <gmd:URL>http://www.fao.org/ag/AGL/aglw/aquastat/watresafrica/
index.stm</gmd:URL>
        </gmd:linkage>
        <gmd:protocol>
          <gco:CharacterString>WWW:LINK-1.0-http--link</gco:CharacterS-
tring>
        </gmd:protocol>
        <gmd:name>
          <gco:CharacterString/>
        </gmd:name>
        <gmd:description>
          <gco:CharacterString>Online link to the 'Water Resources and
Irrigation in Africa'- website</gco:CharacterString>
        </gmd:description>
      </gmd:CI_OnlineResource>
    </gmd:onLine>
    <gmd:onLine>
      <gmd:CI_OnlineResource>
        <gmd:linkage>
          <gmd:URL>https://localhost:8080/geonetwork/srv/api/records/
da165110-88fd-11da-a88f-000d939bc5d8/attachments/basins.zip</gmd:URL>
        </gmd:linkage>
        <gmd:protocol>
          <gco:CharacterString>WWW:DOWNLOAD-1.0-http--download</gco:Cha-
racterString>
        </gmd:protocol>
        <gmd:name>
          <gco:CharacterString>basins.zip</gco:CharacterString>
        </gmd:name>
        <gmd:description>
          <gco:CharacterString>Hydrological basins in Africa (Shapefile
Format)</gco:CharacterString>
        </gmd:description>
      </gmd:CI_OnlineResource>
    </gmd:onLine>
    <gmd:onLine>
      <gmd:CI_OnlineResource>
        <gmd:linkage>
          <gmd:URL>http://data.fao.org/maps/wms</gmd:URL>
        </gmd:linkage>
        <gmd:protocol>
          <gco:CharacterString>OGC:WMS</gco:CharacterString>
        </gmd:protocol>
        <gmd:name>
          <gco:CharacterString>GEONETWORK:basins_296</gco:Character-
String>

```

```

        </gmd:name>
        <gmd:description>
            <gco:CharacterString>Hydrological basins in Africa</gco:Charac-
terString>
        </gmd:description>
        </gmd:CI_OnlineResource>
        </gmd:onLine>
        </gmd:MD_DigitalTransferOptions>
        </gmd:transferOptions>
        </gmd:MD_Distribution>
        </gmd:distributionInfo>
        <gmd:dataQualityInfo>
            <gmd:DQ_DataQuality>
                <gmd:scope>
                    <gmd:DQ_Scope>
                        <gmd:level>
                            <gmd:MD_ScopeCode codeList="./resources/codeList.xml#MD_ScopeCode"
codeListValue="dataset"/>
                        </gmd:level>
                    </gmd:DQ_Scope>
                </gmd:scope>
                <gmd:lineage>
                    <gmd:LI_Lineage>
                        <gmd:statement>
                            <gco:CharacterString>The linework of the map is obtained by delin-
eating drainage basin boundaries from an hydrologically corrected digital ele-
vation model with a resolution of 1 * 1 km.</gco:CharacterString>
                        </gmd:statement>
                    </gmd:LI_Lineage>
                </gmd:lineage>
            </gmd:DQ_DataQuality>
        </gmd:dataQualityInfo>
    </gmd:MD_Metadata>

```

5.3.3.2 Disseminação de conhecimentos e novas funcionalidades do Visão junto a potenciais usuários/parceiros

As funcionalidades descritas nas seções anteriores encontram-se em fase de teste para posterior disponibilização em ambiente de produção. Concomitante à fase de validação do que foi desenvolvido no contexto do projeto de pesquisa, estão sendo realizadas ações de disseminação dos resultados obtidos, algumas delas foram:

- Aula na Escola Nacional de Ciências Estatísticas, na disciplina de SIG I, apresentando a plataforma do Visão;
- Reunião com a Rede Nacional de Pesquisas (RNP), com o intuito de promover a utilização do Visão, tanto pela rede quanto por seus parceiros.

5.4 Considerações finais sobre a Meta 4

As atividades descritas anteriormente evidenciam o andamento satisfatório de obtenção de resultados vinculado à Meta 4 do projeto de pesquisa Ecossistema de informação governamental. Os avanços obtidos até março de 2023 entraram em fase de validação para posterior disponibilização e acesso no sistema Visão nos próximos

meses de execução da pesquisa. As atividades e os avanços alcançados até então corroboram para o avanço das questões voltadas ao aprimoramento de uso de tecnologias abertas para dados e informações georreferenciadas no âmbito do governo eletrônico.

6. RESULTADOS META 5

Tão importante quanto a produção do conhecimento é a sua divulgação, de modo que ocorra a contribuição para a construção de uma ciência colaborativa e acumulativa. Comunicar e disseminar resultados técnicos e científicos são atividades fundamentais para a manutenção do ciclo de produção e validação do conhecimento e permitem maior transparência e ampliação de trocas de informações entre os pares e outros interessados nas temáticas e questões pesquisadas.

Na perspectiva de Lara e Conti (2003), disseminar informações supõe tornar pública a produção de conhecimentos gerados ou organizados por uma instituição. O Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), enquanto unidade de pesquisa vinculada ao Ministério da Ciência, Tecnologia e Inovação (MCTI), apoia e incentiva a disseminação e comunicação dos avanços e resultados obtidos a partir de investigações realizadas no âmbito dos projetos de pesquisa coordenados pelo Ibict juntamente com outras instituições parceiras. Dessa forma, o projeto *Ecosistema de Informação Governamental* propôs a Meta 5, voltada para *disseminar resultados de pesquisa*.

O presente relatório tem o intuito de apresentar uma sumarização sobre resultados alcançados, até março de 2023, para as ações da Meta 5 que já foram iniciadas e encontram-se em fase de andamento. Logo, as informações descritas nas próximas seções deste relatório indicam resultados parciais para as atividades de: 1. desenvolvimento de documentação técnica e científica sobre os estudos efetuados nas metas de 1 a 4 do projeto e 2. apresentação dos resultados de estudo em eventos.

No plano de trabalho do projeto de pesquisa em questão, os indicadores de resultado previstos para a Meta 5 englobam os seguintes tipos de produção técnico-científica: relatórios, livros, guias, cartilhas e artigos e apresentações em eventos. O conteúdo do relatório indica e descreve algumas dessas tipologias de disseminação do conhecimento elaboradas no contexto das ações desenvolvidas pelos pesquisadores do projeto *Ecosistema de Informação Governamental*.

6.1 Objetivo geral da Meta 5

Disseminar resultados de pesquisa.

6.2 Objetivo específico da Meta 5

- Desenvolvimento de documentação técnica sobre os estudos
- Desenvolvimento de documentação científica sobre os estudos
- Apresentação dos resultados de estudo em eventos

6.3 Descrição de resultados da Meta 5

As seções subsequentes, seguindo a mesma descrição de atividades previstas no plano de trabalho do projeto de pesquisa Ecosistema de Informação Governamental, apresentam os resultados parciais de disseminação do conhecimento fruto das investigações desenvolvidas até o primeiro trimestre de 2023.

