



UNIVERSIDADE DE BRASÍLIA
Faculdade de Ciência da Informação
Curso de Graduação em Biblioteconomia

VICTÓRIA DE ABREU E SILVA

PRESERVAÇÃO DIGITAL:
um estudo sobre padrões de metadados

Brasília - DF
2018

VICTÓRIA DE ABREU E SILVA

PRESERVAÇÃO DIGITAL:

um estudo sobre padrões de metadados

Monografia apresentada como requisito parcial das exigências para obtenção do título de Bacharel em Biblioteconomia pela Faculdade de Ciência da Informação (FCI), da Universidade de Brasília (UnB).

Orientador: Professor Doutor Márcio Bezerra da Silva



Título: Preservação digital: um estudo sobre padrões de metadados.

Aluna: Victória de Abreu e Silva.

Monografia apresentada à Faculdade de Ciência da Informação da Universidade de Brasília, como parte dos requisitos para obtenção do grau de Bacharel em Biblioteconomia.

Brasília, 11 de julho de 2018.

Marcio Bezerra da Silva - Orientador
Professor da Faculdade de Ciência da Informação (UnB)
Doutor em Ciência da Informação

Milton Shintaku - Membro externo
Doutor em Ciência da Informação

Ernani Rufino dos Santos Júnior - Membro externo
Bibliotecário da Câmara dos Deputados
Mestre em Ciência da Informação

Dados Internacionais de Catalogação na Publicação (CIP)

S624e Silva, Victória de Abreu e.
Preservação digital : um estudo sobre padrões de metadados / Victória de Abreu e Silva. – 2018.
76 f. : il.

Monografia (Bacharelado em Biblioteconomia) –
Universidade de Brasília, Faculdade de Ciência da
Informação, Curso de Biblioteconomia, 2018.

Orientador: Prof. Dr. Márcio Bezerra da Silva

1. Memória. 2. Preservação digital. 3. Metadados.
4. Metadados de preservação digital. I. Título.

CDD 025.84

Dedico este trabalho à minha mãe Mônica e ao meu falecido pai, Carlos, que se empenharam incansavelmente para tornar a mim e minhas irmãs felizes e aptas a conquistarem nossos sonhos.

AGRADECIMENTOS

Agradeço, em primeiro lugar, não apenas pelo encerramento desta importante etapa, a graduação, mas sim por toda minha trajetória educacional, aos meus amados pais Mônica e Carlos, que sempre me estimularam a estudar, ler e a me interessar pelo conhecimento das coisas, e que me proporcionaram todo o suporte, apoio e amor necessários para que eu construísse meu caminho até aqui. Agradeço também às minhas irmãs Débora e Eduarda que apesar das brigas, comuns aos irmãos em “fase de crescimento”, sempre foram minhas amigas e parceiras ao longo da vida.

Agradeço a todos os amigos que fiz ao longo da jornada, que foi a graduação. Inicialmente aos amigos que fiz em minha primeira escolha de curso na UnB, da qual não pude concluir devida a falta de “química” entre nós, mas foi um momento em que eu fiz importantes amigos. Aline, Carol, Clarissa, Débora, Jacqueline, Jéssika, Larissa e Paulo, os meus sinceros agradecimentos pelo apoio e pela amizade, que tornaram o caminho muito mais prazeroso. Quanto aos amigos que fiz em minha escolha de coração, a Biblioteconomia, onde também tive a sorte de conhecer pessoas importantes, que foram primordiais na minha caminhada, tanto com o apoio nos trabalhos acadêmicos e estágios da vida, quanto pela amizade. Cristina, Emanuelle, Denise, Fabíola, Larissa, Paola e Thais, e a todos os outros, meu eterno agradecimento por tornarem essa etapa ainda melhor devido a toda amizade e apoio que recebi.

Agradeço a uma importante pessoa que entrou em minha vida, em um momento difícil, a perda de meu pai, e que foi de extrema importância para que eu superasse essa perda. Ele trouxe consigo o amor necessário para que eu continuasse perseverando. A ele, meu amor, Hélio, agradeço do fundo do meu coração pelo apoio, pelo amor, pela parceria e pela amizade. Você não é somente um namorado, mas o meu melhor amigo e parceiro. Agradeço também a sua família, me recebendo de braços abertos, com muito carinho e amor.

Ao meu orientador, Márcio Bezerra da Silva, agradeço imensamente por tornar esse processo, comumente temido por todos, tranquilo e até prazeroso, com sua orientação atenciosa e empolgada que me incentivaram e animaram a prosseguir.

Por fim, sou grata à Deus por me permitir ter todas essas queridas pessoas a agradecer e por me trazer até aqui, com saúde, capacidade e força.

*“Em um estado sombrio nós nos encontramos...
um pouco mais de conhecimento iluminar nosso
caminho pode.”*

Mestre Yoda, Star Wars

RESUMO

Discute a preservação em meio digital, identificando e apresentando padrões de metadados pelo viés da Ciência da Informação. Fundamenta a presente investigação segundo uma literatura que traz conceitos de memória e de preservação digital, especialmente elencando métodos de preservação como a adoção de padrões e os metadados de preservação, que por sua vez discorre sobre metadados, exemplificados pelo esquema *dublin core*. Resulta, conforme uma metodologia constituída pelo método dedutivo e pelas técnicas de pesquisa exploratória e bibliográfica, na identificação de padrões de metadados para preservação dos dados de objetos digitais e na aplicação e uso desses padrões em ambientes digitais, como bibliotecas digitais, repositórios digitais e catálogos eletrônicos, a partir da iniciativa de preservação OAIS e dos esquemas de metadados PREMIS, MODS e METS. Conclui-se que a utilização de padrões de metadados para descrever e armazenar documentos, bem como realizar a guarda dos dados de forma padronizada e comunicável entre máquinas, é uma importante etapa no processo de preservação digital, facilitando a comunicação entre os centros de informação/Instituições em redes de informação via o conceito de interoperabilidade.

Palavras-chave: Memória. Preservação digital. Metadados. Metadados de preservação.

ABSTRACT

Discusses the preservation in a digital environment, identifying and presenting methods of implementation and use by the bias of Information Science. Bases the present investigation according to a literature that brings concepts of memory and digital preservation, especially by naming preservation methods such as the adoption of standards and preservation metadata, which in turn talks about metadata, exemplified by the Dublin Core scheme. The result according to a methodology consisting of the deductive method and the techniques of exploratory and bibliographic research is the identification of digital preservation strategies, emphasized by applications and use of the following preservation metadata schemes: OAIS, PREMIS, MODS and METS. Concluded that there is a longing for the use of standards to describe and store documents, as well as to carry out guarding and preservation in a standardized and communicable way between machines, an aspiration that would facilitate communication between information centers/Institutions in information networks via the concept of interoperability.

Keywords: Memory. Digital preservation. Metadata. Metadata for preservation.

LISTA DE FIGURAS

Figura 1	Elementos-chave na atividade da preservação digital	23
Figura 2	Níveis de um objeto digital na preservação digital	24
Figura 3	Métodos de preservação digital	25
Figura 4	Registro simples na BDM	32
Figura 5	Registro qualificado na BDM	35
Figura 6	Registro em DC/HTML com destaque nos qualificadores	36
Figura 7	O elemento <i>meta name</i> no código em HTML	37
Figura 8	Registro em DC/XML	38
Figura 9	O elemento <i>date</i> no código em XML	39
Figura 10	Modelo de informação no OAIS	45
Figura 11	Modelo funcional	47
Figura 12	Modelo PREMIS	52
Figura 13	Exemplo de um relatório de objetos no PREMIS	54
Figura 14	Exemplo de um registro em XML MODS no Copac	58
Figura 15	Registro em METS na <i>Brown University Library</i>	61

LISTA DE QUADROS

Quadro 1	Elementos DC	30
Quadro 2	Qualificadores do padrão DC	33
Quadro 3	Elementos de um registro de eventos no PREMIS	53
Quadro 4	Elementos de um relatório de dados no PREMIS	54
Quadro 5	Elementos do esquema MODS	56
Quadro 6	Elementos de padrões de metadados para preservação digital	62

LISTA DE ABREVIATURAS E SIGLAS

AACR2	<i>Anglo American Cataloguing Rules, second edition</i>
CI	Ciência da Informação
BDM	Biblioteca digital de monografias da UnB
BND	Biblioteca Nacional Digital de Portugal
CCSDS	<i>Consultive Committee for Space Data Systems</i>
CD	<i>Compact Disc</i>
CDR	<i>Carolina Digital Repository</i>
Cedi	Centro Documentação e Informação da Câmara dos Deputados
CURL	<i>Copac Uniform Resource Locator</i>
Copac	<i>UK & Irish Research Library Catalogue</i>
DC	<i>Dublin Core</i>
DCMI	<i>Dublin Core Metadata Initiative</i>
DDC	<i>Dewey Decimal Classification</i>
DLC	<i>Digital Library Collections</i>
DVD	<i>Digital Video Disc</i>
HTML	<i>HyperText Markup Language</i>
IBICT	Instituto Brasileiro de Informação em Ciência e Tecnologia
ISBN	<i>International Standard Book Number</i>
ISO	<i>International Organization for Standardization</i>
LC	<i>Library of Congress</i>
LCC	<i>Library of Congress Classification</i>
LCSH	<i>Library of Congress Subject Headings</i>
LIBER	Laboratório de Tecnologia do Conhecimento
LOCKSS	<i>Lots of Copies Keeps Stuff Save</i>
MARC	<i>Machine Readable Cataloging</i>
MeSH	<i>Medical Subject Headings</i>
METS	<i>Metadata Encoding and Transmission Standard</i>
MIX	<i>Metadata for Images in XML</i>
MODS	<i>Metadata Object Description Schema</i>
NISO	<i>National Information Standards Organization</i>
OAIS	<i>Open Archival Information System</i>
OCLC	<i>Online Computer Library Center</i>
PAM	Perfil de aplicação de metadados
PREMIS	<i>PREservation Metadata Implementation Strategies</i>
RDA	<i>Resource Description and Access</i>
RFC	<i>Request For Comments</i>
RLG	<i>Research Libraries Group</i>
SAAI	Sistema de arquivamento aberto de informação
SO	Sistema operacional
TCC	Trabalho de conclusão de curso
TGN	<i>Getty Thesaurus of Geographic Names</i>
TIC	Tecnologias de informação e comunicação
UDC	<i>Universal Decimal Classification</i>
UFPE	Universidade Federal de Pernambuco
URI	<i>Uniform Resource Identifier</i>
XML	<i>Extended Markup Language</i>

SUMÁRIO

1 INTRODUÇÃO	14
1.1 PROBLEMATIZAÇÃO	17
1.2 JUSTIFICATIVA	17
1.3 OBJETIVOS	18
2 FUNDAMENTAÇÃO TEÓRICA.....	19
2.1 MEMÓRIA.....	19
2.2 PRESERVAÇÃO DIGITAL	21
2.3 MÉTODOS DE PRESERVAÇÃO DIGITAL	25
2.3.1 Metadados	27
2.3.1.1 Dublin Core	29
3 METODOLOGIA.....	40
3.1 CARACTERIZAÇÃO DA PESQUISA.....	40
3.2 CAMPO DA PESQUISA	41
3.3 ETAPAS DA PESQUISA	41
4 APLICAÇÕES E USO DE MÉTODOS DE PRESERVAÇÃO DIGITAL.....	43
4.1 OAIS	43
4.2 PREMIS.....	50
4.3 MODS.....	55
4.4 METS	58
5 CONSIDERAÇÕES FINAIS.....	67
REFERÊNCIAS	69

1 INTRODUÇÃO

A informação figurou-se, e continua figurando, como uma importante ferramenta de auxílio no desenvolvimento da humanidade, ao longo da história. Detê-la e (saber) usá-la representa uma vantagem nas dinâmicas das sociedades, tornando-a um elemento de papel importante à compreensão de mundo frente ao homem enquanto um ser social. Neste sentido, o ato de passar a informação, de geração em geração, no auxílio a tomadas de decisões, na produção de conhecimento e na construção de identidades individuais e coletivas, evidencia a importância e a necessidade de se guardá-la enquanto um processo perpetuação, influenciando tanto na redescoberta do passado, como no ato de futuros julgamentos.

A supracitada contextualização vai ao encontro do conceito de memória, ao qual pode sofrer alterações conforme o ponto de vista da área do conhecimento em que se aplica. Em termos gerais, “[...] memória como propriedade de conservar certas informações, remete-nos em primeiro lugar a um conjunto de funções psíquicas, graças às quais o homem pode atualizar impressões ou informações passadas, ou que ele representa como passadas” (LE GOFF, 2013, p. 387). Porém, tal conceito está mais atrelado às concepções de memória nas áreas da biologia, medicina, psicologia, psiquiatria etc.

Le Goff (2013) aponta também a existência de outro tipo de concepção para a memória, ou seja, na perspectiva coletiva, ou social, compreendida como aquela cuja formação vem por meio de produções realizadas pela humanidade como, por exemplo, a escrita e seus produtos, a música, as artes, as ciências dentre outras. Dito isto, o coletivo apresenta-se atrelado às ciências sociais e traz consigo o conceito de identidade coletiva.

No campo da Ciência da Informação (CI), o conceito de memória, em seu sentido coletivo, promove discussões que evidenciam às “[...] possibilidades de vínculos entre memória e informação, memória e documento, memória e conhecimento, memória e preservação, memória e acesso” (OLIVEIRA, 2010, p. 14). Assim sendo, inclinando as investigações promovidas na CI sobre memória, na presente pesquisa, compreende-se que “[...] o estudo do tema memória [...] pode prover os fundamentos teóricos adequados à sustentação das práticas relacionadas à preservação e ao acesso às informações registradas” (OLIVEIRA, 2010, p. 15).

Entender o conceito de memória é alinhar-se à representação de algo ao longo dos anos, tanto oralmente e materialmente, quanto fisicamente e digitalmente. Com isso, bibliotecas, museus e arquivos apresentam-se como espaços que podem resguardar essa memória, a partir da realização de ações de armazenamento de informação com fins de preservação, sendo esta última definida por Cunha e Oliveira (2008) como “medidas empreendidas com a finalidade de

proteger, cuidar, manter e reparar ou restaurar os documentos”. Entretanto, as finalidades supracitadas transcenderam ao universo digital, trazendo a intitulada preservação digital. Neste caso, vale salientar que a preservação, na perspectiva digital, não tem por objetivo a integridade física do material, conforme pode ser verificado em conceitos mais amplos de preservação. Na ótica digital, a preocupação está na integridade do conteúdo informacional, visto que as tecnologias em meio digital sofrem constantes mudanças.

Tomando como base o atual cenário de constante produção de informação, a preservação digital vem sendo levada em consideração ao funcionamento dos sistemas, pois assume “[...] uma importante etapa no gerenciamento de objetos digitais, [...] garantido o emprego de mecanismos que permitem o armazenamento em repositórios de objetos digitais que, garantem a autenticidade e a perenidade dos seus conteúdos” (MUNHOZ, 2011, p. 24), contexto que se mostra discursivo na CI e aplicado em seus campos de atuação como a biblioteconomia, a arquivologia e a museologia.

Outrossim, deve-se ter em mente que a memória do homem continua sendo desenvolvida. Neste universo de produção informacional são gerados e apresentados conteúdos via suportes como livros e revistas, no entanto, com o advento da computação¹, principalmente a partir da Internet, a humanidade se deparou com um novo tipo de produção informacional, agora por meio de *websites*. Se antes os periódicos ofereciam os seus artigos científicos em papel, agora, esses compêndios passaram a ser oferecidos de maneira digital, criando-se um novo cenário, formalizado pela produção crescente e constante de conteúdo. Para Arellano (2004), esses conteúdos são tanto produzidos originalmente em computador, quanto podem ser uma versão de um conteúdo originalmente físico e que migrou para uma versão digital. Entrando nesta questão do documento, Munhoz (2011, p. 21) aponta que, no âmbito da biblioteconomia, documento é conceituado como sendo “qualquer unidade significativa de informação que tenha sido registrada em um suporte que permita seu armazenamento e sua posterior recuperação”. Desta maneira, um documento digital, ou objeto digital², permanece com tais características, porém seu armazenamento e/ou criação passa a ser realizado em digitalmente.

¹ Trata-se de “[...] um corpo de conhecimentos formado por uma infraestrutura conceitual e um edifício tecnológico onde se materializam o hardware e o software” (FONSECA FILHO, 2007, p. 13).

² A fim de evitar exclusões de conceitos devido à questão semântica da palavra documento, que pode remeter o leitor a um objeto apenas textual, optou-se por utilizar, neste trabalho, o termo objeto digital, que engloba documentos digitalizados, documentos criados em meio digital, *websites*, *blogs*, fóruns dentre outros.

Entretanto, independentemente dos tipos de objetos, a preocupação para com a preservação dos mesmos pode ser pensada a partir dos métodos que visem a representação e o acesso a esses materiais por intermédio da preservação dos dados que os identificam, para que assim ocorra a efetiva preservação da memória. Para tanto, faz-se uso de padrões de metadados, recurso lógico conceituado pela *National Information Standards Organization* (NISO, 2004, p. 5) como “dados sobre dados ou informações sobre informações” e configurado como um “elemento que descreve, explica, localiza e facilita a recuperação e o uso de um recurso de informação” (NISO, 2004, página). Entre os exemplos, cita-se o padrão de metadados *Dublin Core* (DC)³, criados e implementados via linguagens *web* de marcação como o *HyperText Markup Language* (HTML) e *eXtensible Markup Language* (XML), dentre outros padrões.

No caso das iniciativas de preservação digital, a mais tradicional se chama *Open Archival Information System* (OAIS). Para Arellano (2004), a estratégia OAIS é um padrão referência de metadados de preservação para arquivos que necessitam de uma salvaguarda em longo prazo, inclusive, considerada uma *International Organization for Standardization* (ISO) desde 2003 e atestada pelo *Consultive Committee for Space Data Systems* (CCSDS). Ainda conforme Arellano (2004), vale salientar que as estratégias também devem abordar questões como migração e conversão de formatos, recuperação de documentos etc.

Além de preocupar-se com os objetos digitais, quanto ao suporte, as iniciativas também promovem a preservação dos dados que representam esses objetos, em diferentes necessidades de armazenamento e acesso, frente às evoluções tecnológicas discutidas e assumidas nos mais diversos campos do conhecimento, como os presentes na CI, por exemplo.

Conforme o contexto supracitado, este trabalho constrói-se da seguinte forma: no capítulo um (1) apresenta-se a introdução, incluindo os problemas, a justificativa e os objetivos; em seguida, no capítulo dois (2), discute-se teoricamente as bases que constituem o presente trabalho a partir dos assuntos memória, preservação digital e metadados; na sequência, no capítulo três (3), os aportes metodológicos são expostos; no capítulo quatro (4) são interpretados os dados extraídos da literatura levantada, centrados em métodos de preservação digital; e, por fim, no capítulo cinco (5) são apresentadas as considerações finais.

³DC é um “padrão de dados para catalogação de recursos eletrônicos da WEB” (GRÁCIO, 2002, p. 36).

1.1 PROBLEMATIZAÇÃO

Compreende-se que a produção de objetos digitais se configura como uma realidade de destaque na atualidade, fazendo com que venha à tona questões sobre armazenamento, tratamento, acesso e preservação de informação. Neste viés, os padrões de metadados, enquanto o foco do presente estudo no certame da preservação digital, assumem um espaço discursivo no mundo digital, especialmente na Internet, como vem ocorrendo, tanto sobre o alcance da informação à posterioridade, como na superação aos limites da obsolescência. Em frente a este cenário, o problema da pesquisa é: *Quais são os padrões tecnológicos em metadados adotadas para a realização da preservação dos dados de objetos digitais?*

1.2 JUSTIFICATIVA

Olhando para o cenário atual de grande produção informacional no mundo digital, disponibilizada em ambientes como as bibliotecas digitais, os repositórios digitais e catálogos eletrônicos, surgiu a necessidade de reflexão, especificamente na CI, sobre os métodos que vem sendo adotados acerca da preservação dos dados, descritivos e representativos, dos objetos digitais. Neste sentido, o interesse da presente pesquisa deu-se início após um período de estágio obrigatório ao currículo de Biblioteconomia, realizado no Centro de Documentação e Informação da Câmara dos Deputados (Cedi).

No período de estágio, concomitantemente foi cursada a disciplina “Redes de Informação e Transferência de Dados”, momento em que se compreendeu sobre a utilização de metadados nas diversas finalidades, entre as quais foi citada a preservação. Consequentemente, o interesse foi ampliado devido a curiosidade em compreender os processos que permeavam a preservação dos objetos digitais presentes no acervo da Biblioteca Digital da Câmara dos Deputados e no Portal da Câmara dos Deputados⁴, principalmente na forma como estão organizados para a futura recuperação. Enquanto o acervo da Biblioteca Digital é composto de materiais como textos, áudios e vídeos, especialmente nas áreas das Ciências Sociais e Direito, no Portal estão disponibilizados processos legislativos, leis, propostas, e suas tramitações, dentre outros conteúdos. Ambos os acervos são de acesso livre e gratuito, e se configuram como elementos que auxiliariam na construção da memória coletiva da nação legislativa do país. Portanto, pensando no acesso futuro, especialmente frente as constantes mutações das

⁴ Website: <http://www2.camara.leg.br/>

tecnologias de informação e comunicação (TIC), o presente trabalho, segundo uma investigação literária, se justifica pelo interesse em compreender o conceito de preservação digital e descobrir padrões de metadados usados para a preservação dos dados que descrevem e representam os objetos em meio digital e, conseqüentemente, permitir a recuperação de informações à posterioridade.