6.3.1 Desenvolvimento de documentação técnica sobre os estudos

6.3.1.1 Livro - Ecosistemas e inovações tecnológicas [recurso eletrônico]: da construção às boas práticas

A obra apresenta como apoio conceitual e aplicado para o projeto de pesquisa em Ecosistema de Informação Governamental, desenvolvido pelo Instituto Brasileiro de Informação em Ciência e Tecnologia, contando com a parceria, no caso dessa publicação, do Portal de Periódicos Eletrônicos Científicos do Sistema de Bibliotecas da Universidade de Campinas (Unicamp). Com isso, contribui para a discussão sobre os diversos sistemas de informação que podem compor ecossistemas. A obra é dividida em duas partes, que se desdobram em 15 capítulos. A parte I traz discussões sobre a teoria e concepção dos sistemas de informação, enquanto, na Parte II, são relatadas as experiências e boas práticas em relação aos ecossistemas apresentados no conteúdo do e-book. Sistemas de informação, aqui, são entendidos como um conjunto que envolve teoria e prática no uso de softwares, nesse caso, de livre e de código aberto. Assim, primeiro é apresentada a base teórica utilizada para uso das ferramentas e, depois, é exibido um caso de uso de sucesso. Para tanto, o livro conta com textos desenvolvidos por pesquisadores reconhecidamente expoentes em suas áreas de atuação, usuários de softwares livres, voltados à gestão da informação. Ao mesmo tempo, os organizadores convidaram pesquisadores oriundos de várias instituições, de forma a enriquecer o panorama a ser exposto. Com isso, a obra apoia o projeto com a base conceitual e prática, apresentando sistemas de informações utilizados por unidades de informação de governo, ou não, em suas atividades direcionadas à gestão da informação. Ademais, colabora, de forma aplicada, com a discussão sobre tecnologias em um mundo no qual ela se faz cada vez mais presente, visto que a informação tem ganhado destaque e precisa de sistemas que apoiem a sua gestão.

A obra está disponível na plataforma de Documentação Técnica da Coordenação Geral de Tecnologias de Informação e Informática⁴, com acesso público.

4 Disponível em: <https://labcotec.ibict.br/omp/index.php/edcotec>.

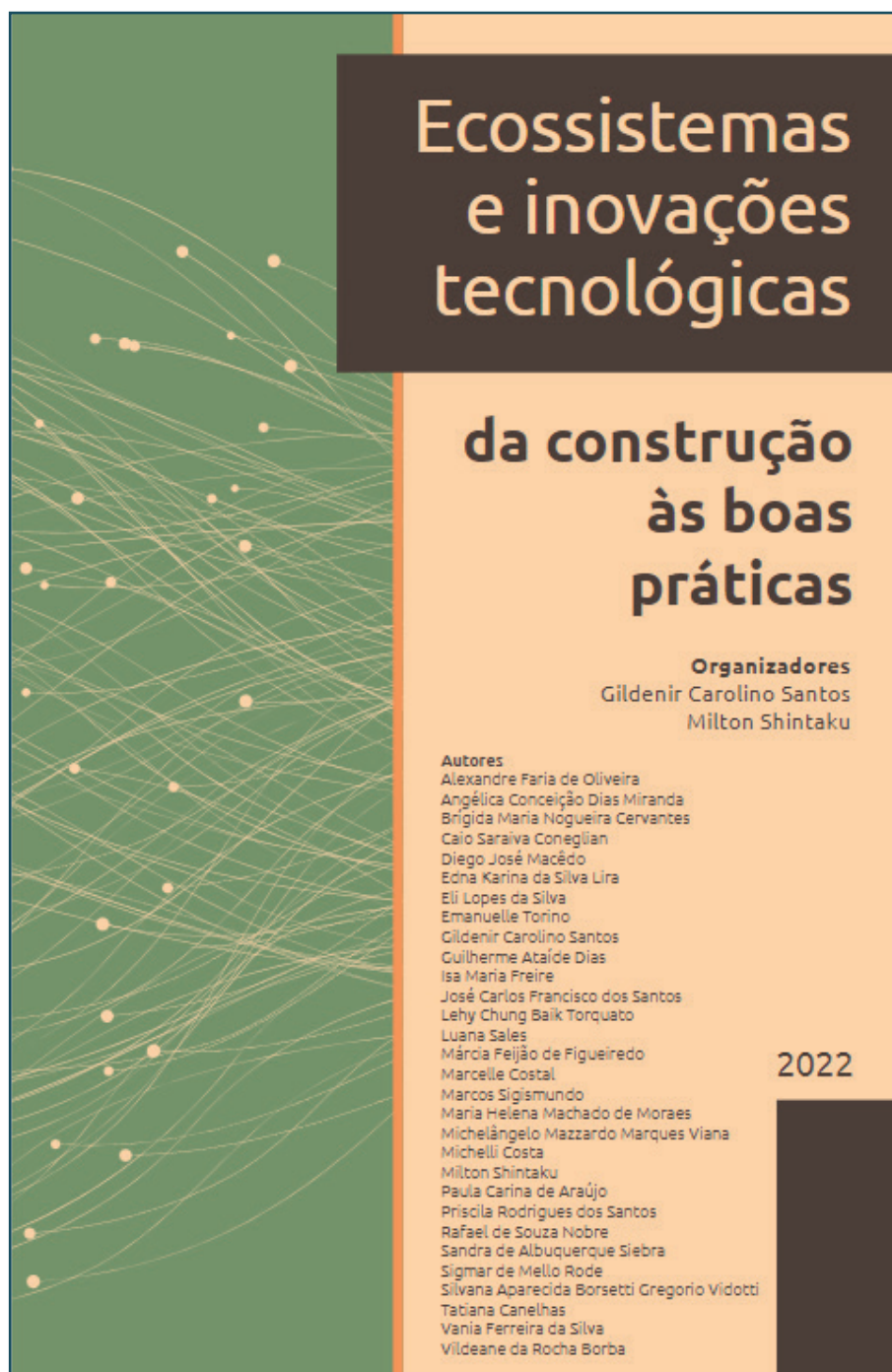


Figura 14 - Capa do livro Ecosistemas e inovações tecnológicas: da construção às boas práticas

Fonte: Santos e Shintaku (2022, capa).

6.3.1.2 Livro - Hipátia: modelo de preservação para repositórios arquivísticos digitais confiáveis

Associada à temática da Meta 2 do projeto de pesquisa Ecossistema de Informação Governamental, que trata sobre os estudos para criação de modelo de preservação de dados e informações governamentais, a obra intitulada *Hipátia: modelo de preservação para repositórios arquivísticos digitais confiáveis* propõe-se a apresentar um pouco da história da preservação digital no âmbito do Instituto Brasileiro de Informação em Ciência e Tecnologia

(Ibict), bem como situar os leitores sobre as propostas conceituais e as estratégias adotadas na consolidação do modelo Hipátia. Em sete capítulos, esse livro apresenta as primeiras percepções quanto à prospecção, o estudo, a proposição e o desenvolvimento do Hipátia, que é um modelo capaz de se adaptar a qualquer sistema que faça gestão de processos e objetos digitais. Temáticas como a preservação de documentos digitais e os Repositórios Arquivísticos Digitais Confiáveis (RDC-Arq) permeiam todas as discussões apresentadas pelos autores que participaram dos processos que resultaram no Modelo Hipátia.



Figura 15 - Capa do livro Hipátia: modelo de preservação para repositórios arquivísticos digitais confiáveis

Fonte: Braga e Mádero Arellano (2022, capa).

Os capítulos que compõem o conteúdo do livro, são:

1. Preservação de documentos em formato digital;
2. O Ibict e a preservação da memória científica;
3. Os principais normativos relacionados ao RDCArq;
4. O modelo Hipátia: a proposta do Ibict para a preservação digital arquivística;
5. Arquitetura do modelo Hipátia;
6. Casos de sucesso de implementação do Modelo RDCArq e Hipátia;
7. A visão do Ibict para um ecossistema nacional de preservação digital.

A obra pode ser acessada e baixada de forma aberta e gratuita na plataforma de Documentação Técnica da Coordenação Geral de Tecnologias de Informação e Informática⁵

6.3.2 Desenvolvimento de documentação científica sobre os estudos

6.3.2.1 Artigo de periódico publicado em periódicos nacionais

SHINTAKU, Milton. Análise da cadeia de valor de órgão de governo como apoio à organização de biblioteca digital. **AtoZ: novas práticas em informação e conhecimento**, [S.l.], v. 12, p. 1 - 11, fev. 2023. ISSN 2237-826X. Disponível em: <https://revistas.ufpr.br/atoz/article/view/84584>. Acesso em: 17 mar. 2023. doi:<http://dx.doi.org/10.5380/atoz.v12i0.84584>.

6.3.3 Apresentação dos resultados de estudo em eventos

6.3.3.1 Conversa com especialistas: Tecnologias de Informação e Informática apoiadas pelo Instituto Brasileiro de Informação em Ciência e Tecnologia

O Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) promoveu o “Conversa com especialistas: Tecnologias de Informação e Informática apoiadas pelo Instituto Brasileiro de Informação em Ciência e Tecnologia”, evento paralelo do XXII Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (Enancib).

5 Disponível em: <https://labcotec.ibict.br/omp/index.php/edcotec/catalog/book/livrohipatia>.

O evento presencial aconteceu no dia 11 de novembro, das 14h às 18h, no campus da Universidade Federal do Rio Grande do Sul (UFRGS).

O objetivo foi auxiliar participantes do ecossistema usuário das tecnologias apoiadas pelo Ibict, visando esclarecer dúvidas e apoiar na solução de problemas por meio de um bate-papo com especialistas, que puderam apresentar soluções e orientações, contribuindo assim, para o melhor uso das diversas ferramentas disponibilizadas pelo Instituto.

Durante o encontro, foram abordados temas e tecnologias como DSpace (repositórios e bibliotecas digitais), Tecnologias *Public Knowledge Project* (PKP), *Open Journal Systems* (OJS), *Open Conference System* (OCS) e *Open Monograph System* (OMP), Tematres (vocabulários controlados), Direitos autorais, Moodle e tecnologias para ensino, WordPress e publicação da informação, Software Visão e a informação geolocalizada, Modelo Hipátia para interação de sistemas para preservação, Plataforma Cívica e a Ciência Cidadã, Koha, Sistema Integrado de Gestão de Biblioteca e Tainacan.

6.3.3.2 Apresentação de trabalhos no evento XXII Encontro Nacional de Pesquisa em Ciência da Informação (Enancib)

BRAGA, Tiago Emmanuel Nunes; ROCHA, Nathaly Cristine Leite. O papel do arquivista na estruturação de propostas informacionais voltadas para a disponibilização de dados e informações. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 22., 2022, Porto Alegre. **Anais** [...]. Porto Alegre: Associação de Pesquisa e Pós-graduação em Ciência da Informação (Ancib), 2022. Disponível em: <https://enancib.ancib.org/index.php/enancib/xxiienancib/paper/view/1003>. Acesso em: 17 mar. 2023.

SOUSA, Rosilene Paiva Marinho de. SHINTAKU, Milton. Informação legislativa sobre inovação tecnológica. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 22., 2022, Porto Alegre. **Anais** [...]. Porto Alegre: Associação de Pesquisa e Pós-graduação em Ciência da Informação (Ancib), 2022. Disponível em: <https://enancib.ancib.org/index.php/enancib/xxiienancib/paper/view/831>. Acesso em: 17 mar. 2023.

6.3.3.3 Apresentação de trabalhos no evento XIII Encuentro EDICIC

NAKANO, Natália; SANTOS, Flávia Karla Ribeiro; SHINTAKU, Milton. Elementos fundamentales para crear correos electrónicos accesibles. **Revista EDICIC**, [S. l.], v. 2, n. 4, 2022. Disponível em: <https://ojs.edicic.org/index.php/revistaedicic/article/view/120>. Acesso em: 17 mar. 2023.

SOUSA, Rosilene Paiva Marinho de; SHINTAKU, Milton; NAKANO, Natália; VIDOTTI, Silvana Aparecida Borsetti Gregório. Elementos de apoyo legal sobre accesibilidad digital para personas sordas en Brasil. **Revista EDICIC**, [S. l.], v. 2, n. 4, 2022. Disponível em: <https://ojs.edicic.org/index.php/revistaedicic/article/view/118>. Acesso em: 17 mar. 2023.

6.3.3.4 Apresentação de trabalhos no evento Workshop de Informação, Dados e Tecnologia (V WIDaT)

MELO, Ívina Flores; PIGNATARO, Tatiana Canelhas; BRAGA, Tiago Emmanuel Nunes. Modelo de preservação hipatia: metodologia de estudo de metadados para extração. In: WORKSHOP DE INFORMAÇÃO, DADOS E TECNOLOGIA. **Anais** [...], 2022. Vitória: Universidade Federal do Espírito Santo, 2022. v. 1. p. 147-152. Disponível em: <https://widat2022.ufes.br/wp-content/uploads/2023/04/widat-2022-anais.pdf>. Acesso em: 17 mar. 2023.

6.3.3.5 Apresentação de trabalho no evento Congresso Nacional de Arquivologia (CNA 2022)

PIGNATARO, Tatiana Canelhas; BRAGA, Tiago Emmanuel Nunes. Modelo Hipátia: análise dos requisitos arquivísticos e computacionais. In: CONGRESSO NACIONAL DE ARQUIVOLOGIA, 9. 2022, Florianópolis. **Anais** [...]. Florianópolis: AAESC, 2022. v. 1. p. 43-45.

6.3.4 Reconhecimento pelo Selo Nacional de Modernização do Estado - 2022

Sistema Aberto de Observatórios para Visualização de Informações (VISÃO) e o modelo de preservação Hipátia receberam o Selo Nacional de Modernização do Estado, ambos serviços mantidos pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) e desenvolvidos em parceria com outras instituições a partir de projetos de pesquisa. O objetivo do prêmio, que foi entregue em 12 de dezembro no Palácio do Planalto, é identificar, certificar e incentivar as iniciativas de modernização em âmbito nacional, seguindo premissas da Política Nacional de Modernização do Estado - Moderniza Brasil, cuja finalidade é direcionar esforços governamentais para aumentar a eficiência e modernizar a administração pública, a prestação de serviços e o ambiente de negócios para melhor atender às necessidades dos cidadãos (GUEDES, 2022).

Abaixo, encontram-se apresentados certificados dos Selos de Modernização recebidos pelo Ibict:



Figura 16 - Selo Prata de Modernização do Estado - Modelo Hipátia



Figura 17 - Selo Bronze de Modernização do Estado - VISÃO

6.4 Considerações finais sobre a Meta 5

Tendo em vista os resultados apresentados no decorrer do presente relatório, nota-se que a Meta 5 de disseminação do conhecimento, fruto de pesquisa realizada no âmbito do projeto de pesquisa Ecossistema de Informação Governamental, ainda em execução, apresenta andamento satisfatório com relação à produção de registros documentais da pesquisa (artigos, livros, resumos expandidos) e disseminação de descobertas por meio da participação dos pesquisadores em apresentações realizadas em eventos científicos. O presente relatório indicou somente as produções consolidadas, todavia, faz-se válido ressaltar que existem livros e outras tipologias de disseminação de resultados previstas para início e em fase de elaboração, tendo como norte para finalização o período de vigência do projeto de pesquisa.