A capacidade das recentes TIC, ao possibilitar a expansão (alcance) no acesso às informações, independentemente do tempo e espaço, oferta subsídios a realização da preservação da memória de um povo, de um conglomerado, enquanto uma temática que se apresenta discutida no terreno da CI.

1.3 OBJETIVOS

De forma geral, este estudo objetiva identificar padrões de metadados de preservação digital apresentados na literatura. No que se refere à perspectiva específica, a investigação busca elencar métodos de preservação dos dados de objetos digitais, listar exemplos de ambientes digitais que fazem uso de padrões de metadados para preservação digital e apontar atributos em metadados nos exemplos de ambientes digitais identificados.

2 FUNDAMENTAÇÃO TEÓRICA

A fundamentação teórica, ou referencial teórico, dentro de uma pesquisa científica, tem por objetivo dar embasamento conceitual à mesma, dando-a substância e suporte para a construção dos objetivos pretendidos. A partir do levantamento bibliográfico, entre textos e autores que coincidem com a linha de investigação proposta pelo pesquisador, realiza-se um diálogo entre eles, gerando conceitos fundamentais à compreensão das etapas da pesquisa. Conforme Prodanov e Freitas (2013), nesta etapa do trabalho, o pesquisador realiza uma análise das produções científicas que tratam dos assuntos abordados, fornecendo o embasamento teórico e metodológico para o desenvolvimento do projeto de pesquisa.

A partir da compreensão deste processo, correspondente ao presente trabalho, a fundamentação teórica traz uma base necessária à compreensão envolta de dois conceitos básicos que serão necessários para a compreensão da importância da salvaguarda de informações, principalmente na atualidade: memória e preservação digital, com destaque aos tipos, aplicações e usos de métodos.

2.1 MEMÓRIA

Dentro de uma percepção cognitiva, o homem realiza processos mentais para seu desenvolvimento e perpetuação da espécie. Neste caso, transmitir informações necessárias à sobrevivência deflagra tais processos. Interiorizada a essas atividades mentais está a memória, que por sua vez configura-se como um elemento constituído da capacidade do homem de reter e guardar informações, tanto internamente, por meio de lembranças e pensamentos, quanto externamente, via escrita, desenho, histórias contadas dentre outros. Ademais, tais informações, transformadas em memória, registradas ou não, podem ser repassadas.

Conforme Le Goff (2013), a memória pode ser definida de diferentes formas. Por exemplo, na biologia, na fisiologia humana, na psicologia, na psiquiatria e na medicina, a memória é percebida como a “propriedade de conservar certas informações [...] [construídas a partir de] [...] um conjunto de funções psíquicas, graças às quais o homem pode atualizar impressões ou informações passadas, ou que ele representa como passadas” (LE GOFF, 2013, p. 387), estando tal conceito, portanto, ligado às questões internas e cognitivas do homem.

No escopo da CI, esta percepção cognitiva é levada em consideração na formalização de três abordagens conceituais sobre a memória, apresentadas por Oliveira (2010, p. 59) a partir da seguinte compreensão:

[...] a primeira, a memória como capacidade humana, relacionada ao processo cognitivo na recuperação da informação através de associações mentais; a segunda, associada ao uso de tecnologias que viabilizam a criação de memórias auxiliares à memória humana; e a terceira, a memória social, relacionada à uma memória compartilhada, através da tradição e da preservação dos registros da informação.

Portanto, quando estudada a partir das ciências sociais, como no caso da CI, a memória pode ser compreendida como memória individual e memória coletiva/social⁵. Ao posicionar o olhar ao indivíduo, tem-se a memória compreendida como “um fenômeno individual e psicológico” (LE GOFF, 2013, p. 387) que está ligado ao meio social em que o indivíduo se insere. Segundo Julio e Botelho (2015, p. 31), “a memória individual se dá a partir da construção de lembranças pessoais e autobiográficas, e se apoia na memória coletiva, que é bem mais ampla”. Com base nas teorias de Halbwachs⁶, não se pode compreender a memória individual distanciada da memória coletiva. Dessa maneira, a memória constitui-se por indivíduos em interação, por meio de grupos sociais, tendo por resultado as lembranças individuais (ARAÚJO; SANTOS, 2007).

A memória coletiva/social concentra-se no grupo e ou sociedade aos quais àquele indivíduo faz parte, assumindo características que o posicionarão dentro de classes sociais e que traz luz aos acontecimentos ocorridos em uma data específica ou um fato histórico. Nesta linha, Le Goff (2013, p. 435) expressa a sua visão sobre a memória coletiva da seguinte maneira:

A evolução das sociedades, na segunda metade do século XX, elucida a importância do papel que a memória coletiva desempenha. Exorbitando a história como ciência e como culto público, ao mesmo tempo a montante, enquanto reservatório (móvel) da história rico em arquivos e em documentos/monumentos, e aval, eco sonoro (e vivo) do trabalho histórico, a memória coletiva faz parte das grandes questões das sociedades em vias de desenvolvimento, das classes dominantes e das classes dominadas, lutando todas pelo poder ou pela vida, pela sobrevivência e pela promoção.

Diante da supracitada compreensão, sobre a memória coletiva, o homem começa a perceber, historicamente, as limitações de sua memória, levando-o a buscar estratégias auxiliares como as memórias artificiais, por exemplo (OLIVEIRA, 2010). Para tanto, “[...] foi preciso conceber um sistema de utilização de signos por meio dos quais as ideias eram fixadas em um suporte físico, o que deu origem à escrita” (OLIVEIRA, 2010, p. 12). Entretanto, os registros não se limitam à escrita de palavras, ou seja, externalizar a memória por meio de

⁵ Vale salientar que “o critério tradicional de distinção entre memória coletiva e memória social, fundamentado na ausência [coletiva] ou presença [social] da escrita” (GONDAR, 2008), não será levado à risca neste trabalho. O termo coletivo poderá ser usado como sinônimo de social;

⁶ Maurice Halbwachs (1877-1945), sociólogo francês e pioneiro nas discussões sobre memória coletiva (VIEIRA, 2015).

imagens, musicas, imagens em movimento e dentre outras (OLIVEIRA, 2010). Em complemento, Melo e Konrad (2015, p. 34) afirmam que a memória, quando exteriorizada, “[...] é representada por meio de registros de informação, qualquer que seja o suporte em que está contida e que seja passível de recuperação (acesso); aos fatos ou às ações registradas se atribui um significado o qual constitui a memória desses acontecimentos”.

As sociedades adotam documentos enquanto registros que possibilitam averbar e transmitir a informação, entre livros, partituras, cartas, relatos e uma infinidade de produções propiciadas pelo uso dos signos e da escrita. Para tanto, faz-se uso de ambientes que objetivam a conservação, a preservação e a propagação desses documentos, tanto no prisma individual, quanto na ótica coletiva, culminando nos chamados lugares de memória.

Tantos locais físicos como museus, bibliotecas, coleções, arquivos, cemitérios, galerias de arte, entre outros, quanto elementos representativos como uma bandeira, um hino, festas, aniversários, entre outros cumprem a função de **Lugar de Memória**, isto é, são lugares onde se cristalizam memórias coletivas, onde salvaguardam elementos importantes à identidade de uma nação. (MAGALHÃES, 2014, p. 26, grifo da autora)

Os referidos ambientes são os locais onde a memória se desenvolve, se cristaliza e/ou é resguardada, pensamento que vem ganhando respaldo a partir da produção constante de informação, que por sua vez está imbricada aos recentes adventos tecnológicos. Desta maneira, pensando em como acessar essas informações, nos lugares de memória, a qualquer momento, por qualquer pessoa e por qualquer suporte de uso, faz-se necessária a adoção de processos que promovam a preservação digital.

2.2 PRESERVAÇÃO DIGITAL

Tendo-se em mente a importância da memória no processo de desenvolvimento do homem e da sociedade onde vive, compreende-se que a produção informacional⁷, mediante os mais variados tipos de registro, configura-se como processo pertencente à construção da memória humana. Pensando neste processo, de produção informacional, destacam-se três momentos históricos: o advento da escrita, anteriormente citado, a prensa móvel de Gutemberg⁸

⁷ Produção informacional, no presente trabalho, representa a produção de informação realizada pelo homem e pelas sociedades. É toda e qualquer informação produzida, por meio de qualquer meio e em qualquer formato, podendo ser uma partitura, um livro, um manual, um desenho dentre outros.

⁸ Gensfleisch von Guttenberg, xilogravador e ourives por profissão, aperfeiçoou as técnicas de impressão por tipos móveis, introduzindo o uso das ligas metálicas, sendo estas mais resistentes e duráveis. Guttenberg inventou também a primeira máquina de impressão, revolucionando a imprensa (FERNANDES, 2001).

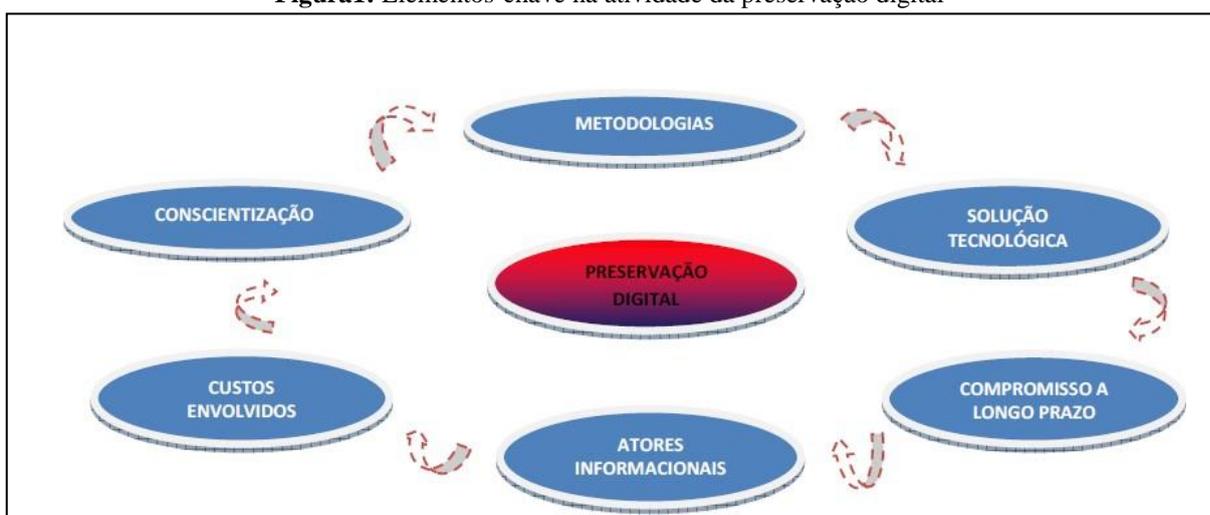
e a segunda guerra mundial, ao qual apresentou o conseqüente aumento na produção científica, sendo esta produção o resultado de pesquisa, que pode ser traduzido como um texto científico, produzido em uma área do conhecimento, que tem coerência, consistência, originalidade e objetividade [...], tendo este que ser publicado, para ser validado e julgado pelos pares. (ALVES, 2009).

Ao passo que o período da segunda guerra mundial proporcionou o avanço das ciências e das tecnologias, houve um aumento, em larga escala, no número de informações produzidas pelo homem, nas mais diversas formas, delineando uma espécie de nova explosão informacional. O avanço científico e tecnológico não se refreou, pelo contrário, vem aumentando exponencialmente até os dias de hoje, apresentando “a explosão da informação [...] como um problema social que teve seu início com o desenvolvimento das ciências, e hoje estendesse para todas as atividades humanas” (SARACEVIC, 1995, p. 2).

Em tempos recentes, proporcionado pelo avanço científico e tecnológico, define-se um novo ambiente, ou seja, o universo digital, instaurado, conforme Munhoz (2011), por causa do avanço das TIC, aos quais facilitam, cada vez mais, a produção, o acesso, a troca e o armazenamento de objetos digitais. Segundo Ferreira (2006), tais objetos podem ser entendidos como informações representadas por meio de uma sequência de dígitos binários, englobando, tanto os objetos criados originalmente em ambiente digital, quanto aqueles que passaram de um suporte analógico para o digital, por meio da digitalização, por exemplo. Ferreira continua abordando sobre os objetos digitais ao afirmar que os mesmos podem estar representados em forma de “[...] documentos de texto, fotografias digitais, diagramas vectoriais, bases de dados, sequências de vídeo e áudio, modelos de realidade virtual, páginas Web e aplicações de software” (FERREIRA, 2006, p. 21).

A fim de conservar, preservar e prolongar o alcance dos objetos digitais vai-se ao encontro da chamada preservação digital, ação apresentada por Grácio e Fadel (2009) como um conjunto de atividades com vistas a garantir o acesso no futuro à informação digital, assegurando o “[...] armazenamento adequado, o acesso perene e a usabilidade dos objetos digitais para a posteridade (ARAUJO; SOUZA; ARELLANO, 2015, p. 5)”.

Para que a preservação digital transcenda do conceito à realidade de execução, algumas questões devem ser levadas em consideração. Neste caso, a figura 1 mostra os elementos-chave no processo de preservação digital, a saber:

Figura1: Elementos-chave na atividade da preservação digital

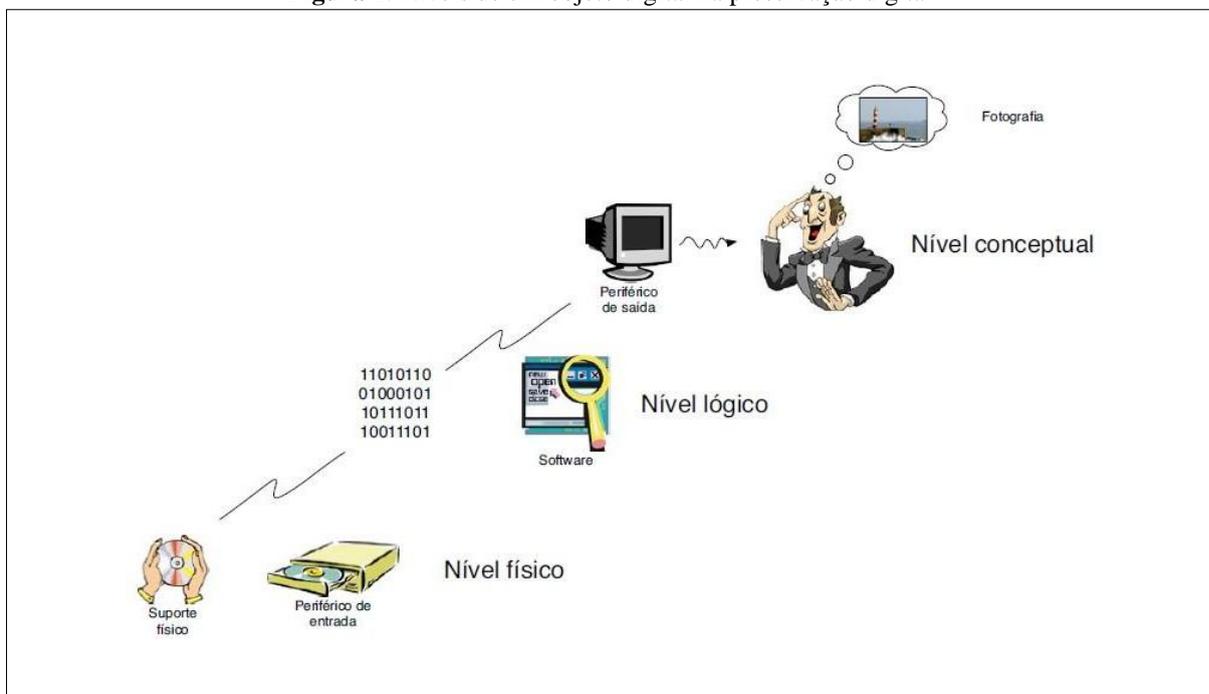
Fonte: Araújo; Souza; Arellano (2015, p. 5).

Os supracitados elementos, conforme Araújo, Souza e Arellano (2015), permeiam o processo de preservação digital quando se pensa em realizá-la. Primeiramente, tem-se a metodologia, que abarca as decisões que devem ser tomadas pela Instituição a fim de realizar a preservação digital. São elas: formulação da política de preservação, como ocorrerá a garantia da proteção à propriedade intelectual, utilização das normas internacionais que regiram o processo, padrões e formatos a serem utilizados e métodos de preservação a serem escolhidas. Na solução técnica deve-se escolher a tecnologia utilizada para a realização do processo, inclusive sugerindo-se uma que faça parte de uma iniciativa já estabelecida de preservação digital. Quanto à questão de compromissos a longo prazo, visa ser uma responsabilidade dos atores informacionais em garantir a continuidade do processo de preservação, independente de quaisquer problemas da Instituição. Logo, os atores informacionais são aqueles que interligam as demais ações, pois estes deverão responsabilizar-se pelo processo, pôr em prática cada etapa e garantir a continuidade da ação de preservação. Ademais, são também de responsabilidade dos atores os custos envolvidos no processo, pois deve-se pensar nos investimentos contínuos para permanência das atividades. Por fim, a conscientização vem para reforçar a necessidade das atividades que contempla o processo, pois a preservação dos objetos digitais faz parte do procedimento de preservação da memória. (ARAÚJO; SOUZA; ARELLANO, 2015).

A preservação digital objetiva, portanto, garantir que a informação digital esteja acessível, com qualidade e autenticidade, e podendo ser interpretada no futuro a partir de plataformas tecnológicas diferentes das utilizadas em sua criação (FERREIRA, 2006). Para tanto, a preservação digital leva em consideração três níveis, aos quais definirão um objeto

digital e a forma como este será preservado. Conforme a figura 2 é possível observar os níveis, conceitual, lógico e físico no processo de preservação digital.

Figura 2: Níveis de um objeto digital na preservação digital



Fonte: Ferreira (2006, p. 24).

Conforme a figura 2, no nível conceitual (intelectual), a preservação tem como foco a garantia da integridade e autenticidade da obra, preocupando-se, portanto, com mecanismos que asseguram este nível de preservação (JESUS; KAFURE, 2010), isto é, que a obra seja representada de maneira fidedigna e compreensível a quem interessar nos processos de busca. No nível lógico, a preservação utiliza-se de métodos que objetivam converter o formato, dos objetos digitais, em versões mais acessíveis e atuais, evitando assim a obsolescência dos mesmos, devido à falta de manutenção dos recursos de *software* e *hardware*. Quanto ao nível físico, a preservação preocupa-se com a conservação das mídias físicas, como o *Digital Video Disc* (DVD), o *Compact Disc* (CD) e os disquetes (JESUS; KAFURE, 2010), além de cartões de memória, *pen drives*, *blu-ray* etc.

Os níveis supracitados farão uso de uma ou mais técnicas/métodos para que a preservação digital ocorra, pois se fará necessário “[...] na maioria dos casos, mudar, recriar, renovar: mudar formatos, renovar mídias, hardware e software” (SAYÃO, 2005, p. 118). Desta forma, a fala do autor configura-se nos métodos de preservação, atividades que visam preservar o objeto digital, por intermédio de diferentes ações como, por exemplo, conversão de formatos, transferência de suportes e digitalizações.

2.3 MÉTODOS DE PRESERVAÇÃO DIGITAL

Existem, dentro da literatura especializada, duas concepções sobre a preservação dos objetos digitais, como tratado anteriormente. Em outras palavras, existem iniciativas que visam à preservação ou atualização a nível do *hardware*, enquanto outros visam preservar o conteúdo dos objetos digitais, isto é, o *software* e os seus dados. Arellano (2004) aponta que, tanto o dado, quanto a mídia que suporta a informação, devem estar em um estado que se permitam reproduzi-los a qualquer momento a partir de métodos estruturais e operacionais (figura 3).

Figura 3: Métodos de preservação digital

Métodos Estruturais	Métodos Operacionais
Adoção de padrões	Conservação de <i>software</i> / <i>hardware</i>
Elaboração de normas	Migração de suporte
Metadados de preservação digital	Conversão de formatos
Montagem de infra-estrutura	Emulação
Formação de consórcios	Preservação do conteúdo

Fonte: Arellano (2004, p. 18).

Os métodos estruturais são aqueles utilizados na descrição e na guarda dos objetos digitais, ou seja, medidas preventivas de preservação. Já os métodos operacionais são aspectos que garantem a reprodução da informação em formas diferentes de sua criação, tomando medidas que preservam e atualizam sua estrutura lógica. Em outras palavras, atualizando, transferindo e simulando formatos e suportes.

Na direção dos métodos estruturais, a **adoção de padrões** é o estabelecimento de formas uniformizadas de descrição e armazenamento de informações com o objetivo de facilitar os processos de recuperação e aplicação de métodos de preservação. A **elaboração de normas** vai ao encontro da adoção de padrões, pois são instruções, definidas pela unidade de informação, que visam certa padronização na descrição dos objetos digitais, bem como armazenamento, preservação, transferências, conversões dentre outras ações. Os **metadados de preservação digital** são as iniciativas base no processo de salvaguarda, pois são eles que descreverão o objeto digital de forma que este possa ser recuperado futuramente por meio dos dados que o identificam, além de ações voltadas a própria preservação como migrações, conversões dentre outras. Quanto à **infraestrutura**, tratam-se dos suportes estruturais, das unidades de informação, responsáveis por armazenar os objetos digitais com vistas à preservação, aos quais

afetarão diretamente na longevidade, ou seja, na conservação dos objetos digitais, a partir dos seguintes pontos: armazenamento em ambiente estável/controlável; atualizações frequentes para cópias em novas mídias; cópias de preservação, inclusive assumindo licenças e permissões; procedimentos de manuseio adequados; e mídia de armazenamento padrão para transferência de conteúdo. Por fim, a **formação de consórcios** é a aglutinação de instituições/unidades de informação que fazem uso de um modelo de preservação em comum e assim garantir um padrão universal de preservação, bem como tornar a comunicação entre todos eficaz e duradoura (ARELLANO, 2004).

Com relação à discussão dos métodos operacionais, a **conservação de software/hardware** sugere a adoção de uma mídia estável para o armazenamento dos objetos digitais para que, caso necessário, possa ser realizado uma cópia de dados para uma nova mídia (*refreshed*). No entanto, a cópia deve estar associada ao *software* original de criação do objeto digital e a um sistema operacional (SO) equivalente. A **migração de suporte** é um método bastante utilizada por instituições/unidades de informação que se preocupam com a preservação, pois consiste na migração periódica da informação para o funcionamento em recursos de *hardware* e *software* mais recentes, ao passo que garantirá a integridade da informação. Na **conversão de formatos**, como o próprio título sugere, os objetos digitais são convertidos para formatos mais recentes, auxiliando e acompanhando a atualização nas migrações, inclusive atuando em consonância a migração de suporte. Quanto a **emulação**, expressa-se em técnicas que preservam e acessam o objeto digital em sua forma original, por meio de programas que simulam o funcionamento do *hardware* e *software* utilizados em sua criação, com o SO da época, por exemplo. Para tanto, é sugerido que haja um encapsulamento⁹ do objeto digital junto ao *software* utilizado em sua criação para facilitar a simulação. Por fim, a **preservação do conteúdo**, como o próprio nome sugere, preocupa-se com a integridade do conteúdo do objeto digital a partir de todas os métodos supracitadas, pois buscam a preservação do conteúdo, isto é, aos quais resumem-se em técnicas de armazenamento, migração conversões e recuperação (ARELLANO, 2004).

Entre os métodos elencados, um se destaca por auxiliar na preservação dos dados descritivos e representativos do objeto digital, tanto à nível estrutural, quanto operacional. Ou seja, destaca-se aqui os metadados de preservação, pois realizam a representação, permitem e

⁹ Baseia-se na ideia de que os objetos preservados devem ser auto descritos e encapsulados em estruturas físicas ou lógicas com todas as informações necessárias para que seja decifrado e compreendido no futuro. (SAYÃO, 2010, p. 10)

garantem a autenticidade dos dados e possibilitam o acesso aos dados entre sistemas, mesmo de características de *software* operacionais diversas.

2.3.1 Metadados

Os metadados são elementos essenciais na preservação digital pois, por meio da descrição dos dados que compõem o objeto digital, como formato, assunto, direitos autorais dentre outros, auxiliam no seu armazenamento, colaborando conseqüentemente na recuperação. Previamente, para compreender o seu conceito, torna-se prudente levar em consideração o contexto de momento, “[...] principalmente pelos desafios do mundo digital, [aos quais] foram redesenhando a ideia puramente descritiva de metadados, criando expansões para o seu conceito com o intuito de abrigar novos propósitos e funções” (SAYÃO, 2010, p. 4). Tais elementos, comumente chamados de dados sobre dados, ou informação sobre informação (NISO, 2004), podem ser compreendidos como simples dados catalográficos ou bibliográficos (CAMPOS, 2007) devido à função de descrição de informações. Sendo assim, de maneira ampla, um metadado é uma informação estruturada que descreve, explica e localiza um objeto digital (NISO, 2004). Em outras palavras, é possível descrever as características de um recurso de informação, abordando produtos e serviços. (SOUZA, 2000).

Destrinchando os conceitos supracitados, os metadados podem ser identificados por tipos, descritivos, estruturais ou administrativos. Sayão (2010, p. 5, grifos do autor) os apresenta conforme os seguintes discursos:

- **Metadados descritivos:** é a face mais conhecida dos metadados, são eles que descrevem um recurso com o propósito de descoberta e identificação; podem incluir elementos tais como título, autor, resumo, palavras-chave e identificador persistente;
- **Metadados estruturais:** são informações que documentam como os recursos complexos, compostos por vários elementos, devem ser recompostos e ordenados. Por exemplo, como as páginas de um livro, digitalizadas separadamente, são vinculadas entre si e ordenadas para formar um capítulo.
- **Metadados administrativos:** fornecem informações que apoiam os processos de gestão do ciclo de vida dos recursos informacionais. Incluem, por exemplo, informações sobre como e quando o recurso foi criado e a razão da sua criação. Nessa categoria, estão metadados técnicos que explicitam as especificidades e dependências técnicas do recurso; inclui também os metadados voltados para apoio à gestão dos direitos relacionados ao recurso.

Segundo Lourenço (2007), outra forma de categorizar os metadados tem como a compreensão de Gilliland-Swetland, ao considerá-los como administrativos, descritivos, de preservação, técnico e de uso. Com base em Gilliland-Swetland (2002), Lourenço (2007, p. 68, grifos nossos) esclarece os tipos de metadados da seguinte forma:

- [...] os **metadados administrativos** são especificamente as informações a respeito da forma de aquisição do recurso informacional pela biblioteca digital, os direitos de reprodução relativos à permissão ou não de cópias e os critérios de seleção para digitalização relativos a acervos físicos que foram digitalizados para formarem uma biblioteca digital;
- [...] os **metadados descritivos** são os metadados utilizados para o tratamento dos recursos informacionais, utilizando para isso o conjunto de técnicas de catalogação, classificação e indexação da Ciência da Informação;
- [...] **metadados de preservação** [...] constituem os metadados utilizados para informar o usuário da biblioteca digital sobre o original físico daquele recurso digitalizado: seu estado de conservação, sua localização física, etc. Com relação aos recursos informacionais gerados na própria *web*, as informações de preservação serão os dados relativos ao organismo e/ou pesquisador responsável pelo seu conteúdo e outras formas de aquisição, no caso deste recurso informacional vir a não estar disponível na biblioteca digital;
- [...] aos **metadados técnicos**, pode-se identificá-los pelas próprias linguagens de marcação que são os softwares para estruturação de bibliotecas digitais. Constituem as informações técnicas de *hardware* e *software* necessárias para desenvolvimento, organização, acesso e visualização dos recursos informacionais de uma biblioteca digital. São os metadados de documentação, de estruturação e de esquematização de uma biblioteca digital;
- [...] **metadados de uso** forneceriam informações a respeito dos usuários que poderiam utilizar esse recurso, sobre quantas vezes esse mesmo recurso foi acessado e por quantos usuários.

A partir das óticas conceitual, funcional e tipológica, indo ao encontro da preservação digital, autores como Arellano (2004) e Sayão (2007) defendem que os metadados de preservação, além de pertencerem ao discurso administrativo, abordam técnicas e meios que proporcionam estocar a informação à posterioridade, ou seja, ao longo do tempo, como é a proposta do padrão OAIS.

O modelo de referência OAIS, traduzido em português pela expressão sistema de arquivamento aberto de informação (SAAI) e desenvolvido pela CCSDS, é uma norma ISO 14721:2003¹⁰ que estabelece diretrizes genéricas de alto nível, cabíveis a diferentes situações, e garante a confiabilidade do que está sendo preservado. Além disso, a norma estabelece um vocabulário padrão (metadados) que afiança a comunicação entre os sistemas que o utilizam, permitindo, inclusive, atingir a maior quantidade de repositórios possíveis. Essas características apresentam o OAIS como sistema aberto, estando a serviço das mais diversas comunidades (Saramago, 2004). Entretanto, deve-se salientar que o modelo OAIS “não é um plano de

¹⁰ Website: <https://www.iso.org/standard/24683.html>.

implementação, mas provê um esquema para a arquitetura e operacionalidade de um repositório digital e definição dos seus metadados” (ARELLANO, 2008, slide 6). Sendo assim, partindo-se do princípio de que o OAIS é um modelo de aplicação lógico-conceitual é possível compreendê-lo como “[...] uma infraestrutura conceitual que descreve o ambiente, as interfaces externas, os componentes funcionais e os objetos de informação, associados com um sistema responsável pela preservação de longo prazo de materiais digitais” (SAYÃO, 2010, p. 13).

A atuação dos metadados, na preservação digital, pode ser identificada a partir da identificação e citação de dois padrões encontrados na literatura sobre o assunto, entre outros, inclusive ambos combinados por vezes: DC, direcionado à descrição e recuperação; e OAI, em linha com o ato de preservar digitalmente.

2.3.1.1 Dublin Core

Os metadados tem papel importante para a descrição dos objetos digitais, com vistas a deixá-los recuperáveis e acessíveis. O *Dublin Core Metadata Initiative* (DCMI), ou apenas iniciativa DC, vem com um propósito voltado aos recursos da *web*, descrevendo e realizando a preservação dos objetos digitais. Conforme Souza (2000, p. 93), DC “[...] pode ser definido como sendo o conjunto de elementos de metadados planejado para facilitar a descrição de recursos eletrônicos”. Trata-se de um esquema simples composto por elementos essenciais que permitem serem mapeados em outros formatos, estruturando uma linguagem padrão, de fácil utilização, e que permite a interoperabilidade¹¹ entre sistemas (SAYÃO, 2010).

De maneira ampla, a expectativa com a implementação do padrão DC seria a de que as pessoas, sem conhecimento sobre catalogação¹², pudessem fazer a descrição dos recursos eletrônicos de forma a permitir uma melhor recuperação pelos sistemas de busca na *web*. (SOUZA, 2000) a partir do uso de 15 elementos¹³ básicos, aos quais, em alcance internacional, objetivam simplicidade, flexibilidade, semântica e interoperabilidade (GRÁCIO, 2002). Usando como base o quadro 1 é possível identificar os elementos do DC e a sua tradução em português, a saber:

¹¹ A interoperabilidade pode ser “[...] entendida como a capacidade que os sistemas de hardware e software têm para se comunicar e operar com outros sistemas no intercâmbio de dados” (MUCHERONI; SILVA, 2011, p. 5).

¹² “Catalogação, ou representação bibliográfica, consiste em um conjunto de informações que simbolizam um registro do conhecimento” (MEY; SILVEIRA, 2009, p.7) a partir de regras como o *Anglo-American Cataloguing Rules* (AACR).

¹³ O esquema é escrito originalmente em língua inglesa.

Quadro 1: Elementos DC

<i>Element</i>	Elemento
<i>Title</i>	Título
<i>Subject</i>	Assunto
<i>Description</i>	Descrição
<i>Language</i>	Idioma
<i>Source</i>	Fonte
<i>Relation</i>	Relação
<i>Coverage</i>	Cobertura
<i>Creator</i>	Criador/Autor
<i>Contributor</i>	Colaborador
<i>Publisher</i>	Publicador/Editor
<i>Rights</i>	Direitos
<i>Date</i>	Data
<i>Type</i>	Tipo
<i>Format</i>	Formato
<i>Identifier</i>	Identificador

Fonte: Adaptado de Rocha (2013).

Apresentando os elementos do DC, **title** (título) é basicamente o nome dado ao recurso, normalmente o qual o objeto é reconhecido, como o título de um livro. O elemento **subject** (assunto) representa as palavras-chave que definem o tema do conteúdo do objeto descrito. Quanto ao elemento **description** (descrição), insere-se uma espécie de relato sobre o conteúdo abordado no objeto descrito, como um resumo. O elemento **language** (idioma) apresenta o idioma do objeto, previamente estabelecido pela *Request For Comments* 1766 (RFC)¹⁴ e ISO 639 e suas variações (639-1; 639-2; 639-3 e 639-5), isto é, uma formalização de duas letras do país e, opcionalmente, o seu código¹⁵. No caso do elemento **source** (fonte) é utilizado quando se deseja referenciar a origem de onde a informação foi extraída como, por exemplo, um livro que contém uma imagem adotada no objeto digital descrito. O elemento **relation** (relação) é adotado quando o objeto descrito tem relação com outro recurso como, por exemplo, ao se descrever um filme baseado em um livro. Sobre o elemento **coverage** (cobertura), contextualiza-se o objeto com relação ao espaço e tempo, como cidades, anos, contextos históricos etc., aos quais influenciaram a produção daquele objeto. O elemento **creator** (criador/autor) nomeia a entidade responsável pela criação do objeto como, por exemplo, pessoas, entidades coletivas, *websites* dentre outros. Quanto ao elemento **collaboration** (colaborador), se identifica entidades responsáveis por auxiliar e fazer contribuições na criação do objeto, como coautores, tradutores etc. O elemento **publisher** (publicador/editor) seria a entidade responsável por tornar o objeto conhecido, ou seja, edita e torna pública a obra. **Rights** (direitos) vão ao encontro dos direitos

¹⁴ Website: <https://www.ietf.org/rfc/rfc1766.txt>

¹⁵ Exemplo: Português ou *portuguese* (inglês): 639-1 pt; 639-2/639-5: por; 639-3: por (SIL, 2015).

ligados ao objeto como, por exemplo, o *copyright*¹⁶ de um recurso. O elemento *date* (data) é responsável por trazer a data relacionada a algum período no ciclo do objeto, como sua data de criação ou publicação, por exemplo. O *type* (tipo) está relacionado com a essência de existência do objeto com fim de especificá-lo como, por exemplo, fotografias, pinturas e desenhos enquanto tipos de imagem. No caso do elemento *format* (formato) refere-se a extensão digital do objeto, ao qual representa a sua apresentação como, por exemplo, jpeg¹⁷, png¹⁸ enquanto extensões de imagens. Por último o *identifier* (identificador) é uma informação não ambígua que identifica o objeto em um determinado contexto, de maneira única, como, por exemplo, o *International Standard Book Number (ISBN)*¹⁹ de um livro (GRACIO, 2002).

Tomando como base a compreensão de Sayão (2007, *slide* 6), os metadados estão inseridos no “[...] objeto digital – inscritos na sua codificação, como é comum nos documentos HTML; ou podem estar armazenados separadamente, estruturados em base de dados e vinculados ao objeto que eles descrevem, facilitando a busca e a recuperação”. A representação da mencionada codificação, no caso dos elementos do DC, pode ser observada em uma estrutura *web*, especificamente de um registro da biblioteca digital de monografias da UnB (BDM)²⁰ referente à um trabalho de conclusão de curso (TCC) intitulado “O uso do tesouro na arquitetura da informação em *Websites*” (figura 4). Conforme a figura 4, sete campos estão objetivados a guardar e identificar o supracitado TCC.

¹⁶ Entendido por direito do autor, ou direito automático, o copyright é o direito conferido ao indivíduo ou Instituição sobre sua criação, representado pelo símbolo Copyright©, seguido do ano de publicação e do nome do autor ou empresa responsável pelo depósito (COPYRIGHT, 2017).

¹⁷ *Joint photographic experts group*.

¹⁸ *Portable Network Graphics*.

¹⁹ O ISBN “[...] é um sistema internacional padronizado que identifica numericamente os livros segundo o título, o autor, o país, a editora, individualizando-os inclusive por edição. Utilizado também para identificar software, seu sistema numérico é convertido em código de barras, o que elimina barreiras linguísticas e facilita a circulação e comercialização das obras” (ISBN, 2017).

²⁰ Website: <http://bdm.unb.br/>

Figura 4: Registro simples na BDM²¹

Título:	O uso do tesouro na arquitetura da informação em Websites
Autor(es):	Rodrigues, Agostinha Maria
Orientador(es):	Silva, Márcio Bezerra da
Assunto:	Tesouros Arquitetura da informação Sites da Web Representação da informação
Data de apresentação:	5-Dez-2014
Data de publicação:	22-Mai-2015
Referência:	RÓDRIGUES, Agostinha Maria. O uso do tesouro na arquitetura da informação em Websites. 2014. 156 f., il. Monografia (Bacharelado em Biblioteconomia)—Universidade de Brasília, Brasília, 2014.
Resumo:	<p>Verifica a utilização do tesouro na representação e organização da informação na Web pelos arquitetos da informação. Contribui para a discussão sobre o uso do tesouro na Arquitetura da Informação (AI) em Website indagando sobre o perfil do arquiteto da informação, como se dá o uso do tesouro na AI em Website e a popularidade do tesouro entre os profissionais em questão. Disserta o trabalho a partir de uma fundamentação teórica que abrange a tríplice dado, informação e conhecimento, representação temática da informação, linguagens documentárias (LD), tesouro e AI em Website com ênfase nos sistemas de rotulação e organização. Constitui-se de um estudo exploratório no qual foi adotada a técnica de pesquisa bibliográfica e abordagem de coleta de dados quantitativa e qualitativa. Confirma que a AI em Website é uma área multidisciplinar, formalizando um perfil profissional concentrado nos campos da tecnologia da informação (TI), designer, marketing e organização e representação da informação. Consta-se ainda que o perfil é de um profissional jovem com aprendizado em AI adquirida no ambiente de trabalho e com pouco tempo de experiência, atuando na área a menos de cinco anos. Destaca que as LD não são utilizadas por boa parte dos arquitetos da informação com predileção pela taxonomia, reflexo do pouco conhecimento sobre tesouro, tanto à metodologia quanto pelo uso, especialmente na elaboração dos sistemas de rotulação e organização da AI em Website. Conclui-se que o tesouro não é um instrumento popular entre os arquitetos da informação, exigindo integralização das LD nos cursos de AI, assim como mais discussões sobre a própria AI nos cursos de graduação em Biblioteconomia, acreditando que a intercepção gerada pela nossa proposta resultará em estudos e pesquisas empíricas que contribuirão no aprimoramento de ambos os sistemas aqui enfocados e na construção de tesouros em Websites, não mais exclusivos das bibliotecas, mas projetados com fins acadêmicos, profissionais e/ou comerciais.</p> <p>_____ ABSTRACT</p> <p>Checks the use of the thesaurus in the representation and organization of information on the Web for information architects. Contributes to the discussion on the use of the thesaurus in Information Architecture (IA) in Website inquiring about the information architect profile, how is the use of the thesaurus in AI in the Website and the popularity of the thesaurus among professionals concerned. Dissertation work from a theoretical foundation that covers the triple data, information and knowledge, thematic representation of information, documentary languages (DL), thesaurus and AI in Website with emphasis on labeling and organization systems. It consists of an exploratory study that adopted the literature technique and approach a collection of quantitative and qualitative data. Confirms that the AI in Website is a multidisciplinary formalizing a professional profile in which several areas contribute and concentrated in the fields of area concentrated in the fields of information technology (IT), designer, marketing and organization and representation of information. Noted also that the profile the profile is of a young professional with learning in AI acquired in the workplace and soon experience, working in less than five years. Notes that the DL is not used for much of the information architects mainly on the taxonomy, reflecting the lack of knowledge about the thesaurus, both the methodology and the use, particularly in the context of labeling systems and organization of AI in Website. Concludes that the thesaurus is not a popular instrument among information architects, requiring insertion of LD in AI courses, as well as further discussion of the AI itself in undergraduate courses in Library, believing that interception generated by our proposal will result in studies and empirical research that will contribute to the improvement of both systems focused here and the construction of thesauri in Websites, no longer unique to libraries, but designed with academic purposes, professional and/or commercial.</p>
Informações adicionais:	Monografia (graduação)—Universidade de Brasília, Faculdade de Ciência da Informação, Curso de Graduação em Biblioteconomia, 2014.
Aparece na Coleção:	Biblioteconomia

Fonte: BDM (2018)

Na representação do TCC, os campos são: **título**, local onde insere-se o nome do trabalho; **autores**, campo que recebe os responsáveis pela elaboração do TCC; **orientador**, destinado ao nome do orientador do trabalho; **assunto**, ou seja, expressões que representam os principais assuntos fidedignos ao recurso; **datas de apresentação e de publicação**, sendo a primeira representativa ao dia da apresentação do TCC e a segunda correspondente ao dia em que o trabalho foi publicado na BDM; **referência bibliográfica**, ao qual apresenta as referências literárias adotadas para a fundamentação e escrita do TCC; **informações adicionais**, isto é, inclusão de informações extras como o nome do curso, a faculdade/departamento responsável e o ano de conclusão do curso; e **aparece na coleção**, sendo um campo que indica em qual coleção (área do conhecimento) este recurso pertence dentro da BDM.