7. CONSIDERAÇÕES FINAIS

Diante dos resultados das cinco metas vinculadas e sumarizadas no presente relatório parcial, torna-se evidente que a execução do projeto de pesquisa *Ecossistema de Informação Governamental*, durante o período de realização, alcançou relevantes frutos no que se refere: à proposição de modelos de integração de sistemas informacionais governamentais; ao modelo de preservação digital para dados e informações de caráter governamental; ao modelo para criação e conversão de documento digitais de forma acessível; ao desenvolvimento de estudos para aprimoramento de possibilidades de apresentação de dados e informações geolocalizadas e; à disseminação de resultados, conforme previsto pelo objetivo geral definido para a investigação.

O andamento do projeto demonstra resultados que corroboram para uma proposição mais factível de um contexto integrado e estruturado no que concerne às informações e aos dados governamentais. Isso posto, e com os avanços da pesquisa, é possível afirmar que as proposições previstas inicialmente, a partir do plano de trabalho do projeto, foram alcançadas por meio da execução normal de todas as metas e atividades previstas até então.

REFERÊNCIAS

ABDF. PORTAL NORMAS.LEG.BR. Uma nova experiência na consulta à legislação. **Youtube**, 2021. Disponível em: <https://www.youtube.com/watch?v=yh3LiTvPxxM>. Acesso em: 7 mar. 2022.

ALAM, Sawood; WEIGLE, Michele C.; NELSON, Michael L.; KLEIN, Martin; SOMPEL, Herbert Van de. Supporting web archiving via web packaging. **ArXiv**:1906.07104 [cs][open-access archive], 2019. Disponível em: <http://arxiv.org/abs/1906.07104>. Acesso em: 7 mar. 2022.

ALBAGLI, Sarita; PY, Hesley; IWAMA, Allan Yu. Geovisualização de dados e Ciência aberta e cidadã-a experiência da Plataforma LindaGeo. **Digital Humanities Quarterly**, v. 14, n. 2, 2020.

ALBANO, Cláudio Sonáglio. Open government data: a value chain model proposal. *In*: DG.O 2013: 14TH ANNUAL INTERNATIONAL CONFERENCE ON DIGITAL GOVERNMENT RESEARCH, 17 jun. 2013. **Proceedings** [...]. Quebec Canada: ACM, 17 jun. 2013. p. 285–286. DOI 10.1145/2479724.2479775. Disponível em: <https://dl.acm.org/doi/10.1145/2479724.2479775>. Acesso em: 17 mar. 2023.

ALPERIN, Juan Pablo; WILLINSKY, John; OWEN, Brian; MACGREGOR, James; SMECHER, Alec; STRANACK, Kevin. **The Public Knowledge Project: reflections and directions after two decades**. [S. l.]: PKP; Libraries at University of Nebraska-Lincoln, 2018(Copyright, Fair Use, Scholarly Communication, etc., 73). Disponível em: <https://digitalcommons.unl.edu/scholcom/73/>. Acesso em: 17 mar. 2023.

AMERICAN NATIONAL STANDARDS INSTITUTE. **ANSI/NISO Z39.19-2005 (R2010)**: guidelines for the construction, format, and management of monolingual controlled vocabularies. Baltimore, Maryland: ANSI, 2010.

ANDRADE DA LUZ, Ana Javes. **Comunicação pública e memória comunicacional: revelações e apagamentos sobre o governo da presidenta Dilma Rousseff**. 2021, 253f. Disponível em: <https://lume.ufrgs.br/handle/10183/235405>. Acesso em 24 fev. 2022.

ANDRADE, Morgana Carneiro de; CERVANTES, Brígida Maria Nogueira. A contribuição da organização do conhecimento para a interoperabilidade semântica: alternativas para repositórios institucionais. **Informação@Profissões**, v. 1, n. 1/2, p. 151-169, 2012. DOI 10.5433/2317-4390.2012v1n1/2p151. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/infoprof/article/view/14593>. Acesso em: 17 mar. 2023.

ANDRADE, Morgana Carneiro de; SHINTAKU, Milton; BARROS, Patricia Pacheco de. Proposta de elementos de metadados para representação e recuperação de memória técnica: o caso da Rede Ufes-Rio Doce. **Cader-**

nos BAD, n. 1, p. 41–58, 2018. Disponível em: <https://www.bad.pt/publicacoes/index.php/cadernos/article/view/1918>. Acesso em: 17 mar. 2020.

ANGUES, I. A importância do modelo OAIS enquanto meio de preservação digital. 25 fev. 2019. **Dados Pesquisa Abertos**, [2021?]. Disponível em: <https://dadosdepesquisa.rnp.br/?p=358>. Acesso em: 7 jan. 2022.

ARQUIVO NACIONAL DO BRASIL. **Política de preservação digital**. Versão 2. Dezembro de 2016. Disponível em: http://www.siga.arquivonacional.gov.br/images/an_digital/and_politica_preservacao_digital_v2.pdf. Acesso em: 05 jan. 2022.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR 16.452**: Acessibilidade na comunicação: audio-descrição: referências. Rio de Janeiro: ABNT, 2016. Disponível em: <https://www.prefeitura.sp.gov.br/cidade/secretarias/upload/ABNT%20-%20Acessibilidade.pdf>. Acesso em: 03 mar. 2023

BASEVI, Teresa. BDJur Consortium - Juridical Digital Library: implementing DSpace in the Brazilian Judiciary. In: ELPUB2005 CONFERENCE ON ELECTRONIC PUBLISHING, 9., 2005, Leuven, Belgium. **Proceedings** [...]. Leuven: Katholieke Universiteit de Leuven, 2005. p. 127- 132. Disponível em: <https://elpub.architexturez.net/system/files/pdf/150elpub2005.con-tent.pdf>. Acesso em: 17 mar. 2023.

BOERES, Sonia Araújo de Assis. **Competências necessárias para equipes de profissionais de preservação digital**. 2017. xvi, 293 f., il. Tese (Doutorado em Ciência da Informação) —Universidade de Brasília, Brasília, 2017. Disponível em: <https://repositorio.unb.br/handle/10482/24354>. Acesso em 3 mar. 2022.

BRAGA, Tiago Emmanuel Nunes; MÁRDERO ARELLANO, Miguel Ángel. **Hipátia**: modelo de preservação para repositórios arquivísticos digitais confiáveis. Brasília: Ibict, 2022. DOI 10.22477/9786589167501.

BRAGG, Molly; HANNA, Kristine. **The Web Archiving Life Cycle Model**. WhitePaper. 2013. Disponível em: http://ait.blog.archive.org/files/2014/04/archiveit_life_cycle_model.pdf. Acesso em: 14 out. 2021.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Brasília, DF. Senado Federal: Centro Gráfico, 1988.

BRASIL. Câmara dos Deputados. **Projeto de Lei nº 2431/2015**. Dispõe sobre o patrimônio público digital institucional inserido na rede mundial de computadores e dá outras providências. Brasília, DF: Câmara dos Deputados, 2015. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2255564>. Acesso em: 9 nov. 2021.

BRASIL. Câmara dos Deputados. **Projeto de Lei nº 3050/2020**. Altera o Marco Civil da Internet para proibir a retirada de conteúdos pelas aplicações de internet nos casos em que especifica. Brasília, DF: Câmara dos Deputados, 2020a. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2256518>. Acesso em: 9 nov. 2021.

BRASIL. Câmara dos Deputados. **Projeto de Lei nº 3051/2020**. Acrescenta o art. 10-A à Lei nº 12.965, de 23 de abril de 2014 (Marco Civil da Internet), a fim de dispor sobre a destinação das contas de aplicações de internet após a morte de seu titular. Brasília, DF: Câmara dos Deputados, 2020b. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2254248>. Acesso em: 9 nov. 2021.

BRASIL. Câmara dos Deputados. **Projeto de Lei nº 3395/2020**. Acrescenta o art. 21-A à Lei nº 12.965, de 23 de abril de 2014 (Marco Civil da Internet), proibindo os provedores de plataformas digitais de remover conteúdos publicados por seus usuários, salvo por força de cumprimento de ordem judicial. Brasília, DF: Câmara dos Deputados, 2020c. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2255564>. Acesso em: 9 nov. 2021.

BRASIL. **Lei nº 13.146, de 6 de julho de 2015**. Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência). Brasília: Presidência da República, 2015. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm. Acesso em: 17 mar. 2023.

BRASIL. Câmara dos Deputados. **PL 3573/2020**. Luiz Philippe de Orleans e Bragança. Altera o Marco Civil da Internet para proibir a retirada de conteúdos pelas aplicações de internet nos casos em que especifica. Disponível em: https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra;jsessionid=node0168xurty2eao-450t98qvmahcp1339305.node0?codteor=1908998&filename=PL+3573/2020. Acesso em: 9 nov. 2021.

BRASIL. **Decreto nº 7.724, de 16 de maio de 2012**. Regulamenta a Lei nº 12.527, de 18 de novembro de 2011, que dispõe sobre o acesso a informações previsto no inciso XXXIII do caput do art. 5, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição. Brasília, DF: Presidência da República, 2012. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/decreto/d7724.htm#art76. Acesso em: 9 nov. 2021.

BRASIL. **Lei nº 12.527**, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal [...]. Brasília: Presidência da República, 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. Acesso em: 9 nov. 2021.

BRASIL. **Constituição** (1988). Constituição da República Federativa do Brasil. Brasília, DF: Senado Federal: Centro Gráfico, 1988.

BRASIL. **Lei nº 8.539, de 09 de outubro de 2015**. Dispõe sobre o uso do meio eletrônico para a realização do processo administrativo no âmbito dos órgãos e das entidades da administração pública federal direta, autárquica e fundacional. Disponível em: <https://legislacao.presidencia.gov.br/atos/?tipo=DEC&numero=8539&ano=2015&ato=deegXVE50dZpWT710>. Acesso em: 14 jan. 2022.

BRASIL. **Lei nº 12.965, de 23 de abril de 2014**. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm. Acesso em: 5 nov. 2021.

BRASIL. **Lei nº 12.682, de 9 de julho de 2012**. Dispõe sobre a elaboração e o arquivamento de documentos em meios eletromagnéticos. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12682.htm. Acesso em: 5 nov. 2021.

BRASIL. Lei nº 13.709/2018, **dispõe sobre a proteção de dados pessoais**. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 5 nov. 2021.

BRASIL. Presidência da República Casa Civil Subchefia para Assuntos Jurídicos **LEI Nº 9.610**, DE 19 de fevereiro de 1998. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l9610.htm. Acesso em 22 fev. 2022.

BRASIL. Presidência da República. Casa Civil. Subchefia para Assuntos Jurídicos. **Lei nº 10.994**, de 14 de dezembro de 2004. Dispõe sobre o depósito legal de publicações, na Biblioteca Nacional, e dá outras providências. Disponível em: https://www.youtube.com/watch?v=ZanQR9-WR_E. Acesso em 22 fev. 2022.

BRAYNER, Aquiles. **Preservação do Conteúdo Web Brasileiro**. Mesa-redonda apresentada no 9º Fórum da Internet do Brasil. Manaus, AM. [S.l.: s.n.], 2 outubro de 2019. 1 vídeo (1h 20 min). Publicado pelo canal NICbrvdeos. Disponível em: https://www.youtube.com/watch?v=ZanQR9-WR_E. Acesso em: 7 fev. 2022.

CARPES, Daiana Stockey.; SOSTER, Demétrio de Azeredo. **Manual de audiodescrição para produtos jornalísticos laboratoriais impressos**. Santa Cruz do Sul: Catarse, 2016. Disponível em: https://editoracatarse.com.br/site/wp-content/uploads/2016/11/Manual_de_audiodescrição_para_produtos_jornalísticos_laboratoriais_impressos-1.pdf. Acesso em: 03 mar. 2023.

CARRETA, Antonio Paulo. **Preservação do Conteúdo Web Brasileiro**. Mesa-redonda apresentada no 9º Fórum da Internet do Brasil. Manaus, AM. [S.l.: s.n.], 2 outubro de 2019. 1 vídeo (1h 20 min). Publicado pelo canal NICbrvdeos. Disponível em: https://www.youtube.com/watch?v=ZanQR9-WR_E. Acesso em: 7 fev. 2022.

COSTA, Carla Medeiros da. **Análise das políticas de arquivamento em repositórios institucionais de universidades federais brasileiras**. Orientador: Sonia Aguiar Cruz Riascos. 102 f.: il., fig. Trabalho de Conclusão de Curso (Graduação) - Centro de Artes e Comunicação, Ciência da Informação, Universidade Federal de Pernambuco, Recife. 2015. Disponível em: <https://repositorio.ufpe.br/bitstream/123456789/34905/1/Carla%20Medeiros%20da%20Costa.pdf>. Acesso em: 23 jan. 2022.

CUNNEA, P. *et al.* The influence of legal deposit legislation on the digital collections of the national library of Scotland. In: GOODING, P.; TERRAS, M. (Ed.). **Electronic Legal Deposit: shaping the library collections of the future**. Londres: Facet Publishing, p. 33-53, 2020.

DANKO, David M. Iso/tc211: geographic information-metadata iso 19115. In: **World Spatial Metadata Standards**. Elsevier Science, 2005. p. 535-555.

DAVIS, C. Archiving the web: A case study from the university of victoria. **Code4Lib Journal**, no. 26, 2014. Disponível em: <https://journal.code4lib.org/articles/10015/comment-page-1>. Acesso em: 28 nov. 2021.

DA INFRAESTRUTURA, CINDE-Comitê de Planejamento et al. Plano de Ação para Implantação da Infraestrutura Nacional de Dados Espaciais. 2010.

DAY, Michael. Preserving the fabric of our lives: A survey of web preservation initiatives. In: **Research and Advanced Technology for Digital Libraries: 7th European Conference, ECDL 2003 Trondheim, Norway, August 17-22, 2003 Proceedings 7**. Springer Berlin Heidelberg, 2003. p. 461-472. Disponível em:

DE METADADOS GEOESPACIAIS, Comitê de Estruturação. Perfil de metadados geoespaciais do Brasil: perfil MGB. **Rio de Janeiro: CEMG-Concar**, 2009.

EUROPEAN UNION. Guidelines to make archivable websites. **Publications Office of The European Union**. Disponível em: <https://op.europa.eu/en/web/web-tools/guidelines-to-make-archivable-websites>. Acesso em: out. 2021.