²¹ Website: <http://bdm.unb.br/handle/10483/10476>

A figura 5 mostra um exemplo de registro na ótica do usuário, ou seja, apresentando os campos com rótulos legíveis ao público, entretanto, na visão do sistema (preenchido pelo bibliotecário), os metadados se apresentam de maneira diferente, contendo mais elementos (campos) e mais detalhes, aos quais são importantes para o sistema, mas não são relevantes para a exibição ao usuário. Essa forma diferenciada se caracteriza pela qualificação dos metadados, compreendidos e apresentados basicamente em duas classes:

- 1- Elemento de refinamento: esses qualificadores dão mais especificidade a um elemento, detalhando-o melhor”.
- 2- Esquema de codificação: esses qualificadores identificam esquemas para o valor do elemento. Incluem vocabulário controlado e notações formais de representação (sistemas de classificação). (GRÁCIO, 2002, p. 68)

Conforme o quadro 2 é possível identificar os elementos do padrão DC e os seus correspondentes qualificadores segundo as duas classes anteriormente mencionadas, ou seja, quanto aos elementos de refinamento e quanto a codificação. Os qualificadores são:

Quadro 2: Qualificadores do padrão DC

Elementos	Elementos de refinamento	Elementos do esquema de codificação
<i>Title</i>	<i>Alternative</i> (alternativo - título alternativo ou obra alternativa)	-
<i>Subject</i>	-	<i>Library of Congress Subject Headings (LCSH)</i> <i>Medical Subject Headings (MeSH)</i> <i>Dewey Decimal Classification (DDC)</i> <i>Library of Congress Classification (LCC)</i> <i>Universal Decimal Classification (UDC)</i>
<i>Description</i>	<i>Table Of Contents</i> (tabela de conteúdos - uma lista de subunidades do conteúdo do Recurso); <i>Abstract</i> (resumo).	-
<i>Language</i>	-	ISO 639-1, ISO 639-2, ISO 639-3 e ISO 639-5 (códigos para representação de nomes de linguagens.); RFC 1766 (código para identificação da linguagem que especifica duas letras da norma ISO 639, seguidas opcionalmente de duas letras para o país da norma ISO 3166).
<i>Source</i>	-	<i>Uniform Resource Identifier (URI)</i>
<i>Relation</i>	<i>Is Version Of</i> (é uma versão de); <i>Has Version</i> (tem uma versão de); <i>Is Replaced By</i> (é substituído por); <i>Replaces</i> (substitui); <i>Is Required By</i> (é requerido por); <i>Requires</i> (requer); <i>Is Part Of</i> (e parte de); <i>Has Part</i> (tem parte); <i>Is Referenced By</i> (é referenciado por); <i>References</i> (faz referência); <i>Is Format Of</i> (e formatado do); <i>Has Format</i> (possui o formato).	URI

Coverage	Spatial (espacial - contém características espaciais do conteúdo intelectual do recurso)	DCMI Point (identifica um ponto no espaço usando suas coordenadas geográficas); ISO 3166 ; DCMI Box (identifica uma região do espaço usando seus limites geográficos); <i>Getty Thesaurus of Geographic Names (TGN)</i> ;
	Temporal (temporal - contém características temporais do conteúdo intelectual do recurso)	DCMI Period (uma especificação de limites de um intervalo de tempo); W3C-DTF ²² (regras de codificação para data e hora da W3C ²³ baseadas na norma ISO 8601).
Creator	-	-
Contributor	-	-
Publisher	-	-
Rights	-	<i>Creative Commons</i> ²⁴
Date	Created (criação – data de criação); Valid (validade – intervalo/tempo de validade do recurso); Available (disponível – intervalo/tempo que o recurso se tornou ou está disponível); Issued (distribuição – data de distribuição formal do recurso); Modified (modificado – data de modificação do recurso) etc.	DCMI Period (uma especificação de limite para intervalo de Tempo); W3C-DTF (regras de codificação do W3C para data e hora baseado na norma ISO 8601).
	-	DCMI Type (uma lista de tipos usada para categorizar a natureza ou o gênero do conteúdo do recurso. Essa lista é composta pelos seguintes valores: collection, dataset, event, image, interactive resource, service, software, sound ou text)
Format	Extent (medida – tamanho ou duração de um recurso)	-
	Medium (meio – material ou transporte físico do recurso)	IMT (esquema de tipos de meio de armazenamento da Internet – <i>Internet Media Types</i>) – como: <i>text/html, application/pdf, image/jpegvideo/mpeg</i> etc.)
Identifier	-	URI

Fonte: Adaptado de Grácio (2002), Dublin core metadata initiative²⁵ (2017) e Library of Congress²⁶ (2017).

Após a visualização de um registro simples (figura 4), voltado a ótica do usuário, tem-se na figura 5 o mesmo exemplo de material registrado, porém, apresentando a estrutura do registro na BDM na perspectiva do sistema, ou seja, ilustrando os campos em DC, inclusive alguns estando qualificados para melhorar o refinamento dos mesmos. Os correspondentes campos em DC podem ser identificados na figura 5, sendo eles: dc.description, dc.rights, dc.title e dc.type. Entretanto, vale destacar a presença de campos qualificados, pensando em uma

²² World Wide Web Consortium - Date and Time Formats.

²³ Da expressão World Wide Web Consortium, o W3C é um conglomerado “[...] internacional no qual organizações filiadas, uma equipe em tempo integral e o público trabalham juntos para desenvolver padrões para a Web” (W3C BRASIL, 2011c).

²⁴ Licenças autorais para compartilhamento, complementares ao direito do autor (CREATIVE COMMONS BR, 2018).

²⁵ Da sigla DCMI.

²⁶ Da sigla LC.

representação de informação mais fidedigna a obra e melhoria nos processos de recuperação, destinados aos campos a saber: *dc.contributor.advisor* e *dc.contributor.author*; *dc.identifier.citation* e *dc.identifier.uri*; *dc.description.abstract*; *dc.subject.keywords*; *dc.date.accessioned*, *dc.date.available*, *dc.date.issued* e *dc.date.submitted*; e *dc.language.iso*.

A título de exemplo, usar-se-á o elemento *date* para compreender o significado de campos qualificados. Referente a datas, o item sofre quatro qualificações, ou seja, especificar as datas associadas ao documento pela data de acesso (*dc.date.accessioned*), pela data de avaliação (*dc.date.available*), pela data de publicação (*dc.date.issued*) e pela data de apresentação (*dc.date.submitted*).

Figura 5: Registro qualificado na BDM²⁷

Registro completo		
Campo Dublin Core	Valor	Lingua
dc.contributor.advisor	Silva, Márcio Bezerra da	-
dc.contributor.author	Rodrigues, Agostinha Maria	-
dc.identifier.citation	RODRIGUES, Agostinha Maria. O uso do tesouro na arquitetura da informação em Websites. 2014. 156 f., il. Monografia (Bacharelado em Biblioteconomia)—Universidade de Brasília, Brasília, 2014.	en
dc.description	Monografia (graduação)—Universidade de Brasília, Faculdade de Ciência da Informação, Curso de Graduação em Biblioteconomia, 2014.	en
dc.description.abstract	Verifica a utilização do tesouro na representação e organização da informação na Web pelos arquitetos da informação. Contribui para a discussão sobre o uso do tesouro na Arquitetura da Informação (AI) em Website indagando sobre o perfil do arquiteto da informação, como se dá o uso do tesouro na AI em Website e a popularidade do tesouro entre os profissionais em questão. Disserta o trabalho a partir de uma fundamentação teórica que abrange a trílice dado, informação e conhecimento, representação temática da informação, linguagens documentárias (LD), tesouro e AI em Website com ênfase nos sistemas de rotulação e organização. Constitui-se de um estudo exploratório no qual foi adotada a técnica de pesquisa bibliográfica e abordagem de coleta de dados quantitativa e qualitativa. Confirma que a AI em Website é uma área multidisciplinar, formalizando um perfil profissional concentrado nos campos da tecnologia da informação (TI), designer, marketing e organização e representação da informação. Constata-se ainda que o perfil é de um profissional jovem com aprendizado em AI adquirida no ambiente de trabalho e com pouco tempo de experiência, atuando na área a menos de cinco anos. Destaca que as LD não são utilizadas por boa parte dos arquitetos da informação com predileção pela taxonomia, reflexo do pouco conhecimento sobre tesouro, tanto à metodologia quanto pelo uso, especialmente na elaboração dos sistemas de rotulação e organização da AI em Website. Conclui-se que o tesouro não é um instrumento popular entre os arquitetos da informação, exigindo integralização das LD nos cursos de AI, assim como mais discussões sobre a própria AI nos cursos de graduação em Biblioteconomia, acreditando que a intercepção gerada pela nossa proposta resultará em estudos e pesquisas empíricas que contribuirão no aprimoramento de ambos os sistemas aqui enfocados e na construção de tesouros em Websites, não mais exclusivos das bibliotecas, mas projetados com fins acadêmicos, profissionais e/ou comerciais. _____ ABSTRACT	en
dc.description.abstract	Checks the use of the thesaurus in the representation and organization of information on the Web for information architects. Contributes to the discussion on the use of the thesaurus in Information Architecture (IA) in Website inquiring about the information architect profile, how is the use of the thesaurus in AI in the Website and the popularity of the thesaurus among professionals concerned. Dissertation work from a theoretical foundation that covers the triple data, information and knowledge, thematic representation of information, documentary languages (DL), thesaurus and AI in Website with emphasis on labeling and organization systems. It consists of an exploratory study that adopted the literature technique and approach a collection of quantitative and qualitative data. Confirms that the AI in Website is a multidisciplinary formalizing a professional profile in which several areas contribute and concentrated in the fields of area concentrated in the fields of information technology (IT), designer, marketing and organization and representation of information. Notes also that the profile is of a young professional with learning in AI acquired in the workplace and soon experience, working in less than five years. Notes that the DL is not used for much of the information architects mainly on the taxonomy, reflecting the lack of knowledge about the thesaurus, both the methodology and the use, particularly in the context of labeling systems and organization of AI in Website. Concludes that the thesaurus is not a popular instrument among information architects, requiring insertion of LD in AI courses, as well as further discussion of the AI itself in undergraduate courses in Library, believing that interception generated by our proposal will result in studies and empirical research that will contribute to the improvement of both systems focused here and the construction of thesauri in Websites, no longer unique to libraries, but designed with academic purposes, professional and/or commercial.	en
dc.rights	Acesso Aberto	en
dc.subject.keyword	Tesouros	en
dc.subject.keyword	Arquitetura da informação	en
dc.subject.keyword	Sites da Web	en
dc.subject.keyword	Representação da informação	en
dc.title	O uso do tesouro na arquitetura da informação em Websites	en
dc.type	Trabalho de Conclusão de Curso - Graduação - Bacharelado	en
dc.date.accessioned	2015-05-22T12:08:57Z	-
dc.date.available	2015-05-22T12:08:57Z	-
dc.date.issued	2015-05-22T12:08:57Z	-
dc.date.submitted	2014-12-05	-
dc.identifier.uri	http://bdm.unb.br/handle/10483/10476	-
dc.language.iso	Português	en
Aparece na Coleção:	Biblioteconomia	

Fonte: BDM (2018).

²⁷ Website: <http://bdm.unb.br/handle/10483/10476?mode=full>.

Os campos em DC e exemplificados na BDM, por trás da interface e com fins de visualização do usuário, são construídos conforme o uso de linguagens de marcação, especificamente a partir do *HyperText Markup Language* (HTML), ou seja, “[...] uma linguagem de marcação utilizada para criar páginas acessadas a partir de um navegador” (MILETTO; BERTAGNOLLI, 2014, p. 62). O HTML se trata de uma linguagem que viabiliza a publicação de texto, imagem, vídeo, áudio dentre outros na *web*, tornando-se um padrão (W3C BRASIL, 2011c). Além disso, o HTML “foi criado para ser uma linguagem independente de plataformas, browsers e outros meios de acesso” (W3C BRASIL, 2011c, p. 8), apresentando-se como flexível e acessível. Entre as suas características, destaca-se a criação dos chamados hipertextos²⁸, aos quais viabilizarão a navegação pelas páginas (MILETTO; BERTAGNOLLI, 2014). Quanto a um exemplo de código em HTML, enquanto um registro codificado em linguagem de marcação, a figura 6 ilustra o registro da BDM sobre o TCC “O uso do tesauro na arquitetura da informação em *Websites*”.

Figura 6: Registro em DC/HTML com destaque nos qualificadores

```

<!DOCTYPE html">
<html>
<head>
<title> Registro em HTML/DC </title>
<meta name="DC.Description" content="Monografia (graduação)– Universidade de Brasília, Faculdade de
Ciência da Informação, Curso de Graduação em Biblioteconomia, 2014.">
<meta name="DC.Title" content="O uso do tesauro na arquitetura da informação em Websites">
<meta name="DC.Creator" content="Rodrigues, Agostinha Maria">
<meta name="DC.Subject" content="Tesouros">
<meta name="DC.Contributor" content="Silva, Márcio Bezerra da">
<meta name="DC.Date.accessioned" content="2015-05-22T12:08:57Z">
<meta name="DC.Date.available" content="2015-05-22T12:08:57Z">
<meta name="DC.Date.issued" content="2015-05-22T12:08:57Z">
<meta name="DC.Date.submitted" content="2014-12-05">
<meta name="DC.Type" content="Trabalho de Conclusão de Curso -
Graduação - Bacharelado">
<meta name="DCTERMS.BibliographicCitation" content="RODRIGUES, Agostinha Maria.
O uso do tesauro na arquitetura da informação em Websites. 2014. 156 f., il.
Monografia (Bacharelado em Biblioteconomia)–Universidade de Brasília, Brasília, 2014.">
<meta name="DC.Language" content="Português">
<meta name="DC.Rights" content="Acesso Aberto">
</head>
<body>
<br><br>
Título: O uso do tesauro na arquitetura da informação em Websites<br>
Autor(es): Rodrigues, Agostinha Maria<br>
Resumo: Verifica a utilização do tesauro na representação e organização da informação na Web pelos
arquitetos da informação.
Contribui para a discussão sobre o uso do tesauro na Arquitetura da Informação (AI) em Website
indagando sobre o perfil do arquiteto da informação,[...].<br>
Informações adicionais: Monografia (graduação)–Universidade de Brasília,
Faculdade de Ciência da Informação, Curso de Graduação em Biblioteconomia, 2014.
Aparece na Coleção: Biblioteconomia<br />
</body>
</html>

```

Elaborado pela autora (2018).

²⁸ Hipertextos “[...] são conjuntos de elementos – ou nós – ligados por conexões. Estes elementos podem ser palavras, imagens, vídeos, áudio, documentos etc. Estes elementos conectados formam uma grande rede de informação” (W3C BRASIL, 2011c, p.7).

No escopo dos metadados, e com o intuito de compreender melhor a forma como o código se expressa, faz-se necessário mencionar um elemento básico quando se utiliza o padrão DC na linguagem HTML: *meta*. O elemento *meta* “[aponta] os metadados que não podem ser definidos usando outro elemento HTML” (MDM WEB DOCS. RECURSOS DE DESENVOLVEDORES PARA DESENVOLVEDORES, 2017), ou seja, que não estarão no corpo de informações da página (*body*²⁹). Na sintaxe do código, especificamente no título (*title*) da página, insere-se o elemento *meta* e em seguida o nome do metadado propriamente dito, identificando-o por *name*. Complementando, informa-se o atributo *content*, que por sua vez é destinado a informar o dado em si. Como exemplo de uso do elemento *meta* em um código em HTML, extraído da figura 6, tem-se duas linhas referentes a data (*date*). Observando a figura 7 é possível verificar que a primeira linha traz a data de publicação (*date.issued*) do trabalho, enquanto que a segunda diz respeito a data de apresentação (*date.submitted*). Como verificado nas datas, ambas possuem os elementos *meta name*, seguidos do metadado qualificado (*issued* e *submitted*) e finalizadas com os dados (data/horário) inseridos em *content*.

Figura 7: O elemento *meta name* no código em HTML

```
<meta name="DC.Date.issued" content="2015-05-22T12:08:57Z">
<meta name="DC.Date.submitted" content="2014-12-05">
```

Elaborado pela autora (2018).

Entretanto, o HTML apenas apresenta as informações no sentido visual, sem qualquer sentido semântico, o que pode dificultar na recuperação de informação e impossibilitar a interoperabilidade entre sistemas. Pensando nessa questão, adota-se a linguagem de marcação *Extended Markup Language* (XML). Conforme Almeida (2002) é uma linguagem de marcação mais flexível que a HTML e que, por não ser predefinida, permite ao autor especificar a estruturação dos dados no documento, e ainda o definir semanticamente. De acordo com Pereira (2009) é este o modo recomendado pela W3C para casos em que os dados devam ser organizados lógico-hierarquicamente e de acordo com uma sintaxe semântica básica, tornando a linguagem bastante portátil no sentido de poder ser usada em diferentes ambientes. Sayão (2010) afirma que, desta maneira, os dados não dependerão de um *software* específico, sofrendo menos as consequências das obsolescências tecnológicas. A própria WC3 (2015c) apresenta o XML como um formato superior ao HTML, pois o XML fornece vantagens como uma marcação mais detalhada, facilitando o encontro de erros, uma legibilidade favorável,

²⁹ Na *tag body* estará “[...] todo o conteúdo de um documento HTML, como texto, hiperlinks, imagens, tabelas, listas etc.” (W3SCHOOLS, 2018c, tradução nossa) que será apresentado (visualmente) em uma página web.

facilitando a compreensão do formato por outros *softwares*, por exemplo, tornando-o como o formato mais usado em todo o mundo.

De forma ampla, o XML é uma linguagem de marcação extensível diante da sua capacidade de permitir a definição de seus elementos de marcação de forma semântica, adaptando-se as necessidades do ambiente/sistema. Sendo assim, reescrevendo o código da figura 6, o registro da BDM teria a seguinte versão em XML, apresentada na figura 8:

Figura 8: Registro em DC/XML³⁰

```

<?xml version="1.0" encoding="UTF-8"?>
<metadata
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:dcterms="http://purl.org/dc/terms/">
  <dc:title>O uso do tesouro na arquitetura da informação em Websites</dc:title>
  <dc:creator>Rodrigues, Agostinha Maria</dc:creator>
  <dc:subject>Tesauros</dc:subject>
  <dc:description>Monografia (graduação)-Universidade de Brasilia, Faculdade de
  Ciência da Informação, Curso de Graduação em
  Biblioteconomia, 2014.</dc:description>
  <dc:contributor>Silva, Márcio Bezerra da</dc:contributor>
  <dcterms:dateAccessioned>2015-05-22T12:08:57Z</dcterms:dateAccessioned>
  <dcterms:dateAvailable>2015-05-22T12:08:57Z</dcterms:dateAvailable>
  <dcterms:dateIssued>2015-05-22T12:08:57Z</dcterms:dateIssued>
  <dcterms:dateSubmitted>2014-12-05</dcterms:dateSubmitted>
  <dc:type>Trabalho de Conclusão de Curso - Graduação - Bacharelado</dc:type>
  <dcterms:bibliographicCitation>RODRIGUES, Agostinha Maria. O uso do tesouro na
  arquitetura da informação em Websites. 2014. 156 f., il.
  Monografia (Bacharelado em Biblioteconomia)-Universidade de Brasília, Brasília,
  2014.</dcterms:bibliographicCitation>
  <dc:language>Português</dc:language>
  <dc:rights>Acesso Aberto</dc:rights>
</metadata>
</xml>

```

Elaborado pela autora (2018).

No *script* em XML (figura 8) é possível verificar que a apresentação das *tags* (campos) é definida pelo próprio programador, ainda dentro da estrutura do código, não sendo necessário, por exemplo a utilização da *tag meta name*, assim como exige-se no caso da linguagem de marcação HTML. Portanto, como exemplo, a figura 9 destaca duas linhas de código retiradas do registro em XML (figura 8), correspondentes ao item *date*.

³⁰ Deve-se destacar que no exemplo existem elementos chamados de *namespaces*. Referem-se a “[...] uma coleção de nomes que podem ser usados como nomes de atributo ou elemento em um documento XML. O namespace qualifica nomes de elemento exclusivamente na Web para evitar conflitos entre os elementos com o mesmo nome. O namespace é identificado por um Uniform Resource Identifier (URI), qualquer um localizador URL (Uniform Resource), ou um nome de recurso uniforme (URN) [...]” (MICROSOFT, 2018c). No código da figura 8 existem três XML *namespace* (XMLNS). O primeiro, ou seja, o XMLNS referente ao XML Schema-instance (XSI), destina-se a versão do XMLNS utilizada no registro (`xmlns:xsi=http://www.w3.org/2001/XMLSchema-instance`). O segundo é o XMLNS corresponde aos elementos do DC (`xmlns:dc="http://purl.org/dc/elements/1.1/"`). O terceiro e último também se destina ao DC, neste caso, sobre o vocabulário do padrão (`xmlns:dcterms="http://purl.org/dc/terms/"`).

Figura 9: O elemento *date* no código em XML

```
<dcterms:dateIssued>2015-05-22T12:08:57Z</dcterms:dateIssued>
<dcterms:dateSubmitted>2014-12-05</dcterms:dateSubmitted>
```

Elaborado pela autora (2018).

Para uma melhor representação dos dados, os campos referentes a datas foram qualificados, com fins de que as datas cadastradas sejam confundidas, que por sua vez poderão gerar resultados inconsistentes na recuperação. Neste caso, a partir da flexibilidade do XML, além da criação de *tags* referentes a datas, como são distintas, também foi possível qualificá-las, ou seja, uma *tag* específica para a data de apresentação do trabalho (*date.submitted*) e outra correspondente a data de publicação (*date.issued*) no sistema, que por sua vez estão em consonância com um XMLNS específico. Sendo assim, é possível observar que os elementos em DC (figura 8) estão qualificados, fortalecendo a compreensão de que o desenvolvimento do código em XML não segue uma estrutura rígida, baseando-se no interesse e necessidade do ambiente/sistema. Diante dessa não rigidez, Sayão (2007) afirma que os elementos básicos do DC estruturam um núcleo de representação, que por sua vez permite a interoperabilidade entre vários sistemas, o que incentiva a sua aplicação em um grande número de comunidades via “perfis de aplicação”, inclusive incorporados a outros esquemas de metadados.