EU WEB ARCHIVE. **Publications Office of The European Union. EU Web Archive**, [2021?]. Disponível em: <https://op.europa.eu/en/web/web-tools/euwebarchive>. Acesso em: out. 2022.

FARRELL, M., *et al.* **Web Archiving in the United States**: a 2017 survey: an NDSA report. [s.l.]: National Digital Stewardship Alliance, 2018. DOI: 10.17605/OSF.IO/3QH6N. Disponível em: <https://osf.io/ht6ay/>. Acesso em: 8 mar. 2021.

FERREIRA, Lisiane Braga. **Arquivamento da web e mídias sociais: preservação digital de vídeos da campanha presidencial brasileira de 2018**. Dissertação (Mestrado em Comunicação e Informação) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019. Disponível em: <https://lume.ufrgs.br/handle/10183/194617>. Acesso em: 3 mar. 2022.

FLORES, D. Preservação Digital Sistêmica orientada pelo Modelo OAIS: Gestão, Preservação, Curadoria Digital e Transparência Ativa. [Palestra Slides] Material elaborado para Palestra na IX **Conferência Internacional sobre Bibliotecas e Repositórios Digitais da América Latina**. UNIFESP: São Paulo, SP; 2019. Disponível em: <http://documentosdigitais.blogspot.com>. Acesso em: 1 ago. 2021.

FORTINI, Cristiana; SHERMAM, Ariane. Governança pública e combate à corrupção: novas perspectivas para o controle da Administração Pública brasileira. **Int. Públ. – IP**, v. 19, n. 102, p. 27–44, abr. 2017. Disponível em: <https://www.editoraforum.com.br/wp-content/uploads/2017/05/governanca-combate-corrupcao.pdf>. Acesso em: 17 mar. 2023.

GOMES, D. **Preservação do Conteúdo Web Brasileiro**. Mesa-redonda apresentada no 9º Fórum da Internet do Brasil. Manaus, AM. [S.l.: s.n.], 2 outubro de 2019. 1 vídeo (1h 20 min). Publicado pelo canal NICbrvideos. Disponível em: (106) [9º FórumBR] Programação Salão D - 2/10 - Tarde - YouTube . Acesso em: 7 fev. 2022.

GOMES, D. Preservar a Web: um desafio ao alcance de todos. In: CONGRESSO NACIONAL DE BIBLIOTECÁRIOS, ARQUIVISTAS E DOCUMENTALISTAS, 10., 2010, Guimarães, Portugal, Políticas de Informação na Sociedade em Rede. **Actas...** Disponível em: <https://publicacoes.bad.pt/revistas/index.php/congressosbad/article/view/158>. Acesso em: 8 dez. 2021.

GOMES, Daniel; MIRANDA, João; COSTA, Miguel. A survey on web archiving initiatives. In: **International Conference on Theory and Practice of Digital Libraries**, 15.: 2011: Berlin. [Proceedings of...]. Berlin: Springer-Verlag, 2011. p. 408-420.

GONZÁLEZ DE GÓMEZ, Maria Nélide; MACHADO, Rejane. A ciência invisível: o papel dos relatórios e as questões de acesso à informação científica. **DataGramZero - Revista de Ciência da Informação**, v. 8, n. 5, out. 2007. Disponível em: <http://ridi.ibict.br/handle/123456789/119>. Acesso em: 17 mar. 2023.

GUEDES, Lucas. **Ibict é premiado com Selo Nacional de Modernização do Estado**: O Sistema Aberto de Observatórios para Visualização de Informações (VISÃO) e o modelo de preservação Hipátia receberam Selo Nacional de Modernização do Estado. 12 dez. 2022. Disponível em: <https://www.gov.br/ibict/pt-br/central-de-conteudos/noticias-ibict/dezembro/ibict-e-premiado-com-selo-nacional-de-modernizacao-do-estado>. Acesso em: 23 mar. 2022.

HERITRIX. Heritrix Home page. **Internet Archive**, 2022. Disponível em: <http://crawler.archive.org/index.html>. Acesso em: 15 jan. 2022.

HIPÁTIA. Pesquisa aplicadas. **Hipátia**, 2023?. Disponível em: <https://hipatia.ibict.br/>. Acesso em: 7 jan. 2023.

HOCKX-YU, H. **10 years of archiving the UK web**. British Library. [2013?]. Slide Presentation. 42 f. Disponível em: https://www.dcc.ac.uk/sites/default/files/documents/IDCC15/Presentations%20Day%20IDCC_HHY_Finalv2.pdf. Acesso em: 9 out. 2021.

HOCKX-YU, Helen. How good is good enough? – quality assurance of harvested web resources, 2012. Disponível em: <https://blogs.bl.uk/webarchive/2012/10/how-good-is-good-enough-quality-assurance-of-harvested-web-resources.html>. Acesso em: 07 out. 2021.

IDEA AND CONCEPT. [2013]. Disponível em: <http://mementoweb.github.io/SiteStory>. Acesso em: 7 dez. 2021.

IIPC MEMBERS. IIPC Member Organizations. **IIPC netpreserv.org** [2017]. Disponível em: <http://netpreserve.org/about-us/members>. Acesso em 02 abr. 2017.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA: Livro Aberto: portal do livro aberto em CT&I. **Portal do Livro Aberto**, [2021?]. Disponível em: <https://livroaberto.ibict.br/>. Acesso em: 7 dez. 2021.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. Rede Cariniana. [2022]? **Linhas de Pesquisa e Grupos de Estudo**. Disponível em: <https://cariniana.ibict.br/index.php/institucional/pre-dig>. Acesso em: 7 dez. 2021.

INSTITUTO BRASILEIRO DE GEOGRAFIA. **Website**, 2018. Pesquisa Nacional por amostra de domicílio. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/populacao/9171-pesquisa-nacional-por-amostra-de-domicilios-continua-mensal.html?=&t=o-que-e>. Acesso em: 25 nov. 2022.

INTERNATIONAL CONFERENCE ON THEORY AND PRACTICE OF DIGITAL

LIBRARIES, 7.: 2003: Berlin. [Proceedings of...]. Berlin: Springer-Verlag, 2003.

INTERNATIONAL ORGANIZATION OF STANDARDIZATION. **ISO 25964-1**: information and documentation: thesauri and interoperability with other vocabularies: part 1: thesauri for information retrieval. Geneva: ISO, 2011. Disponível em: <https://www.iso.org/standard/53657.html>. Acesso em: 05 jul. 2022.

INTERNATIONAL ORGANIZATION OF STANDARDIZATION. **ISO 25964-2**: information and documentation: thesauri and interoperability with other vocabularies: part 2: interoperability with other vocabularies. Geneva: ISO, 2013. Disponível em: <https://www.iso.org/standard/53658.html>. Acesso em: 4 ago. 2022.

INTERNATIONAL INTERNET PRESERVATION CONSORTIUM. Web Archiving [site]. **Netpreserv.org** [site]. 2021. Disponível em: <https://netpreserve.org/web-archiving/>. Acesso em: 7 dez. 2021.

INTERNET ARCHIVE. **Nutchwax** [site]. [2021?]. Disponível em: <http://archiveaccess.sourceforge.net/>. Acesso em 3 nov. 2007.

KHAN, Muzammil; RAHMAN, Arif Ur. A systematic approach towards web preservation. *Information Technology and Libraries*, v. 38, n. 1, p. 71-90. 2019. Disponível em: <https://doi.org/10.6017/ital.v38i1.10181>. Acesso em: 20 nov. 2021.

LARA, Marilda Lopes Ginez de; CONTI, Vivaldo Luiz. Disseminação da informação e usuários. **São Paulo Perspec.**, São Paulo, v. 17, n. 3-4, dez. 2003. Disponível em: <https://doi.org/10.1590/S0102-88392003000300004>. Acesso em: 16 mar. 2023.

LEVY, D. C., *et al.* **Using a web-archiving service**: how to ensure your cited web-references remain available and valid. [online]. The University of Sidney, 2017. Disponível em: <http://hdl.handle.net/2123/17210>. Acesso em: 12 out. 2021.

LIBRARY OF CONGRESS COLLECTION DEVELOPMENT OFFICE. **Library of Congress collections policy statements supplementary guidelines**: Web archiving, 2017. Disponível em: <http://www.loc.gov/acq/devpol/webarchive.pdf>. Acesso em: 7 mar. 2021.

LIMA, Davi; LIMA, João; KRAUSS, Peter. **LexML**: Portal do projeto LexML. [2022?]. Disponível em: <https://projeto.lexml.gov.br/documentacao/destaques-lexml#o-que-lexml>. Acesso em: 7 dez. 2022.

LIST of Web archiving initiatives. In: WIKIPÉDIA: a enciclopédia livre. Wikimedia, [2021?]. Disponível em: https://en.wikipedia.org/wiki/List_of_Web_archiving_initiatives. Acesso em: 7 dez. 2021.

LÓSCIO, Bernadette Farias; BURLE, Caroline; CALEGARI, Newton. Implementation of the Best Practices for Data on the Web in Brazil and Costa Rica. IN: **WWW '19: Companion Proceedings of The 2019 World Wide Web Conference** May 2019 Pages 807–810. **REVISTA ACM**. Disponível em: https://dl.acm.org/doi/pdf/10.1145/3308560.3316463?casa_token=SEiR7terfUkAAAAA:vBVHVLKaDKxnl-GWMR96DtE8dyfi6Bi-TWIXS8NYYtziQuWBa8YkFMBkgoaNqTJ4aNLnPKpilrI31. Acesso em 24 fev. 2022.

MÁRDERO ARELLANO, Miguel Angel; SANTOS, Gildenir Carolino (org.); SILVA, Ana Paula Araújo Cabral da (colab.) *et al.* **Bibliografia sobre preservação digital** [recurso eletrônico]: um levantamento nos diversos suportes informacionais. Campinas, SP : BCCL/UNICAMP, 2021. E-book (Coleção Preservação digital; v.1). DOI: <https://doi.org/10.20396/ISBN9786588816110>.

MARTINS, Marina Rodrigues. **Mapeamento de públicos estratégicos para iniciativas brasileiras universitárias de arquivamento da web no âmbito acadêmico**: uma projeção para Universidade Federal do Rio Grande do Sul e seu Programa de Pós-graduação em Comunicação. 2019. Dissertação (Mestrado em Comunicação e Informação) - Programa de Pós-Graduação em Comunicação e Informação UFRGS. Porto Alegre, 2019. Disponível em: <https://lume.ufrgs.br/handle/10183/189170>. Acesso em 24 fev. 2022.

MELO, Jonas Ferrigolo; ROCKEMBACH, Moisés. Arquivabilidade de websites para preservação digital: estudo a partir da área da saúde. **RECIIS - Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, Rio de Janeiro, v. 14, n. 3, p. 529-545, jul./set. 2020. Disponível em: <https://www.arca.fiocruz.br/handle/icict/43726>. Acesso em 20 fev. 2022.

MELO, J. F. **Arquivamento dos websites do governo federal brasileiro**: preservação do domínio gov.br. 2020. 133 f. Dissertação (Mestrado – Programa de Pós-Graduação em Comunicação e Informação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2020. Disponível em: <http://hdl.handle.net/10183/210671>. Acesso em: 20 set. 2021.

MELO, J. F.; NUNES, L. A. N. de O.; ROCKEMBACH, M. Preservação de websites governamentais a partir do arquivamento da web: abordagens e metodologias. In: **XX Encontro nacional de pesquisa em ciência da informação – ENANCIB**, GT-8 – Informação e Tecnologia, Florianópolis, SC, 2019. ISSN 2177-3688. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/202684/001106575.pdf?sequence=1>. Acesso em: 7 out. 2021.

MINOW, Mary. **Digital preservation and copyright by Peter Hirtle**. 2003. Disponível em: https://fairuse.stanford.edu/2003/11/10/digital_preservation_and_copyr/. Acesso em: 06 jan. 2022.

MODELO DE PRESERVAÇÃO DIGITAL HIPÁTIA. [S.l.: s.n.], 2021. 1 vídeo (25 min). Publicado pelo canal Rede Cariniana. Disponível em: <https://www.youtube.com/watch?v=YlfRmstG59U>. Acesso em: 7 mar. 2022.

MOMENTO. About the Memento Project. **Momento**, 2022. Disponível em: <http://mementoweb.org/about/>. Acesso em: 3 jan. 2022.

MORESI, Eduardo Amadeu Dutra. Delineando o valor do sistema de informação de uma organização. **Ciência da Informação**, v. 29, n. 1, p. 14–24, abr. 2000. Disponível em: <https://www.scielo.br/j/ci/a/pzj7MLqjc6jX-5zHLxH5PFwq/?format=pdf&lang=pt>. Acesso em: 17 mar. 2023.

NAKANO, Natalia. Acessibilidade web no ensino a distância na ciência da informação: uma revisão sistemática da literatura brasileira na BRAPCI. **AtoZ: novas práticas em informação e conhecimento**, [S.l.], v. 11, p. 1 - 13, jun. 2022. ISSN 2237-826X. Disponível em: <https://revistas.ufpr.br/atoz/article/view/81992>. Acesso em: 17 mar. 2023. doi:<http://dx.doi.org/10.5380/atoz.v11i0.81992>.

NAKANO, Natália; SANTOS, Flávia Karla Ribeiro; SHINTAKU, Milton. Elementos fundamentais para crear correos electrónicos accesibles. **Revista EDICIC**, [S. l.], v. 2, n. 4, 2022. Disponível em: <https://ojs.edicic.org/index.php/revistaedicic/article/view/120>. Acesso em: 17 mar. 2023.

NORMAS.LEG.BR. **As normas federais na linha do tempo**, [2022?]. Disponível em: <https://normas.leg.br/busca>. Acesso em: 7 mar. 2022.

ORGANISERS #EWA Conference. **Book of Abstracts: #EWAVirtual 2020**. [s.l.]: Zenodo, 2020. Disponível em: <https://zenodo.org/record/4058013>. Acesso em: 7 out. 2021.

OWENS, T. **What do you mean by archive?** genres of usage for digital preservers. 2014. Disponível em: <http://blogs.loc.gov/digitalpreservation/2014/02/what-do-you-mean-by-archive-genres-of-usage-for-digital-preservers/>. Acesso em: 7 out. 2021.

Preservação digital. Rede Cariniana [site]. [2022?]. Disponível em: <https://cariniana.ibict.br/index.php/institucional/pre-dig>. Acesso em: 7 dez. 2021.

PUBLIC KNOWLEDGE PROJECT. Disponível em: <http://pkp.sfu.ca>. Acesso em: 13 mar. 2022.

PUBLIC KNOWLEDGE PROJECT (PKP). Disponível em: <https://docs.pkp.sfu.ca/>. Acesso em: 17 mar. 2023.