Um perfil de aplicação de metadados (PAM) constitui-se em um esquema de metadados composto por vários elementos pertencentes a mais de um *namespace*, ocasionando uma otimização e personalização do esquema elaborado/montado por parte de seus implementadores, a fim de aplicá-lo em um local específico (HEERY; PATEL, 2000). Os PAM são usados “[...] como uma maneira de compreender as diferentes relações que os implementadores e gerentes de *namespaces* têm de acordo ao esquema de metadados e as diferentes maneiras de usar e desenvolver esquemas” (HEERY; PATEL, 2000, tradução nossa).

Em suma, os conceitos, as características e os métodos de preservação digital elencados na presente seção do TCC estão em linha aos padrões de metadados, o que formalizou a base literária necessária para a identificação e posterior compreensão de conceitos, de aplicações e de uso dos métodos de preservação digital possivelmente identificados na pesquisa.

3 METODOLOGIA

A ciência evolui e progride a partir da produção de conhecimentos, por intermédio de indivíduos como professores e pesquisadores. Em uma comum realidade que se dá na academia, essa produção de conhecimentos constrói-se a partir do que se chama de pesquisa científica, definida como um “[...] conjunto de procedimentos sistemáticos, baseado no raciocínio lógico, que tem por objetivo encontrar soluções para problemas propostos, mediante a utilização de métodos científicos” (ANDRADE, 1999, p. 103). Trata-se de uma pesquisa de ordem intelectual e que objetiva a busca por conhecimento podendo tal pesquisa ser denominada “pura” ou “fundamental” (ANDRADE, 1999), estando, normalmente, ligada ao desenvolvimento de trabalhos acadêmicos, por exemplo.

Na presente investigação, enquanto uma pesquisa científica, utiliza-se de um percurso metodológico constituído de um método, técnicas de pesquisa, abordagens de coleta de dados e campos de investigação, com vistas a identificar e compreender o uso de métodos de preservação digital na contemporaneidade, bem como elencar ambientes digitais que fazem uso da preservação digital dos seus dados.

3.1 CARACTERIZAÇÃO DA PESQUISA

Em linha com a presente pesquisa está o método dedutivo, pois este tipo de raciocínio “[...] tem o objetivo de explicar o conteúdo das premissas” (SILVA; MENEZES, 2005, p. 25) ao analisar a estrutura da cadeia lógica de ideias, do geral ao específico (SILVA; MENEZES, 2005). Diferentemente da indução “[...] o método dedutivo procura transformar enunciados complexos, universais, em particulares. A conclusão sempre resultará em uma ou várias premissas, fundamentando-se no raciocínio dedutivo” (OLIVEIRA, 2000, p. 62).

Quanto a sua natureza, este TCC configura-se como uma pesquisa básica já que “objetiva gerar conhecimentos novos úteis para o avanço da ciência sem aplicação prática prevista. Envolve verdades e interesses universais” (SILVA; MENEZES, 2005, p. 20). Com relação aos objetivos, a investigação é classificada como exploratória, pois pretende-se “proporcionar maior familiaridade com o problema com vistas a torná-lo explícito ou a construir hipóteses” (SILVA; MENEZES, 2005, p. 21), subsidiada, em muitos casos, por levantamento bibliográfico e análise de exemplos, com fins de auxiliarem na compreensão do tema (SILVA; MENEZES, 2005). Ou seja, a pesquisa exploratória está em linha com o levantamento bibliográfico ao “[...] proporcionar maiores informações sobre determinado

assunto; definir objetivos ou formular as hipóteses de uma pesquisa ou descobrir novo tipo de enfoque para o trabalho que se tem em mente” (OLIVEIRA, 2000c, p. 106).

Com relação aos procedimentos técnicos, em consonância ao caráter exploratório, adotou-se a pesquisa bibliográfica, justificada pelo desejo de “[...] conhecer as diferentes formas de contribuição científica que se realizaram sobre determinado assunto ou fenômeno” (OLIVEIRA, p.113, 2000).

Sobre a abordagem do problema, o trabalho caracteriza-se como do tipo qualitativa, ao passo que auxilia na construção de uma conclusão própria a partir de interpretações específicas do investigador. Para Andrade (1999), ao contrário da quantitativa, a versão qualitativa não se utiliza de dados estatísticos para a análise de um problema. Ao invés disso, faz uso de uma série de leituras acerca do tema da pesquisa para explicar o que os autores e especialistas estão publicando, e, a partir deste ponto, estabelecer correlações com os objetivos da pesquisa.

3.2 CAMPO DA PESQUISA

Por se tratar de uma pesquisa fundamentalmente teórica, o campo da pesquisa é constituído pelos ambientes onde ocorrem aplicações e usos de métodos de preservação digital identificados na literatura. Para tanto, a coleta de dados ocorreu especificamente em bibliografias a partir das seguintes expressões de busca: “memória”; “preservação”; “preservação digital”; “estratégias de preservação digital”; “métodos de preservação digital!; “metadados”; “metadados de preservação”; “ciência da informação e preservação digital”; “ciência da informação e metadados”; “modelos de preservação digital”; “modelo OAIS”; “OAIS e aplicações”; “PREMIS”; “modelo PREMIS”; “modelo MODS”; “MODS *library of congress*”; “METS”; “METS *library of congress*”; e “METS” AND “aplicação”.

3.3 ETAPAS DA PESQUISA

Primeiramente, foram pesquisados conceitos (básicos) na crença de que ajudariam na definição e compreensão do tema da pesquisa e, conseqüentemente, na contextualização dos objetivos propostos. Neste caso, as pesquisas se deram sobre memória; preservação e conservação; explosão da informação; e avanço das TIC. Em seguida, de posse de uma compreensão sobre o tema da pesquisa, que é a preservação digital, foi possível realizar o levantamento bibliográfico a partir de discussões sobre memória, bem como sobre conceitos e métodos de preservação digital, especialmente sobre metadados, como um método estrutural

de preservação digital. A partir desse levantamento teórico foi admissível identificar aplicações e usos de métodos de preservação digital, formalizando a coleta de dados. Na sequência, ocorreram o tratamento, a análise e a interpretação dos dados coletados a fim de responder aos problemas da pesquisa. Enfim, resumidamente, é possível observar que a pesquisa percorreu o seguinte caminho:

1. Definição do tema;
2. Levantamento bibliográfico e construção da fundamentação teórica mediante pesquisa bibliográfica, basicamente em livros, artigos científicos e apresentações (*slides*) de especialistas sobre o tema em congressos;
3. Coleta de dados sobre aplicações e usos de métodos de preservação digital, identificados em artigos científicos, apresentações (*slides*) de especialistas sobre o tema e páginas *web*, com destaque aos padrões de metadados para a preservação dos dados de objetos digitais;
4. Tratamento, análise e interpretação dos dados coletados.

4 APLICAÇÕES E USO DE MÉTODOS DE PRESERVAÇÃO DIGITAL

Com base na pesquisa literária que formalizou a presente investigação, centrada em livros (4), artigos científicos (20), apresentações (*slides*) de especialistas sobre o tema (2) em congressos (5) e páginas *web* especializadas (36), como *websites* oficiais de iniciativas/projetos de preservação digital e *websites* governamentais, foram identificados exemplos de métodos de preservação digital, especificamente: o modelo de referência OAIS, que faz uso de padrões de metadados, anteriormente citado e identificado em uma quantidade significativa entre os materiais recuperados; o padrão de metadados *Preservation Metadata: Implementation Strategies* (PREMIS); o padrão de metadados *Metadata Object Description Schema* (MODS); e o padrão de metadados *Metadata Encoding and Transmission Standard* (METS).

No momento de levantamento bibliográfico, expressões como “preservação digital”, “estratégias de preservação digital” e “métodos de preservação digital” foram utilizadas no processo de busca, recuperando-se materiais como livros, artigos e apresentações (*slides*). Em um segundo momento, partindo dos exemplos encontrados na primeira etapa, foi feito um refinamento na busca, incidindo na recuperação de outros mais artigos, apresentações (*slides*) e *websites* voltados para pesquisa de padrões de metadados, especificamente quanto a forma de aplicação e uso dos metadados de preservação e do modelo de referência supracitados.

4.1 OAIS

O modelo OAIS, traduzido em português pela sigla SAAI, foi publicado em 2003, “foi desenvolvido pela Consultative Committee for Space Data Systems (CCSDS) no âmbito da NASA” (SARAMAGO, 2004, p.3) e corresponde a norma ISO 14721:2003³¹. O OAIS consiste em uma iniciativa que, por intermédio de um vocabulário padrão (metadados), busca estabelecer um modelo de referência para a preservação de materiais digitais a longo prazo. No caso da norma, ela é responsável por apresentar as diretrizes que garantam a confiabilidade do que está sendo preservado, além de estabelecer o padrão de metadados, que por sua vez afiançará a comunicação entre os sistemas que o utilizam (ARELLANO, 2004; SARAMAGO, 2004; SOUZA *et al.*, 2012). Essas características apresentam o OAIS como sistema aberto, estando a serviço das mais diversas comunidades (SARAMAGO, 2004). Entretanto, deve-se salientar que o modelo OAIS, como afirma Arellano (2008), não é necessariamente uma

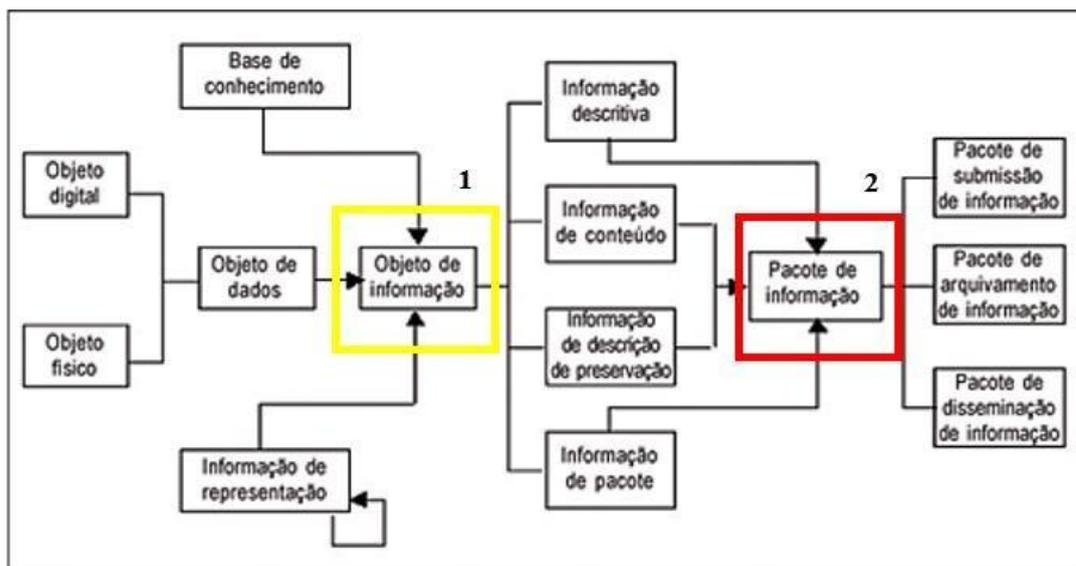
³¹ Website: <https://www.iso.org/standard/24683.html>.

implementação, e sim um modelo e/ou esquema que abrange tanto a arquitetura quanto a operacionalidade de um repositório digital, como anteriormente dito mediante definição e padronização do uso de metadados específicos. Sendo assim, partindo-se do princípio de que o OAIS é um modelo de aplicação lógico-conceitual é possível compreendê-lo como “[...] uma infraestrutura conceitual que descreve o ambiente, as interfaces externas, os componentes funcionais e os objetos de informação, associados com um sistema responsável pela preservação de longo prazo de materiais digitais” (SAYÃO, 2010, p. 13). Quanto aos objetivos, Thomaz e Soares (2004, p. 9) afirmam que o modelo OAIS almeja o seguinte:

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

No que se refere à sua aplicação e funcionamento, basicamente, o modelo possui duas infraestruturas: funcional e de informação. No primeiro modelo são estabelecidas as diretrizes que devem ser seguidas pelos repositórios pertencentes à iniciativa OAIS, entre questões como admissão ou recepção, armazenamento, gestão de dados, planejamento da preservação, administração e acesso. O modelo de informação está ligado às descrições e ao uso dos metadados de preservação, visando assegurar a preservação por longo período, garantir o acesso e apresentar os dados de forma que todas as comunidades interessadas possam interpretá-los (SAYÃO, 2010). Conforme a figura 10 é possível observar o modelo de informações no OAIS segundo o **objeto de informação** e o **pacote de informação**.

Figura 10: Modelo de informação no OAIS



Fonte: Arellano (2004), adaptado pelo autor.

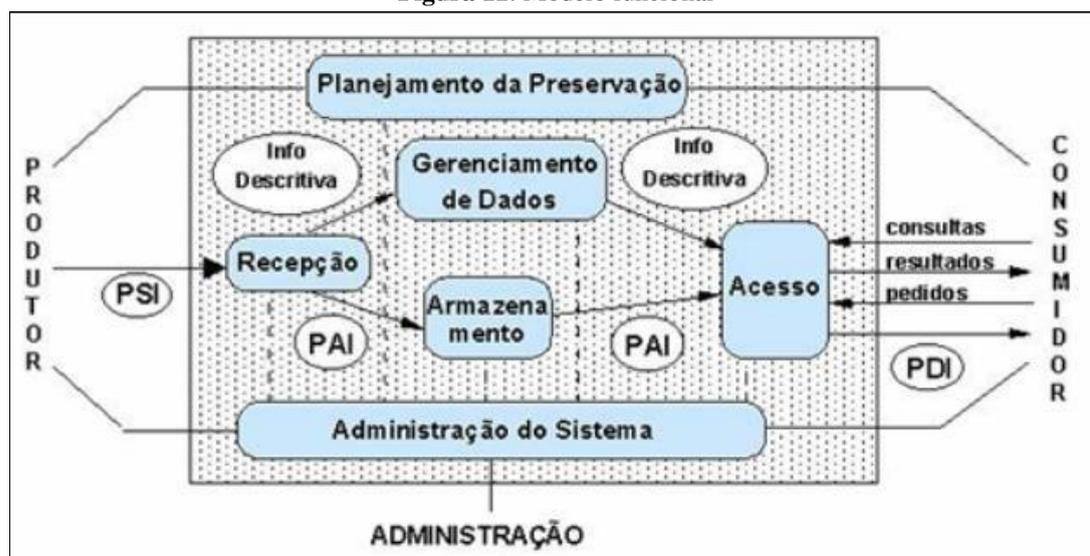
Esmiuçando a figura 10, observa-se o apontamento inicial aos **objetos físico e digital**. A partir da presença de pelo menos um deles forma-se o **objeto de dados**. Ao assumir uma base de conhecimento da comunidade alvo e receber informações pertinentes à sua representação dentro do modelo, o **objeto de dados** transforma-se em um **objeto de informação** (1) (RODRIGUES *et al.*, 2013). Com o **objeto de informação** criado, a sua caracterização se dará a partir de quatro tipos de informação. A primeira, chamada de **informação descritiva**, é responsável por facilitar “[...] o acesso à informação de pacote através de ferramentas de pesquisa e recuperação” (RODRIGUES *et al.*, 2013), ou seja, apresentando as informações básicas que caracterizam o objeto a ser representado. Em seguida tem-se a **informação de conteúdo**, que é o conteúdo principal em si, isto é, a informação alvo das ações de preservação. Na sequência apresenta-se a **informação de descrição de preservação (IDP)**, a qual abarca as informações necessárias para a preservação do conteúdo, subdividida em quatro categorias: “[...] referência (descritores), contexto (relacionamentos com o ambiente), proveniência (proveniência e histórico) e rigidez (informação para a comprovação de integridade e autenticidade)” (RODRIGUES *et al.*, 2013). O quarto tipo é a **informação de pacote**, que reúne a informação de conteúdo com a IDP a fim de torná-la um pacote identificável (descrito) e passível de recuperação de seus conteúdos (RODRIGUES *et al.*, 2013).

Com o **objeto de informação** devidamente representado, o item está preparado e pronto para ser armazenado, passando ao **pacote de informação** (2), seja para o fim de submissão ao sistema, de arquivamento e/ou de disseminação.

Pacote de submissão de informação (PSI), pacote enviado do produtor para o arquivo; pacote de arquivamento de informação (PAI), pacote de informação efetivamente armazenado dentro do arquivo; e pacote de disseminação de informação (PDI), pacote transferido do Arquivo para um consumidor em resposta a uma solicitação. (RODRIGUES, *et al.*,2013)

Mas para que o pacote de informação seja gerado e constituído, fases ocorrem internamente na iniciativa OAIS, o que se chama de modelo funcional (figura 11). Esse modelo nada mais é do que o processo de preservação em si, estruturado entre **produtor**, **administração** e **consumidor**. Primeiramente, faz-se necessário entender que tal processo se divide em duas etapas principais, sendo elas o **planejamento da preservação** e a **administração do sistema**, estas em conjunto realizam o processo do modelo OAIS.

Figura 11: Modelo funcional



Fonte: Rodrigues et al. (2013).

A primeira etapa de funcionamento ocorre quando o **produtor** faz o envio do **PSI** para o arquivo (representação imagética, quadrada e pontilhada em seu interior). Em seguida, ao receber o pacote, na entidade de **recepção** (admissão), o **PSI** sofrerá duas ações, sendo uma correspondente ao **planejamento da preservação**³² e a outra ação referente a **administração do sistema**³³. Enquanto no **planejamento da preservação** ocorrerá o gerenciamento de dados, especificamente a **descrição** (info descritiva) do arquivo, na **administração do sistema** será realizado o **armazenamento** do pacote que, neste caso, diante do seu arquivamento, passa a ser denominado de PAI. Por fim, com o pacote tratado e armazenado, ocorre a **entidade de acesso**, responsável em informar a “existência, descrição, localização e disponibilidade da informação armazenada no OAIS” (RODRIGUES *et al.*,2013), dentre outras ações, com fins de facilitar a localização e recuperação de conteúdos a partir da dinâmica ocorrida entre três ações: **consultas** (busca), **resultados** (ocorrências) e **pedidos** (desejo/recuperação esperada).

Quanto ao uso, a aplicação do modelo OAIS pode ser identificada em propostas realizadas em ambientes de ensino e/ou pesquisa como é o caso do Laboratório de Tecnologia do Conhecimento (LIBER)³⁴, ligado ao Diretório Central dos Estudantes (DCE) da Universidade Federal de Pernambuco (UFPE), é considerado um espaço multiusuário, com foco em estudos (utilização e análise) de tecnologias do conhecimento voltadas à preservação a

³² O planejamento da preservação monitora o ambiente OAIS, além de fornecer informações que permitam a preservação a longo prazo

³³ A administração do sistema gerencia toda a rotina operacional do arquivo PAI, ou seja, realiza as rotinas de armazenar, gerenciar, atualizar, criar cópias do pacote para a entidade de acesso dentre outras rotinas.

³⁴ Website: <http://www.liber.ufpe.br/>

longo prazo. O LIBER realiza as suas pesquisas com base em parcerias e convênios com a UFPE, além de Instituições (LIBER, 2018).

O laboratório LIBER está envolvido em cerca de 72 projetos e extensões de pesquisa. Entre eles está o intitulado “Memórias do Golpe”³⁵, que traz uma coleção de documentos históricos referentes ao período entre 1965 e 1985 no Brasil, entre atos institucionais e recortes de jornais (LIBER, 2018). Outro projeto chama-se “Memorial da Justiça”³⁶, que tem por objetivo preservar e difundir o patrimônio histórico documental de Pernambuco (PE) e que é composto por cerca de 500 processos judiciais criminais do período imperial brasileiro, entre os anos de 1822 e 1889 (TJPE³⁷, 2018). O LIBER também coordena o projeto chamado “Memorial Joaquim Caetano”³⁸, que objetiva dar acesso à documentos históricos ligados à Joaquim Caetano Silva, personalidade brasileira importante nas relações diplomáticas entre Brasil e Holanda referentes aos arquivos históricos que envolviam os Estados Gerais em Haia em meados do século XIX. (LIBER, 2018). O projeto “Frans Post: imagens de um Brasil Holandês”³⁹ se trata de uma coleção de imagens que retrata o Brasil no período de 1637 a 1660 e que foram digitalizadas a partir do catálogo de uma exposição realizada no Museu de arte de São Paulo (MASP), realizada por artistas holandeses, dentre eles, o próprio Frans Post (LIBER, 2018). O trabalho “Rede Memorial de Pernambuco”⁴⁰, realizado da colaboração entre LIBER, Museu da Cidade do Recife, Biblioteca Pública do Estado e Arquivo Público do Estado, objetiva, em suma, a preservação e o acesso ao patrimônio memorial de Pernambuco (REDE MEMORIAL DE PERNAMBUCO, 2018). O projeto “Selos do Brasil”⁴¹, gerenciado pelo sistema Clio⁴² e desenvolvido pela LIBER, traz em seu acervo a "Coleção de Selos do Centenário da Revolução Republicana em Pernambuco" (LIBER, 2018). O Banco de Teses e Dissertações da UFPE⁴³, ambiente onde estão reunidas as publicações acadêmicas e científicas da Universidade, preservadas e disponibilizadas à comunidade, é mais um projeto realizado (LIBER, 2018). O “Projeto Ultramar”⁴⁴ é mais uma pesquisa realizada, especificamente sobre a preservação de documentos históricos do Brasil (LIBER, 2018). Chamado de “Atlas do Brasil Holandês”⁴⁵, trata-se de um projeto sobre a cooperação brasileira-holandesa-espanhola e que

³⁵ Website: <http://www.liber.ufpe.br/ditadura/>.

³⁶ Website: <http://www.tjpe.jus.br/web/memorial-da-justica>.

³⁷ Tribunal de Justiça de Pernambuco. Website: <http://www.tjpe.jus.br/inicio>.

³⁸ Website: <http://www.liber.ufpe.br/jcaetano/>.

³⁹ Website: <http://www.liber.ufpe.br/libvirt/fpost/>.

⁴⁰ Website: <http://redememorialpernambuco.blogspot.com/>.

⁴¹ Website: <http://www.liber.ufpe.br/selos/acervo2.php>.

⁴² Sistema de Gerenciamento de Acervos Digitais, desenvolvido pelo Laboratório LIBER (CLIO, 2018).

⁴³ Website: <https://repositorio.ufpe.br/>.

⁴⁴ Website: <http://www.liber.ufpe.br/ultramar/>.

⁴⁵ Website: <http://www.liber.ufpe.br/atlas/>.

disponibiliza um acervo formado de artigos, imagens, mapas, atlas dentre outros, mostrando a relação entre o Brasil e a Holanda a partir do século 16 (LIBER, 2018). Por fim, a “Oficina Guaianases”⁴⁶ objetiva, por meio de uma rede, preservar parte da história cultural de PE por meio de uma base de dados constituída de 2036 litografias e que se encontram armazenadas na biblioteca do Centro de Artes e Comunicação (CAC) da UFPE (LIBER, 2018).

Outra iniciativa que faz uso do modelo OAIS é a Rede Brasileira de Serviços de Preservação Digital, ou Rede Cariniana⁴⁷. Criada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)⁴⁸, a rede tem por objetivo “salvaguardar os registros da ciência, tecnologia e do patrimônio cultural do Brasil” (ARELLANO, 2012, p. 85). Quanto à composição de sua rede, há a subdivisão entre parceiros integrais e demais parceiros. A Universidade de Campinas⁴⁹, a Universidade Federal de Santa Maria (UFSM)⁵⁰, a Universidade da Paraíba (UFPB)⁵¹, a Universidade de São Paulo (USP)⁵², a Universidade de Brasília (UnB)⁵³, a Universidade Federal do Goiás (UFG)⁵⁴, a Universidade Federal da Bahia (UFBA)⁵⁵, a Universidade Federal do Rio Grande do Norte (UFRN)⁵⁶, a Universidade Federal do Rio Grande do Sul (UFRGS)⁵⁷, a Universidade Federal de Santa Catarina (UFSC)⁵⁸ e o próprio IBICT compõem os parceiros integrais. Quanto às demais parcerias, entre os parceiros institucionais estão o Jardim Botânico de Brasília⁵⁹, a empresa Lepidus Tecnologia⁶⁰ e a iniciativa *Legatum*, desenvolvida pelo grupo de pesquisa Cultura, Representação e Informações Digitais (CRIDI)⁶¹ da UFBA. A iniciativa ainda possuiu, como parceiros internacionais, a rede *Lots of Copies Keeps Stuff Save* (LOCKSS)⁶² e a agência de arquivamento de periódicos eletrônicos chamada The Keepers⁶³.

⁴⁶ Website: <http://www3.ufpe.br/guaianases/>.

⁴⁷ Website: <http://cariniana.ibict.br/>

⁴⁸ Website: <http://www.ibict.br/>

⁴⁹ Website: <http://www.unicamp.br/unicamp/>

⁵⁰ Website: <http://site.ufsm.br/>

⁵¹ Website: <http://www.ufpb.br/>

⁵² Website: <http://www5.usp.br/>

⁵³ Website: <http://www.unb.br/>

⁵⁴ Website: <https://www.ufg.br/>

⁵⁵ Website: <https://www.ufba.br/>

⁵⁶ Website: <http://ufrn.br/>

⁵⁷ Website: <http://www.ufrgs.br/ufrgs/inicial>

⁵⁸ Website: <http://ufsc.br/>

⁵⁹ Website: <http://www.jardimbotanico.df.gov.br>

⁶⁰ Website: <https://lepidus.com.br/>

⁶¹ Website: <http://www.cridi.ici.ufba.br>

⁶² Website: <https://www.lockss.org/>

⁶³ Website: <https://thekeepers.org/agencies>

A Rede Cariniana faz uso da ferramenta/software LOCKSS, utilizada na rede LOCKS supracitada, é desenvolvida pela *Stanford University*⁶⁴ e que se caracteriza pela criação de redes privadas locais (PLN) (ARELLANO, 2012). Oferecendo a estrutura necessária para uma rede de preservação digital, o LOCKSS é um *software* de código aberto, de arquitetura distribuída e que proporciona a integração com outros programas. O LOCKSS é uma ferramenta de baixo custo e que tem por modelo de referência o OAIS. (CARINIANA, 2018). O LOCKSS ao criar as redes, tem por intuito gerar cópias e distribuí-las por esta rede com o objetivo de preservar os materiais distribuídos, sendo este o seu principal método de preservação (Stanford University, 2018).

Ao fazer parte da Aliança LOCKSS, a rede Cariniana, fornece ao seu conglomerado, composto principalmente por Instituições de ensino e de pesquisa, além do compartilhamento de produções, o “apoio técnico, a capacitação e o assessoramento na negociação das permissões de arquivamento das publicações e construção de futuras subredes locais de preservação digital” (ARELLANO, 2012, p. 88). Será justamente no interior desta rede, e de suas subredes, que serão realizadas as cópias e as distribuições das publicações com vistas à preservação digital à longo prazo.

4.2 PREMIS

O PREMIS é um grupo de trabalho internacional criado entre 2003 e 2005, resultado de em uma parceria entre a *Online Computer Library Center (OCLC)*⁶⁵ e a *Research Libraries Group (RLG)*⁶⁶ e atualmente mantido pelo *Network Development and MARC*⁶⁷ *Standards Office* da LC (FORMENTON, 2015). A ideia desta parceria foi o estabelecimento de padrões de metadados para preservação digital (CAMPOS; SARAMAGO, 2007). Em outras palavras, o modelo PREMIS, ou Dicionário de Dados, foi o objeto exposto no relatório final da mencionada parceria, assim apresentado: “[...] define um conjunto de metadados básicos, implementáveis e de ampla aplicação, para apoiar a preservação digital em sistemas de repositórios” (FORMENTON, 2015, p. 81).

⁶⁴ Website: <https://www.stanford.edu/>

⁶⁵ Website: <https://www.oclc.org/en/home.html?redirect=true>

⁶⁶ Website: <http://www.rlg.org/>

⁶⁷ Machine Readable Cataloging. Website: <https://www.loc.gov/marc/ndmsso.html>

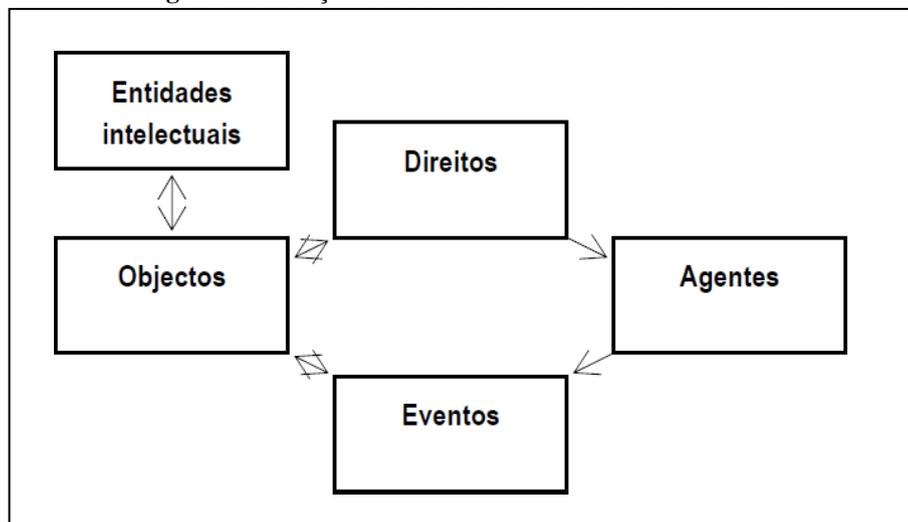
A aplicação do modelo PREMIS é incorporado a uma série de ferramentas e sistemas de preservação digital comerciais, e de código aberto. Tem a sua composição por conta do **Dicionário de Dados**, ou seja, os metadados de preservação. Atualmente na versão 3.0, o PREMIS é representado por um **esquema XML** e possui uma **documentação de suporte**, que auxilia na utilização e entendimento do mesmo, ambos revisados e atualizados pelo comitê de edição do modelo (PREMIS, 2018). Quanto à sua lógica, o PREMIS entende que um recurso digital contém em si várias representações, como textos, gráficos, fotos, áudio dentre outras, semelhante ao que se encontra em uma página *websites*, por exemplo. No que se refere a constituição do recurso a ser preservado, o modelo representa-o conforme cinco tipos de entidades, a saber: objetos, entidades intelectuais, eventos, direitos e agentes (CAMPOS; SARAMAGO, 2007). A entidade **objeto** refere-se à uma unidade de informação em formato digital, subdividida em:

- a) Ficheiros: “[...] sequência de bytes reconhecido por um sistema operativo. Um ficheiro pode ser composto por zero ou mais bytes e tem num formato, tem permissões de acesso e também incorpora alguns dados tais como o tamanho [da] data da última modificação” (CAMPOS; SARAMAGO, 2007, p.4);
- b) Fluxo de Bites (*Bits*): conjunto de dados de um ficheiro, que manifestam informações em comum com fins de preservação, sendo que este conjunto, não possui autonomia se não estiver inserido em uma estrutura (CAMPOS; SARAMAGO, 2007, p.4);
- c) Representação: “[...] conjunto de ficheiros, incluindo os metadados estruturais, necessários para restabelecer a natureza e o aspecto de uma entidade intelectual” (CAMPOS; SARAMAGO, 2007, p.4);

A **entidade intelectual** é uma unidade descrita a partir de um conjunto coerente, ou seja, é uma unidade de informação que pode ser um livro, uma revista, uma imagem e até uma base de dados, por exemplo. Os **eventos** são entidades que, durante o ciclo de vida do recurso, agregam metadados sobre quaisquer ações que o recurso venha a sofrer, inclusive de antes do recurso ser incluso em um repositório (CAMPOS; SARAMAGO, 2007). Os **agentes** “são pessoas, organizações ou aplicações informáticas associadas às acções de preservação de um dado objecto digital” (CAMPOS; SARAMAGO, 2007, p. 4). Por fim, os **direitos** “são as permissões legais de cada objecto ou agente” (CAMPOS; SARAMAGO, 2007, p. 4). Em

complemento, a figura 12 apresenta a forma como as entidades se relacionam na representação de um recurso no modelo PREMIS.

Figura 12: Relações entre as entidades no modelo PREMIS



Fonte: Campos; Saramago (2007).

Relacionado ao uso da aplicação PREMIS, entre as Instituições que a utilizam está a *Carolina Digital Repository* (CDR)⁶⁸, arquivo digital que reúne a produção acadêmica científica dos membros da Universidade da Carolina do Norte⁶⁹ (CDR, 2018a). O CDR utiliza-se dos metadados do PREMIS “[...] como uma medida para garantir a autenticidade dos objetos digitais ingeridos no repositório” (CDR, 2018b, (tradução nossa). São focos do CDR, tanto o registro dos **eventos**, quanto as ações que incidem sobre o **objeto**, gerando-se relatórios do percurso desse objeto dentro do repositório. Alguns dos eventos registrados pelo CDR são:

- Criação - quando o pacote de informações de envio [*submission information package*] (SIP) é criado na mesa de trabalho do Curador;
- Validação - quando o SIP e os objetos que ele contém são verificados em busca de vírus (ou se essa etapa for ignorada);
- Validação - quando o XML do METS que compõe o SIP é validado em relação ao perfil do METS;
- Validação da soma de verificação - quando a soma de verificação XML do METS é verificada;
- Normalização - quando, durante a transformação do SIP para o FOXML⁷⁰, os objetos recebem um identificador persistentemente exclusivo (Fedora PID);

⁶⁸ Website: <https://cdr.lib.unc.edu/>

⁶⁹ Website: <https://www.unc.edu/>

⁷⁰ O *Fedora Object XML* (FOXML) “é um formato XML simples que expressa diretamente o Modelo de Objeto Digital do Fedora” (PRATER, 2014, tradução nossa). O Fedora define um modelo de objeto digital genérico que pode ser usado para manter e fornecer as características essenciais para muitos tipos de conteúdo digital, incluindo documentos, imagens, livros eletrônicos, objetos de aprendizado multimídia, conjuntos de dados, metadados e muitos outros. (DAVIS; WILPER, 2011, tradução do autor)

- Validação - quando o registro descritivo de MODS é validado para vocabulários controlados;
- Normalização - quando os dados do MODS são transformados em Dublin Core (CAROLINA DIGITAL REPOSITORY, 2018b, tradução nossa).

Atualmente, o repositório criou dois tipos de metadados a partir do PREMIS: os **registros de eventos** e os **relatórios de objetos** (CDR, 2018c). Nos **registros de eventos** tem-se informações sobre a criação, validação, normalização dentre outras ações, e o conjunto dessas informações é tido como **registro de eventos**. Para tanto, esses eventos estão relacionados ao objeto digital, especificamente compostos pelos elementos apresentados no quadro 3:

Quadro 3: Elementos de um registro de eventos no PREMIS

Elemento	Tipo de registro	
<i>eventType</i>	Criação, validação ou normalização, conforme listado acima	
<i>eventDateTime</i>	Uma data / hora (ISO 8601)	
<i>eventIdentifier</i>	Um ID para o evento em si, que é um URN ⁷¹ baseado em um UUID ⁷² gerado (ID exclusivo gerado)	
<i>linkingAgentIdentifier</i>	O PID ⁷³ de quaisquer objetos do Fedora que foram agentes no evento (embora às vezes seja apenas um nome, não um PID, para eventos de pré-ingestão)	
	O software pode ser um agente, por exemplo, um scanner de vírus. Alguns agentes de software registrados no repositório são referenciados pelo PID, outros pelo nome. Os agentes de software podem ter várias funções e, às vezes, são o "Iniciador" de um evento	Quando uma pessoa está envolvida em um evento, seu nome (se não estiver registrado no repositório) ou PID (se estiverem registrados no repositório) é registrado. Geralmente uma pessoa é considerada o "Iniciador" de um evento se interagiu com o repositório para causar o evento
<i>linkingObjectIdentifier</i>	O PID e as funções de quaisquer objetos do Fedora e quaisquer fluxos de dados específicos que eram objetos no evento, quando aplicável	
	A função do objeto "Assunto" é registrada, embora seja uma informação redundante	Em eventos de derivação, as funções dos fluxos de dados são registradas como "Dados de origem" e "Dados derivados", com os fluxos de dados identificados pelo PID

Fonte: Adaptado de CDR (2018c).

⁷¹ *Uniform Resource Names* (URNs) são identificadores de recursos, permanentes e independentes de localização, cujo objetivo é facilitar o mapeamento de outros *namespaces*, que compartilham de suas propriedades. (NETWORK WORKING GROUP R. MOATS, 1997).

⁷² *Universally Unique Identifier* (UUID), também conhecidos por *Globally Unique Identifiers* (GUIDs), são identificadores de 128 bits, normalmente utilizados em um URN como parte de sua identificação. (ITU, 2018)

⁷³ Um PID é um identificador único, assim como um ID, e pertence ao objeto digital do modelo Fedora (DAVIS; WILPER, 2011).

Quanto aos **relatórios de objetos**, fornecem a maior parte das informações do objeto PREMIS reconhecidas pelo repositório, gerados conforme demanda. São compostos pela identidade do objeto e de todas as partes que constituem o(s) objeto(s) (CDR, 2018c) a partir de elementos como identificação, nome original, formato, localização entre outros (quadro 4).

Quadro 4: Elementos de um relatório de dados no PREMIS

Elemento	Especificação e composição	
<i>objectIdentifier</i>	Um Fedora PID ou um <i>Fedora Datastream</i> PID	
<i>originalName</i>	Se disponível; nome “amigável” do objeto	
<i>objectCharacteristics</i>	<i>compositionLevel</i> : composição do fluxo de bytes <i>size</i> : o tamanho do arquivo em bytes <i>format</i> : formato do fluxo de bytes; sempre em XML	
<i>storage (repeated)</i>	<i>contentLocation</i> : locais de armazenamento	<i>storageMedium</i> : descrição do meio de armazenamento

Fonte: Adaptado de CDR (2018c).

Quanto a geração dos **relatórios de objetos**, o PREMIS adotado o formato XML, exemplificado pela figura 13 segundo um registro descrito pelos elementos *objectIdentifier*, *originalName* e *storage*, apresentados no quadro 4, além do elemento *format*.

Figura 13: Exemplo de um relatório de objetos no PREMIS⁷⁴

```
<?xml version="1.0" encoding="UTF-8"?>
<premis xmlns="info:lc/xmlns/premis-v2">
  <object xmlns:xsi="http://www.w3c.org/2001/XMLSchema-instance" xsi:type="representation">
    <objectIdentifier>
      <objectIdentifierType>Fedora PID</objectIdentifierType>
      <objectIdentifierValue>uuid:e0521d4f-c8eb-46c1-831a-617a4e35e9b0</objectIdentifierValue>
    </objectIdentifier>
    <originalName>Final Facilities Planning Inventory (drawers 1-5d).xls</originalName>
    [...]
    <format>
      <formatDesignation>
        <formatName>text/xml</formatName>
      </formatDesignation>
      <formatNote>IANA MIME-type</formatNote>
    </format>
    [...]
    <storage>
      <contentLocation xmlns="">
        <contentLocationType xmlns="info:lc/xmlns/premis-v2">iRODS URI</contentLocationType>
        <contentLocationValue xmlns="info:lc/xmlns/premis-v2">
          irods://aalcdl24.lib.unc.edu:3333/cdrZone/home/fedora/datastreams
          /2009/1207/16/21/uuid_e0521d4f-c8eb-46c1-831a-617a4e35e9b0+MD_EVENTS+MD_EVENTS.0
        </contentLocationValue>
        </contentLocation>
        <storageMedium xmlns="">cache - unix file system</storageMedium>
      </storage>
    </object>
  </premis>
```

Fonte: Carolina Digital Repository (2018)

⁷⁴ A versão completa do código suprimido está no *website* <https://blogs.lib.unc.edu/cdr/index.php/about/cdr-development-and-collab/technical-documentation/metadata/premis/premis-object-report-sample/>.

Deve-se salientar que além dos elementos descritivos do PREMIS, também são utilizados metadados do padrão FEDORA e do *software* iRODS⁷⁵ (CDR, 2018c). Conforme a figura 13 é apresentado um exemplo de **relatório de objetivos**, trazendo as seguintes informações: “[...] exemplo foi gerado em 1/7/2010 em uma instalação de teste [e] [...] mostra dois locais de armazenamento para cada arquivo, o que reflete a replicação de arquivos no iRODS” (CDR, 2018c).

4.3 MODS

Seguindo a ideia de iniciativas que auxiliam a preservação de objetos digitais, outra aplicação identificada na literatura chama-se MODS, traduzida para a língua portuguesa como esquema de metadados de descrição de objetos. O MODS foi desenvolvido pela LC em 2002, expresso em linguagem XML, derivado do padrão MARC, este objetiva traspor elementos selecionados do MARC para a descrição de recursos. Atualmente, na versão 3.7, o esquema MODS resume-se a um conjunto de elementos para descrição bibliográfica, usualmente voltado para bibliotecas, com mais simplicidade que um padrão MARC por exemplo (MODS, 2016).

O padrão MODS apresenta algumas características que lhe dão destaque quanto a sua aplicação. Os elementos do esquema possuem uma alta compatibilidade com registros em formato MARC, devido as suas semelhanças semânticas. Na verdade, o esquema MODS é uma versão mais simples do que o padrão MARC já que possui a capacidade de englobar vários elementos MARC em apenas um elemento MODS, outra vantagem é que o esquema dispõe de uma descrição mais completa do que a fornecida pelo padrão DC, por exemplo. Além disso, a sua estrutura em XML lhe permite integrar-se a outros padrões de metadados na mesma linguagem trazendo, portanto, toda a interoperabilidade da estrutura XML. (FORMENTON, 2015; COPAC, 2010).

No que se refere à sua estrutura, o esquema MODS é composto por 20 elementos de descrição principais, podendo alguns serem qualificados. Desta forma, o quadro 5 apresenta os elementos do padrão MODS e suas possíveis qualificações.

⁷⁵ *Software* de gerenciamento de dados de código aberto. Website: <https://irods.org/>.