PY, Hesley. Visualização por meio de mapas como recurso para a democratização dos dados e para a apropriação social da informação. *In*: ENCONTRO NACIONAL DE GOVERNO ABERTO E CIÊNCIA ABERTA DA FIOCRUZ, 1., 2019. Brasília, DF: Fiocruz/VPEIC, 2019. 13 p. Disponível em: <https://www.arca.fiocruz.br/handle/icict/37625>. Acesso em: 13 mar. 2023.

REED, Carl N. The open geospatial consortium and web services standards. In: **Geospatial Web Services: Advances in Information Interoperability**. IGI Global, 2011. p. 1-16.

RILEY, H.; CROOKSTON, M. **Awareness and use of the New Zealand web archive**: a survey of New Zealand academics. [online]. National Library of New Zealand, 2015. Disponível em: <https://natlib.govt.nz/files/webarchive/nzwebarchive-awarenessanduse.pdf>. Acesso em: 13 out. 2021.

ROCKEMBACH, M. Arquivamento da Web: estudos de caso internacionais e o caso brasileiro. **Revista Digital de Biblioteconomia e Ciência da Informação UNICAMP**, v. 16, n. 1, p. 07-24, 2018. Disponível em: <http://eprints.rclis.org/32269/1/8648747-30995-7-PB.pdf>. Acesso em: 8 nov. 2021.

ROCKEMBACH, M. **Preservação do Conteúdo Web Brasileiro**. Mesa-redonda apresentada no 9º Fórum da Internet do Brasil. Manaus, AM. [S.l.: s.n.], 2 outubro de 2019. 1 vídeo (1h 20 min). Publicado pelo canal NICbrvdeos. Disponível em: https://www.youtube.com/watch?v=ZanQR9-WR_E. Acesso em: 7 fev. 2022.

ROCKEMBACH, M. A web brasileira na Covid-19: arquivamento da web e preservação digital. **Liinc Em Revista**, v. 17, n. 1, e5713, 2021. doi: <https://doi.org/10.18617/liinc.v17i1.5713>.

ROSA, Carlos André; CRAVEIRO, Olga; DOMINGUES, Patricio. Open source software for digital preservation repositories: a survey. **International Journal of Computer Science & Engineering Survey (IJCSSES)** Vol.8, No.3, June 2017. Available at: DOI:10.5121/ijcses.2017.8302 21 Accessed in 22 Out. 2021.

SANTOS, Gildeir Carolino; SHINTAKU, Milton (Org.). **Ecossistemas e inovações tecnológicas**: da construção às boas práticas. Campinas: Unicamp/BCCL; Brasília: Ibict, 2022. 376 p. DOI: 10.22477/ISBN9786588816363.

SANTOS, V. B. D. Arquivamento web: legislação correlata. **Revista Brasileira de Preservação Digital**, v. 1, n. 2020, 2000. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/158975>. Acesso em: 23 fev. 2022.

SCHWIRTLICH, A. M. Preserving large-scale cultural heritage: the case for collaboration. In: **CONGRESS OF SOUTHEAST ASIAN LIBRARIANS (CONSAL)**, Bangkok, Thailand, p. 1-12, jun. 2015. This paper was presented at the Conference CONSAL XVI held from 11-13 June 2015 in Bangkok, Thailand. Disponível em: <https://myrepositori.pnm.gov.my/handle/123456789/4015>. Acesso em: 17 out. 2021.

SCHAFER, V.; WINTERS, J. Perspectives on web archive studies: taking stock, new ideas, next steps. **WARCnet Papers**, 2020. Disponível em: <https://bit.ly/3ektICP>. Acesso em: 1 set. 2021.

SILVA, Karina Moura da. **Um modelo de ciclo de vida de dados na web**. 107 f.: il., fig., tab. 2019. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Pernambuco, Recife, 2017. Disponível em: <https://repositorio.ufpe.br/handle/123456789/34147>. Acesso em 28 out.2021.

SOUSA, Rosilene Paiva Marinho de. SHINTAKU, Milton. Informação legislativa sobre inovação tecnológica. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 22., 2022, Porto Alegre. **Anais [...]**. Porto

Alegre: Associação de Pesquisa e Pós-graduação em Ciência da Informação (Ancib), 2022. Disponível em: <https://enancib.ancib.org/index.php/enancib/xxii/enancib/paper/view/831>. Acesso em: 17 mar. 2023.

SOUSA, Rosilene Paiva Marinho de; SHINTAKU, Milton; NAKANO, Natália; VIDOTTI, Silvana Aparecida Borsetti Gregório. Elementos de apoio legal sobre accesibilidad digital para personas sordas en Brasil. **Revista EDI-CIC**, [S. l.], v. 2, n. 4, 2022. Disponível em: <https://ojs.edicic.org/index.php/revistaedicic/article/view/118>. Acesso em: 17 mar. 2023.

SOUZA, A. H. L. R.; OLIVEIRA, A. F.; D'AVILA, R. T.; CHAVES, E. P. S. S. O modelo de referência OAIS e a preservação digital distribuída. **Ciência da Informação**, v. 41, n. 1, 2012. DOI: <https://brapci.inf.br/index.php/res/v/19386>. Acesso em: 10 nov. 2021.

TEMATRES. **TemaTres**: servidor de vocabulários controlados. 2022. Disponível em: <http://r020.com.ar/tema-tres/manual/>. Acesso em: 27 mar. 2022.

TOMAÉL, Maria Inês; SILVA, Teresinha Elisabeth da. Repositórios institucionais: diretrizes para políticas de informação. In: **Encontro nacional de pesquisa em Ciência da informação – WaybackMachine**. [2022?]. Disponível em: <https://archive.org/web/>. Acesso em: 7 dez. 2021.

VECHIATO, Fernando Luiz; VIDOTTI, Silvana Aparecida Borsetti Gregorio. **Encontrabilidade da informação**. São Paulo: Cultura Acadêmica, 2014. 198 p. ISBN 978-85-7983-586-5 (Coleção PROPG Digital).

ZENG, M. L. Interoperability. **Knowledge Organization**, v. 46, n. 2, p. 122-146, 2019.

ZIERAU, E.; SCHULTZ, M. **Criando uma estrutura para aplicação do OAIS à preservação digital distribuída**. 10th International Conference on Preservation of Digital Objects, Lisboa, Portugal, 3-5 set., 2013, p. 78-83. Tradução para o português disponível em: <https://cridi.ici.ufba.br/institucional/arquivos/artigos/artigo-sobre-o-processo-de-criacao-de-uma-Estrutura-para-aplicacao-do-OAIS-a-preservacao-digital-distribuida.pdf>. Acesso em 5 set. 2018

ZIERAU, E.; SCHULTZ, M. Criando uma estrutura para aplicação do OAIS à preservação digital distribuída. Traduzido para o português por Miguel Rio Branco Nabuco de Gouvea. Revisão de Rubens Ribeiro Gonçalves da Silva. Original em inglês publicado em **Proceedings of the 10th International Conference on Preservation of Digital Objects**, Lisboa, Portugal, p. 78-83, 2013. Tradução para o português disponível em: <https://cridi.ici.ufba.br/institucional/arquivos/artigos/artigo-sobre-o-processo-de-criacao-de-uma-Estrutura-para-aplicacao-do-OAIS-a-preservacao-digital-distribuida.pdf>. Acesso em: 5 set. 2018.