Quadro 5: Elementos do esquema MODS

<i>Element</i>	Elemento (português)	Qualificadores
<i>titleInfo</i>	Informação de Título	<i>title; subTitle; partNumber; partName; nonSort</i>
<i>name</i>	Nome	-
<i>typeOfResource</i>	Tipo de recurso	-
<i>genre</i>	Gênero	-
<i>originInfo</i>	Informação de Origem	<i>dateIssued; dateCreated; dateCaptured; dateValid; dateModified copyrightDate; dateOther; edition; issuance; frequency; calendar attribute</i>
<i>language</i>	Idioma	<i>languageTerm; scriptTerm;</i>
<i>physicalDescription</i>	Descrição física	<i>form; reformattingQuality; internetMediaType; extent; digitalOrigin; note</i>
<i>abstract</i>	Resumo	-
<i>tableOfContents</i>	Índice	-
<i>targetAudience</i>	Público alvo	-
<i>note</i>	nota	-
<i>subject</i>	Assunto	<i>topic geographic temporal titleInfo name genre hierarchicalGeographic cartographics geographicCode occupation</i>
<i>classification</i>	Classificação	-
<i>relatedItem</i>	Item relacionado	
<i>identifier</i>	Identificador	
<i>location</i>	Localização	<i>physicalLocation; shelfLocator url; holdingSimple; holdingExternal</i>
<i>accessCondition</i>	Condição de acesso	-
<i>part</i>	Parte	<i>detail extent; date text</i>
<i>extension</i>	Extensão	-
<i>recordInfo</i>	Informação de registro	<i>recordContentSource; recordCreationDate; recordChangeDate; recordIdentifier; recordOrigin; languageOfCataloging descriptionStandard</i>

Fonte: Adaptado de Formeton (2015) e UFRGS (2018).

Em relação a sua aplicação enquanto um método de preservação digital, o esquema MODS apoia outros padrões expressos em XML que visam a preservação, como o METS e o PREMIS, por intermédio de suas informações bibliográficas. Três elementos MODS incidem positivamente acerca da preservação digital. São eles: **informação de origem** (*originInfo*), **item relacionado** (*relatedItem*) e **condição de acesso** (*accessCondition*) (FORMETON, 2015).

Estes elementos registram informações que são úteis para auxiliar os metadados de preservação na corroboração da proveniência, autenticidade e integridade dos objetos digitais, bem como no reconhecimento dos direitos associados ao recurso que podem intervir, sobretudo, na realização das ações de preservação, acesso e utilização dos seus respectivos conteúdos. (FORMENTON, 2015, p. 70)

Entre os exemplos de implementação e uso do esquema MODS está o *Columbia University Digital Library Repository*⁷⁶, ou *Digital Library Collections (DLC)*. O repositório digital das bibliotecas da *Columbia University*⁷⁷ traz em sua coleção imagens digitalizadas, fotografias, pôsteres, manuscritos dentre outros materiais (MODS, 2010; DIGITAL LIBRARY COLLECTIONS, 2018) e utilizam o padrão MODS para descrever e preservar os objetos armazenados em seu repositório.

Outro exemplo de uso da aplicação do esquema MODS chama-se “Copac: UK and *Irish Research Library Catalogue*”⁷⁸. Trata-se de um catálogo de busca que está interligado às bibliotecas do Reino Unido e Irlanda, e que por meio dos catálogos das mais de 100 grandes bibliotecas participantes, traz um rico e único material para livre pesquisa (COPAC, 2018a). Sendo uma iniciativa que disponibiliza materiais de forma gratuita (acesso livre), ao passo que evita questões sobre direitos autorais caso adotassem o formato MARC, o Copac utiliza o esquema XML MODS descrever bibliograficamente os seus registros. A proposta apresenta registros bibliográficos completos, detalhados além de notas, com informações sobre a biblioteca onde o registro está, e de termos que representam os assuntos dos registros (RUDDOCK, 2010; COPAC, 2018b).

Como exemplo de um registro em formato XML MODS no catálogo do Copac, a figura 14 apresenta um registro com destaques aos campos que remetem a detalhes, como os elementos *place*, que apresenta dados sobre o local, editora e ano de publicação, *titleInfo*, que remete a informações sobre o título do registro, *note*, contendo informações que ampliam a descrição e que não estão contemplados em outros campos, e *h:localMolds*, especificamente sobre o tipo de endereço de acesso (CURL⁷⁹), local de hospedagem e a identificação única no catálogo para acesso via CURL.

⁷⁶ Website: <https://academiccommons.columbia.edu/>

⁷⁷ Website: <https://www.columbia.edu/>

⁷⁸ Website: <https://copac.jisc.ac.uk/>

⁷⁹ Copac Uniform Resource Locator.

Figura 14: Exemplo de um registro em XML MODS no Copac

```

- <modsCollection>
- <mods xsi:schemaLocation="http://www.loc.gov/mods/v3
http://www.loc.gov/standards/mods/v3/mods-3-2.xsd http://copac.ac.uk/schemas/mods-copac/v1
http://copac.ac.uk/schemas/mods-copac/v1/mods-copac-v1.xsd http://copac.ac.uk/schemas/holdings/v1
http://copac.ac.uk/schemas/holdings/v1/holdings-v1.xsd">
- <recordInfo>
  <recordIdentifier source="copac">72012701502</RecordIdentifier>
  <recordContentSource>Copac</recordContentSource>
  <recordCreationDate>20061021</recordCreationDate>
</recordInfo>
  <typeOfResource>text</typeOfResource>
- <originInfo>
  <dateIssued encoding="marc">1993</dateIssued>
</originInfo>
- <originInfo>
- <place>
  <placeTerm type="text">London :</placeTerm>
</place>
  <publisher>David Campbell,</publisher>
  <dateIssued>1993.</dateIssued>
</originInfo>
- <titleInfo>
  <title>The decline and fall of the Roman Empire</title>
</titleInfo>
[... ]
- <note>This ed. originally published: 1910. Repr. with new introduction.</note>
<note>Three vols. in slip case.</note>
<note>Bibliography: Vol.1, p.xcix. - Includes index.</note>
<note>Dates of available copies: 1993, 1994.</note>
- <physicalDescription>
  <extent>3 v. : general.tables,maps ; 21 cm.</extent>
</physicalDescription>
  <identifier type="isbn">1857150953</identifier>
  <classification authority="ddc" edition="20">937.06</classification>
[... ]
- <h:localHolds>
  <h:org type="CURL" displayName="British Library">BLI</h:org>
  <h:objId>72012701502</h:objId>
[... ]
</mods>
</modsCollection>

```

Fonte: Adaptado de Copac, 2018b.

Sobre o elemento *note* (nota), apresentado no registro da figura 14, traz informações sobre a publicação original da obra, datas de avaliação das cópias entre outros dados. O mencionado campo é um exemplo do quanto um registro em MODS pode ser detalhado, trazendo maior riqueza descritiva ao registro, tornando o MODS um esquema de metadados favorável quando se pensa na preservação a longo prazo de obras bibliográficas, por exemplo.

4.4 METS

Outro método identificado na literatura sobre a preservação de objetos e recursos digitais denomina-se METS, criado pela LC em 2001. O esquema METS, mediante uso da linguagem XML, objetiva gerir objetos digitais em bibliotecas e repositórios a partir de metadados, descritivos, administrativos e estruturais. No uso desses metadados, o METS contempla as diversas e complexas ligações entre os objetos digitais, oferecendo assim uma codificação mais flexível (ARELLANO, 2010). Para tanto, ao contemplar as diferentes partes

de um objeto digital, a aplicação do esquema METS fica por conta de uma estrutura composta por sete seções principais. Essas seções são:

1. **Cabeçalho METS:** onde se encontram os metadados que descrevem o documento em si como, por exemplo, autor, criador, editor, datas entre outros;
2. **Metadados descritivos:** podem ser tanto metadados descritivos externos ao documento METS, quanto metadados do próprio documento, inseridos separadamente ou juntos;
3. **Metadados administrativos:** informações referentes a questões administrativas do documento como a forma que foram criados, quando foram armazenados, direitos de propriedade intelectual, informações sobre possível objeto que deu origem, além de informações sobre os arquivos que compõe o objeto, externas ou internas ao documento;
4. **Secção de arquivos:** onde são listados os arquivos que possuem uma versão do objeto digital descrito;
5. **Mapa estrutural:** é a parte central do documento METS, apresentando uma estrutura hierárquica para o objeto digital, ligando os elementos dessa estrutura aos arquivos e metadados que correspondam à cada elemento;
6. **Ligações estruturais:** onde os criadores METS registram se há hiperligações entre os elementos da hierarquia do mapa estrutural do documento, como também arquivam *websites*;
7. **Comportamento:** utiliza-se para fazer associações entre um código executável (um programa ou um serviço) com um conteúdo do objeto (documento) METS, permitindo que esse conteúdo tenha uma característica específica, ou seja, um comportamento, além de estar disponível em fontes exteriores ao código METS (GRACIO, 2009; RODRIGUES, 2008).

Com as sete seções, o METS garante, além da descrição do recurso para armazenamento, a interoperabilidade entre repositórios. (GRACIO, 2009). O METS oferece ainda a associação do objeto digital com comportamentos e serviços (SAYÃO, 2007). A atualização do esquema fica por conta do seu conselho editorial, tendo a última atualização ocorrida em sete de janeiro de 2015, gerando a versão 1.11 (METS, 2017).

Com relação ao uso, o esquema METS é adotado pela Biblioteca Nacional Digital de Portugal (BND)⁸⁰, instituição que possui um acervo *online* e gratuito com cerca de 25.000 documentos (BIBLIOTECA DIGITAL NACIONAL, 2018). A preservação deste acervo ocorre a partir de um plano de estratégias (aplicações em conjunto), entre eles o OAIS e o PREMIS. Em outras palavras, a Biblioteca Nacional Digital portuguesa definiu o seu modelo de aplicação ao considerar modelos de dados, perfis de metadados, opções de implementação etc. (PATRÍCIO, 2015). Sua composição se deu basicamente em três estruturas, na definição do modelo de informação foi implementado o OAIS, enquanto que as definições de tipologia e estrutura dos objetos digitais da BND foram embasadas nos esquemas PREMIS e METS. (PATRÍCIO, 2015). Especificamente sobre o METS, a Instituição utiliza-o na descrição dos arquivos tanto de uma forma geral, em sua totalidade (ficheiro *metsItems.xml*), quanto de forma específica (ficheiros *mets.xml*) (PATRÍCIO, 2015). Além disso, devido ao METS, os objetos digitais da BND são estruturados de forma positiva a migração de obras físicas para digitais (BORBINHA et al 2005).

Outro exemplo de Instituição que faz uso do esquema METS em seus registros é a *Brown University Library*⁸¹, especificamente em seu repositório digital, funcionando como um centro físico, virtual e intelectual de apoio aos discentes e docentes da Universidade. (BROWN UNIVERSITY LIBRARY, 2018). Em sua coleção especial de manuscritos e documentos históricos, por exemplo, o *Brown Archival & Manuscript Collections Online*⁸², a biblioteca disponibiliza de forma *online* a busca por diversos documentos históricos, além da cópia do registro do material no formato METS, conforme o exemplo de um manuscrito da coleção supracitada, extraído diretamente do *website* da biblioteca e apresentado na figura 15.

⁸⁰ Website: <http://purl.pt/index/geral/PT/about.html>

⁸¹ Website: <https://library.brown.edu/>

⁸² Website: https://repository.library.brown.edu/studio/collections/id_589/

Figura 15: Registro em METS na *Brown University Library*⁸³

```

▼<METS:mets xmlns:METS="http://www.loc.gov/METS/" xmlns:ir="http://dl.lib.brown.edu/md/irdata" xmlns:mods="http://www.loc.gov/mods/v3"
xmlns:rights="http://cosimo.stanford.edu/sdr/metsrights/" xmlns:sdr="http://rs.tdwg.org/dwc/xsd/simpledarwincore/" xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd
http://www.loc.gov/standards/mods/v3/ http://www.loc.gov/standards/mods/v3/mods-3-3.xsd" OBJID="bdr:211414"
▼<METS:dmdSec ID="DM1">
▼<METS:mdWrap MDTYPE="MODS">
▼<METS:xmlData>
▼<mods:mods ID="manu004231" xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-3.xsd">
▼<mods:titleInfo>
<mods:title>Wing, Nell to Sue Windows</mods:title>
<mods:subTitle>May 24, 1976</mods:subTitle>
</mods:titleInfo>
▼<mods:name type="personal">
<mods:namePart>Wing, Nell</mods:namePart>
▼<mods:role>
<mods:roleTerm type="text">creator</mods:roleTerm>
</mods:role>
</mods:name>
▼<mods:name type="personal">
<mods:namePart>Windows, Sue Smith</mods:namePart>
▼<mods:role>
<mods:roleTerm type="text">recipient</mods:roleTerm>
</mods:role>
</mods:name>
▼<mods:name type="corporate">
<mods:namePart>General Service Office of A.A.</mods:namePart>
▼<mods:role>
<mods:roleTerm type="text">contributor</mods:roleTerm>
</mods:role>
</mods:name>
<mods:typeOfResource>text</mods:typeOfResource>
<mods:genre authority="aat">letters (correspondence)</mods:genre>
▼<mods:originInfo>
<mods:dateCreated encoding="w3cdtf" keyDate="yes">1976-05-24</mods:dateCreated>
</mods:originInfo>
▼<mods:language>
<mods:languageTerm authority="iso639-2b" type="code">eng</mods:languageTerm>
</mods:language>
▼<mods:physicalDescription>

```

Fonte: *Brown University Library* (2018).

Nota-se, a partir da figura 15, devido a característica de integração entre os esquemas de metadados pelo METS, a presença de outro importante método citado no presente trabalho, ou seja, o modelo MODS, responsável pela descrição da exemplificada obra e presente no catálogo da *Brown University Library*. Desta forma, infere-se que, em conjunto, ambas as estratégias realizam as ações necessárias para a descrição e preservação digital a longo prazo dos materiais catalogados na coleção disponível no repositório digital da Instituição.

⁸³ A versão completa do código suprimido está no *website* <https://repository.library.brown.edu/studio/item/bdr:211414/>.

4.5 ELEMENTOS DE METADADOS PARA PRESERVAÇÃO DIGITAL

Os padrões de metadados identificados na pesquisa variam em finalidades e alguns objetivam questões além da preservação digital, como o padrão MODS, que tem por objetivo principal a descrição bibliográfica de um recurso, e o esquema METS, normalmente utilizado em repositórios digitais, codificando dados descritivos, administrativos e estruturais de forma a torná-los comunicáveis. Por outro lado, o modelo PREMIS, entre os identificados na literatura, como o único que originalmente objetiva às estratégias de preservação digital. Porém, tendo o seu cunho destinado ou não à preservação digital, esses padrões de metadados apresentam elementos que acabam por auxiliar na preservação digital, permitindo-os serem incluídos entre os métodos utilizados no processo de preservação dos dados dos objetos digitais. Para tanto, apresenta-se a relação dos campos/elementos de cada padrão que auxiliam na estruturação dos dados do recurso e que, conseqüentemente, permitem a preservação digital desses dados (quadro 6).

Quadro 6: Elementos de padrões de metadados para preservação digital

Padrão	Características	Elementos de metadados úteis para a preservação digital		Características
		<i>Element</i>	Elemento	
PREMIS	<ul style="list-style-type: none"> - Esquema XML de metadados para suporte a preservação digital em repositórios. - É mantido pela LC, focaliza o sistema e sua gerência e define unidades semânticas. - Pode ter apoio do DC, MODS, EAD e MIX na verificação da autenticidade, integridade, procedência ou direitos relativos aos objetos mantidos em um repositório digital. 	<i>Intellectual Entities</i>	Entidade Intelectual	Unidade descrita a partir de um conjunto coerente de dados, ou seja, é uma unidade de informação que pode ser um livro, uma revista, uma imagem e até uma base de dados, por exemplo.
		<i>Objects</i>	Objeto	Unidade de informação em formato digital, subdividida em ficheiros, fluxo de <i>bytes</i> (bits) e representação.
		<i>Events</i>	Eventos	Entidades que, durante o ciclo de vida do recurso, agregam metadados sobre quaisquer ações que o recurso venha a sofrer.
		<i>Agents</i>	Agentes	São pessoas, organizações ou aplicações informáticas associadas às ações

				de preservação de um dado objeto digital.
		<i>Rights</i>	Direitos	Permissões legais de cada objeto ou agente.
MODS	<p>- Esquema XML aplicável aos objetos de bibliotecas digitais e mantido pela LC;</p> <p>- Elementos descritivos herdam a semântica dos elementos do MARC 21;</p> <p>- É mais simples que o MARC completo e mais rico que o DC qualificado.</p>	<i>originInfo</i>	Informação de Origem	Trata da origem do recurso como dados do local de origem, edição, datas relacionadas, direitos autorais entre outras informações.
		<i>physicalDescription</i>	Descrição Física	Apresenta características físicas do recurso, dados de origem, dimensões, duração, formato entre outros.
		<i>tableOfContents</i>	Índice	Descreve os conteúdos do recurso, por exemplo, a indicação do conjunto de capítulos incluídos num livro ou em um documento textual. ⁸⁴
		<i>relatedItem</i>	Item Relacionado	Traz as associações do recurso, ou seja, itens atrelados a ele, como referências, outras versões do recurso, outros formatos do mesmo entre outras informações.
		<i>identifier</i>	Identificador	Código ou número padrão único que identificará o recurso. Por exemplo, o ISBN.
		<i>location</i>	Localização	Informações sobre o mantenedor (Instituição) do recurso, além da localização eletrônica (URL) do mesmo.

⁸⁴ (FORMENTON, 2015, p. 69);

		<i>part</i>	Parte	Partes estruturais do recurso, a numeração de parte do recurso etc. Por exemplo: volume, edição, capítulo, seção, parágrafo entre outros, e suas respectivas medições como a página, minuto etc.
		<i>accessCondition</i>	Condição de Acesso	Apresenta as restrições ou a falta destas para o acesso, o uso e a reprodução do recurso. Por exemplo: direitos de propriedade intelectual, limitações de distribuição entre outros.
		<i>recordInfo</i>	Informação de Registro	Informações necessárias ao gerenciamento e a interpretação dos metadados do registro como a data da última modificação, a origem ou proveniência do registro, o idioma, as regras para descrição (AACR2 ⁸⁵ , RDA ⁸⁶ etc.) e qual instituição o criou ou alterou a sua originalidade.
METS	<p>- Esquema XML mantido pela LC para gestão de objetos de biblioteca digital;</p> <p>- Pode implantar e padronizar os pacotes PSI, PAI ou PDI no modelo SAAI;</p> <p>- Organiza e expressa relações hierárquicas e hiperligações entre os objetos e suas partes;</p> <p>- Inclui DC, MODS, EAD e MIX na seção de</p>	<i>mets Header</i> < <i>metsHdr</i> >	Cabeçalho METS	Seção onde estão os metadados que descrevem o documento em si como, por exemplo, autor, criador, editor, datas entre outros.
		<i>Structural Map section</i> < <i>structMap</i> >	Mapa Estrutural	Parte central do documento METS, apresentando uma estrutura hierárquica para o objeto digital e ligando os elementos dessa estrutura aos arquivos e metadados que

⁸⁵ *Anglo American Cataloguing Rules, second edition*

⁸⁶ *Resource Description and Access.*

	Metadados Descritivos, e pode ter o PREMIS na seção de Metadados Administrativos.			correspondam à cada elemento.
		<i>Structural Link section</i> <structLink>	Ligações Estruturais	Seção onde os criadores METS registram se há hiperligações entre os elementos da hierarquia do mapa estrutural do documento, como também arquivam <i>websites</i> .
		<i>Behavior Section</i> <behaviorSec>	Comportamento	Utiliza-se para fazer associações entre um código executável (um programa ou um serviço) com um conteúdo do objeto (documento) METS, permitindo que esse conteúdo tenha uma característica específica, ou seja, um comportamento, além de estar disponível em fontes exteriores ao código METS.

Fonte: Adaptado de Formenton (2015); Inclusões do autor (2018).

Os padrões de metadados (quadro 6) apresentam alguns elementos que são semelhantes quanto às atuações na descrição, na coleta e na preservação dos dados, como é o caso dos elementos *recordInfo*, do modelo MODS, e do *mets Header*, do esquema METS, avizinados quanto ao gerenciamento e a interpretação dos metadados do registro, neste caso, referentes a proveniência desse registro, a data de entrada dos dados no sistema, as possíveis alterações no registro, a data de modificações no registro, os responsáveis pelas edições no registro etc. Os padrões de metadados elencados também se aproximam em outro ponto importante na preservação digital, ou seja, nas possíveis relações entre os registros, entre os recursos, como, por exemplo, relacionar documentos de um mesmo fim e que serão preservados em conjunto. Entre os metadados estão o *relatedItem*, do padrão MODS, e o *structLink*, do esquema METS. Em relação ao modelo PREMIS, o mesmo tem por uma de suas características a separação dos

elementos de um documento a fim de preservá-lo ao invés de usar um metadado para várias etapas assim como os demais modelos.

Outros elementos, de cada padrão, buscam diferentes objetivos, porém são úteis a preservação digital do recurso em suas distintas maneiras, como por exemplo o elemento de condição de acesso (*accessCondition*) do padrão MODS que é importante quando se pensa na recuperação do recurso e o elemento de comportamento (*Behavior Section*) do modelo METS que traz informações que os demais padrões não contemplam como a possível associação à um programa ou serviço.

Nota-se, portanto, que os padrões elencados apresentam características que favorecem o registro e o armazenamento correto dos dados de um objeto digital, sendo utilizados deste modo em iniciativas para a descrição e preservação em ambientes digitais como bibliotecas digitais, repositórios digitais e catálogos eletrônicos. Sendo assim, conforme os dados coletados na literatura, o padrão de metadados PREMIS está empregado em um repositório digital, enquanto que o esquema MODS é adotado em um repositório digital e em um catálogo eletrônico, ao passo que o modelo METS foi identificado em duas bibliotecas digitais.

5 CONSIDERAÇÕES FINAIS

A informação produzida pelo homem, encontradas em formas de conhecimentos, de memórias, de atividades entre outras dimensões, corre o risco de se perder ou de não ser corretamente resguardada à posteriores, especialmente pelos recentes adventos das TIC. A mencionada preocupação vai ao encontro da temática memória, compreendida aqui como o resultado histórico de indivíduos e da coletividade, que tem na produção documental uma das principais maneiras de desenvolvê-la.

Se por um lado as TIC facilitam e intensificam a produção e divulgação de informação e subsidiam soluções, essa intensificação ocorre de forma desordenada, alimentando uma espécie de nova explosão informacional. Ainda na perspectiva das TIC, outro ponto a ser observado refere-se à migração das informações, como de formatos físicos aos digitais, como de versões antigas às atualizadas. Desta forma, na tentativa de resguardar informações, alerta-se para o cerne discursivo, ou seja, a forma como as informações estão sendo conservadas, o que pode gerar, ou não, em dificuldades quanto a guarda e, principalmente, quanto a recuperação.

Com este cenário em vislumbre, identificaram-se Instituições, que lidam com a produção, guarda e disseminação informacional, engajadas na busca e desenvolvimento de estratégias que auxiliem na efetiva conservação de informações em meio digital como o uso de padrões, entre eles o de descrição de registros, de inserção de conteúdo na *web*, de guarda de informações, de objetos digitais etc., o que, teoricamente, representam exemplos que facilitam e melhoram o processo de preservação de informação em meio digital.

Neste trabalho, buscou-se trazer a importância da salvaguarda e da preservação digital de informações enquanto resquícios da memória humana, tanto de conteúdos produzidas no próprio meio eletrônico, quanto os que são originários da forma física e que se transformam em digitais, por meio de métodos como os de migração, conversão e emulação. O certame literário e discursivo permitiu identificar medidas de preservação digital a nível de preservação de metadados que já estão sendo discutidas com fins de desenvolverem-se métodos em meio digital. Para tanto, o presente TCC destacou os métodos estruturais, pois são medidas preventivas de preservação. Em outras palavras, a ênfase aqui discursiva esteve em linha a adoção de padrões e aos metadados de preservação digital, preocupados na descrição de conteúdos e na guarda das informações que cerceiam os objetos digitais. Ou seja, trata-se de uma especificidade embasada nas características dos metadados, que são dados responsáveis

pela descrição de outros dados e informações, tendo como grande exemplo o esquema DC, padrão de metadados usualmente utilizado na descrição de objetos digitais na *web*.

A importância em se discutir métodos de preservação digital alimenta o encargo sobre a recente explosão informacional, que a sociedade atual se faz presente, pois traz à tona questões sobre o armazenamento, a preservação e o acesso de informação, independente do espaço e do tempo, enquanto ações realizadas pelos cidadãos e que subsidiam a construção da memória da ciência e da própria humanidade. Neste sentido, a OIAS se apresenta como um modelo de referência na estruturação de ações que objetivam a preservação digital a partir de instruções em níveis, entre administrativos e de descrição dos objetos digitais. Somando-se ao OAIS, tem-se os padrões de metadados, como o modelo PREMIS, que representa um conjunto de elementos básicos e aplicáveis principalmente a repositórios digitais que objetivam a preservação digital, por meio de campos como o *Events*, além de ser o único com foco nessa ação entre os identificados na pesquisa. O modelo MODS, outro padrão, traz consigo um conjunto de metadados flexíveis e que permitem um maior detalhamento nas descrições, mas que oferece, mesmo não sendo o seu principal objetivo, elementos à preservação digital. Por fim, tem-se o modelo METS que, assim como os demais exemplos, apresenta um conjunto de metadados descritivos, estruturais e administrativos, mediante linguagem XML, mas que, apesar de não ser o seu maior intuito, também oferece elementos que permitem a preservação digital, neste caso, dos dados dos objetos digitais, como o *Structural Map section*.

Ao analisar os métodos identificados na busca bibliográfica, especificamente a iniciativa OAIS e os formatos PREMIS, MODS e METS, conclui-se que existe o anseio pela utilização de padrões para descrever e armazenar documentos, bem como realizar a guarda e a preservação de forma padronizada e comunicável entre máquinas, aspiração que facilitaria a comunicação entre os centros de informação/Instituições em redes de informação via o conceito de interoperabilidade.

O presente elenco de exemplos de aplicações e uso de padrões de metadados para preservação dos dados de objetos digitais representam apenas realidades de um cenário maior, o que culmina na expectativa de que mais estudos sejam realizados e, conseqüentemente, outros métodos com tais finalidades de padronização sejam identificados, apresentados e exemplificados, o que contribuiria ainda mais na compreensão das formas como os objetos digitais são tratados, quanto aos dados que os descrevem e representam, e disponibilizados para o acesso, tanto no agora, como à posteriores.

REFERÊNCIAS

ALMEIDA, Maurício Barcellos. Uma introdução ao XML, sua utilização na Internet e alguns conceitos complementares. **Ciência da Informação**, Brasília, v. 31, n. 2, p.5-13, maio/ago. 2002. Disponível em: <<http://www.scielo.br/pdf/%0D/ci/v31n2/12903.pdf>>. Acesso em: 02 maio 2018.

ALVES, Bruno Henrique. Abordagens métricas: análise da produção científica de artigos e rede de colaboração científica dos docentes do programa de pós-graduação em ciência da informação, na linha de pesquisa organização da informação da UNESP/Marília. **Revista de Iniciação Científica da FFC**, São Paulo, v. 9, n. 2, p.104-115, 2009. Disponível em:<<http://www2.marilia.unesp.br/revistas/index.php/ric/article/view/248>>. Acesso em: 25 out. 2017.

ANDRADE, Maria Margarida de. **Introdução à metodologia do trabalho científico**. 4. ed. São Paulo: Atlas, 1999.

ARAÚJO, Maria Paula Nascimento; SANTOS, Myrian Sepúlveda dos. História, memória e esquecimento: Implicações políticas. **Revista Crítica de Ciências Sociais**, v. 79, p.95-111, 2007. Disponível em: <<http://www.ces.uc.pt/publicacoes/rccs/artigos/79/RCCS79-095-111-MPNascimento-MSepulveda.pdf>>. Acesso em: 18 set. 2017.

ARAUJO, Priscilla Mara Bermudes; SOUZA, Rosali Fernandez de; ARELLANO, Miguel Ángel Márdero. Preservação digital e os periódicos científicos eletrônicos brasileiros em ciência da informação. In: Encontro Nacional de Pesquisa em Ciência da Informação (XVI ENANCIB), 16, 2015, João Pessoa. **Anais**. Disponível em: <<http://www.ufpb.br/evento/lti/ocs/index.php/enancib2015/enancib2015/paper/viewFile/2829/1177>>. Acesso em: 25 out. 2017.

ARELLANO, Miguel Ángel Márdero. Cariniana: uma rede nacional de preservação digital. **Ciência da Informação**, Brasília, v. 41, n. 1, p.83-91, jan./abr. 2012. Disponível em: <http://eprints.rclis.org/23074/1/Artigo_miguel_CI.pdf>. Acesso em: 25 maio 2018.

_____. Preservação de documentos digitais. **Ciência da Informação**, Brasília, v. 33, n. 2, p.15-27, maio 2004. Disponível em: <<http://www.scielo.br/pdf/ci/v33n2/a02v33n2.pdf>>. Acesso em: 09 ago. 2017.

_____. **Preservação digital em repositórios confiáveis**. Rio de Janeiro: Mct/ibict, 2008. 41 slides, color. Segunda parte da apresentação do workshop sobre preservação digital em repositórios confiáveis no II CIPECC no Rio de Janeiro, dia 17 de novembro de 2008. Disponível em:<<https://pt.slideshare.net/gemireki/preservao-digital-em-repositrios-confiveis-part-ii-presentation>>. Acesso em: 25 out. 2017.

BIBLIOTECA NACIONAL DIGITAL. **Apresentação**. 2018. Disponível em: <<http://purl.pt/index/geral/PT/about.html>>. Acesso em: 03 jun. 2018.

BORBINHA, José Luís. *et al.* A gestão de obras digitalizadas na BND. In: 3ª CONFERÊNCIA NACIONAL XATA,2005, Braga, Portugal. **XML: Aplicações e Tecnologias Associadas**. Braga: Universidade do Minho, 2005. p. 167 - 175. Disponível em: <<http://xata.fe.up.pt/2005/papersfinal/43.pdf>>. Acesso em: 05 jul. 2018.

BROWN UNIVERSITY LIBRARY. **Welcome**. 2018. Disponível em: <<https://library.brown.edu/>>. Acesso em: 07 jul. 2018.

CAMPOS, Luiz Fernando de Barros. Metadados digitais: revisão bibliográfica da evolução e tendências por meio de categorias funcionais. **Encontros Bibli**, n. 23, 1. sem. 2007. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2007v12n23p16/390>>. Acesso em 24 abr. 2017.

CAMPOS, Fernanda Maria Guedes; SARAMAGO, Maria de Lurdes. Preservação digital de longo prazo em instituições patrimoniais: reutilização e adaptação de metadados. **Actas dos Congressos Nacionais de Bibliotecários, Arquivistas e Documentalistas**, Lisboa, n. 9, p. 1-7, 2007. Disponível em: <<http://www.bad.pt/publicacoes/index.php/congressosbad/article/view/540/330>>. Acesso em: 28 maio 2018.

CARINIANA. **Cariniana, Rede Brasileira de Serviços de Preservação Digital**: Brasília: Ibict, [20?]. 70 slides, color. Disponível em: <https://www.bc.ufg.br/up/88/o/Apresentação_CARINIANA_IBICT.pdf>. Acesso em: 25 maio 2018.

CAROLINA DIGITAL REPOSITORY. **About the Repository**. 2018. Disponível em: <<https://blogs.lib.unc.edu/cdr/>>. Acesso em: 30 maio 2018a.

_____. **PREMIS Overview**. 2018. Disponível em: <<https://blogs.lib.unc.edu/cdr/index.php/about/cdr-development-and-collab/technical-documentation/metadata/premis/premis-overview/>>. Acesso em: 30 maio 2018b.

_____. **Metadados de Preservação (PREMIS)**. 2018. Disponível em: <<https://blogs.lib.unc.edu/cdr/index.php/about/cdr-development-and-collab/technical-documentation/metadata/premis/preservation-metadata-premis/>>. Acesso em: 30 maio 2018c.

CLIO. **Gerenciamento de acervo digital**: apresentação. 2018. Disponível em: <<http://www.liber.ufpe.br/selos/index.php>>. Acesso em: 05 jul. 2018.

COLUMBIA DIGITAL LIBRARY. **Welcome**. 2018. Disponível em: <<https://dlc.library.columbia.edu/>>. Acesso em: 02 jun. 2018.

COPAC. **About us**. 2018a. Disponível em: <<https://copac.jisc.ac.uk/about/>>. Acesso em: 02 jun. 2018.

_____. **Exporting records**. 2018b. Disponível em: <<https://copac.jisc.ac.uk/help/export-help.html>>. Acesso em: 05 jul. 2018.

COPYRIGHT. **Copyright**. 2017. Disponível em: <<http://copyright.com.br/copyright.html>>. Acesso em: 27 dez. 2017.

CREATIVE COMMONS BR. **Sobre**: O que é Creative Commons. 2018. Disponível em: <<https://br.creativecommons.org/sobre/>>. Acesso em: 21 ago. 2018.

CUNHA, Murilo Bastos da. Das bibliotecas convencionais às digitais: diferenças e convergências. **Perspectivas em Ciência da Informação**, Brasília, v. 23, n. 1, p.2-17,

jan/abr. 2008. Disponível em:

<<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/221>>. Acesso em: 19 out. 2017.

_____.; OLIVEIRA, Cordélia Robalinho de. **Dicionário de biblioteconomia e arquivologia**. Brasília: Briquet de Lemos, 2008.

DAVIS, Daniel; WILPER, Chris. **Fedora Digital Object Model**. 2011. Disponível em:

<<https://wiki.duraspace.org/display/FEDORA34/Fedora+Digital+Object+Model#FedoraDigitalObjectModel-TheFedoraDigitalObjectobject>>. Acesso em: 30 maio 2018.

DICIONÁRIO AURÉLIO. **O Dicionário Online de Português**. 2018. Disponível em:

<<https://dicionariodoaurelio.com/>>. Acesso em: 02 maio 2018.

FERREIRA, Miguel. **Introdução à preservação digital: conceitos, estratégias e actuais consensos**. Guimarães, Portugal: Escola de Engenharia da Universidade do Minho, 2006.

Disponível em: <<http://repositorium.sdum.uminho.pt/handle/1822/5820>>. Acessado em: 17 out. 2017.

FERNANDES, Amaury. Notas Sobre a Evolução Gráfica do Livro. **Comum**. Rio de Janeiro: Faculdades Integradas Hélio Alonso, 2001. v.6, nº 17. p. 126 / 148. Disponível em:

<<http://www.producao Grafica.com.br/textos/NotEvGrafLiv.pdf>>. Acesso em: 11 out. 2017.

FONSECA FILHO, Clézio. **História da Computação: O caminho do pensamento e da tecnologia**. Rio Grande do Sul: Edipucrs, 2007.

FORMENTON, Danilo. **Identificação de padrões de metadados para preservação digital**. 2015. 103 f. Dissertação (Mestrado) - Curso do Programa de Pós-graduação em

Ciência, Tecnologia e Sociedade, Centro de Educação e Ciências Humanas, Universidade Federal de São Carlos, São Carlos, 2015. Disponível em:

<<https://repositorio.ufscar.br/handle/ufscar/7221>>. Acesso em: 28 maio 2018.

GONDAR, Jô. Memória individual, memória coletiva, memória social. **Morpheus: Revista Eletrônica em Ciências Humanas**, Rio de Janeiro, v. 13, n. 8, 2008. Disponível em:

<<http://www4.unirio.br/morpheusonline/numero13-2008/jogandar.htm>>. Acesso em: 18 set. 2017.

GRÁCIO, José Carlos Abbud. **Metadados para a descrição de recursos da Internet: o padrão Dublin Core, aplicações e a questão da interoperabilidade**. 2002. 127 f. Dissertação (Mestrado) - Curso de Ciência da Informação, Universidade Estadual Paulista, Marília, 2002.

Disponível em: <https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/gracio_jca_dr_mar.pdf>. Acesso em: 22 ago. 2017.

_____.; FADEL, Bárbara. Metadados para a preservação da informação digital. In: Encontro de Pesquisadores, 10, 2009, Franca, **Anais**, p. 354-362. Disponível em:

<<http://legacy.unifacfe.com.br/novo/3fem/Encontro/Arquivos/Jose%20Gracio%20e%20Barbara%20Fadel.pdf>>. Acesso em: 24 out. 2017.

HEERY, Rachel; PATEL, Manjula. **Application Profiles: Mixing and Matching Metadata Schemas**. 2000. Disponível em: <<http://www.ariadne.ac.uk/issue25/app-profiles/>>. Acesso em: 22 maio 2018.

INTERNATIONAL STANDARD BOOK NUMBER, Agência Brasileira do. **ISBN**. 2017. Disponível em: <<http://www.isbn.br/website/o-que-e-isbn>>. Acesso em: 27 dez. 2017.

INTERNATIONAL TELECOMMUNICATION UNION. **Universally Unique Identifiers (UUIDs)**. 2018. Disponível em: <<https://www.itu.int/en/ITU-T/asn1/Pages/UUID/uuids.aspx>>. Acesso em: 30 maio 2018.

JESUS, Joana D'arc Pereira de; KAFURE, Ivette. PRESERVAÇÃO DA INFORMAÇÃO EM OBJETOS DIGITAIS. **Biblionline**, João Pessoa, v. 6, n. 2, p.29-43, 2010. Disponível em: <periodicos.ufpb.br/index.php/biblio/article/download/7532/5516>. Acesso em: 17 out. 2017.

JULIO, Cristiane Aparecida; BOTELHO, Gabriela Evangelista. **Preservação de imagens fotográficas digitais: um estudo de caso na “Casa da Memória” da Biblioteca Pública “Vó Philomena” do Núcleo Bandeirante**. 2015. 64 f., il. Monografia (Bacharelado em Biblioteconomia) -Universidade de Brasília, Brasília, 2015. Disponível em: <<http://bdm.unb.br/handle/10483/11202>>. Acesso em: 05 set. 2017.

LABORATÓRIO DE TECNOLOGIA DO CONHECIMENTO. **Laboratório LIBER**. 2018. Disponível em: <<http://www.liber.ufpe.br/>>. Acesso em: 27 maio 2018.

_____. **Resumo**. 2018. Disponível em: <<https://br.linkedin.com/in/liber-ufpe-79929511b>>. Acesso em: 30 abr. 2018.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. Fundamentos de metodologia científica. 5. ed. São Paulo: Atlas, 2003. Disponível em: <https://docente.ifrn.edu.br/olivianeta/disciplinas/copy_of_historia-i/historia-ii/china-e-india>. Acesso em: 11 out. 2017.

LE GOFF, Jacques. Memória. In: LE GOFF, J. **História e memória**. 7. ed. Campinas: Editora da UNICAMP, 2013.

LOURENÇO, Cíntia Azevedo. Metadados: o grande desafio na organização da web. **Informação & Sociedade: Estudos**, João Pessoa, v. 17, n. 1, p.65-72, jan. 2007. Disponível em: <<http://www.periodicos.ufpb.br/ojs/index.php/ies/article/view/466/1466>>. Acesso em: 28 ago. 2018.

MAGALHÃES, Sara Alencar. **Rede da memória virtual brasileira: reflexão sobre os lugares de memória no ambiente virtual**. 2014. 104 f., il. Monografia (Bacharelado em Biblioteconomia) -Universidade de Brasília, Brasília, 2014. Disponível em: <<http://bdm.unb.br/handle/10483/8635>>. Acesso em: 05 set. 2017.

MDM WEB DOCS. **Lista de elementos do HTML5**. 2017. Disponível em: <https://developer.mozilla.org/pt-BR/docs/Web/HTML/HTML5/HTML5_element_list>. Acesso em: 24 mar. 2018.

MERLO, Franciele; KONRAD, Glaucia Vieira Ramos. Documento, história e memória: a importância da preservação do patrimônio documental para o acesso à informação. **Informação & Informação**, v. 20, n. 1, p.26-42, 22 mar. 2015. Universidade Estadual de Londrina. Disponível em:

<http://www.uel.br/revistas/uel/index.php/informacao/article/viewFile/18705/pdf_43>.

Acesso em: 05 set. 2017.

MEY, Eliane Serrão Alves; SILVEIRA, Naira Cristofoletti. **Catálogo no plural**. Brasília: Briquet Lemos, 2009.

MILETTO, Evandro Manara; BERTAGNOLLI, Silvia de Castro (Org.). **Desenvolvimento de Software II: Introdução ao desenvolvimento web com HTML, CSS, JAVASCRIPT e PHP**. Rio Grande do Sul: Bookman, 2014.

METADATA OBJECT DESCRIPTION SCHEMA. **MODS: Uses and Features**. 2016.

Disponível em: <<http://www.loc.gov/standards/mods/mods-overview.html>>. Acesso em: 02 jun. 2018.

_____. **MODS Implementation Registry**. 2010. Disponível em:

<<http://www.loc.gov/standards/mods/registry.php>>. Acesso em: 02 jun. 2018.

MSDN MICROSOFT. **Usar Namespaces XML**. 2018c. Disponível em:

<[https://msdn.microsoft.com/pt-br/library/ms754539\(v=vs.85\).aspx](https://msdn.microsoft.com/pt-br/library/ms754539(v=vs.85).aspx)>. Acesso em: 18 maio 2018c.

MUCHERONI, Marcos Luiz; SILVA, José Fernando Modesto da. A interoperabilidade dos sistemas de informação sob o enfoque da análise sintática e semântica de dados na web. **Ponto de Acesso**, Salvador, v. 5, n. 1, p.03-18, abr. 2011. Disponível em:

<<https://portalseer.ufba.br/index.php/revistaici/article/view/3622/3661>>. Acesso em: 03 nov. 2017.

MUNHOZ, Augusto Maciel. **Preservação digital de acervos: uma análise das políticas e estratégias utilizadas pelas universidades do Rio Grande do Sul**. 2011. 63 f. Monografia (Graduação) - Curso de Biblioteconomia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2011. Disponível em:<<http://www.lume.ufrgs.br/handle/10183/37473>>. Acesso em: 19 ago. 2017.

NATIONAL INFORMATION STANDARDS ORGANIZATION. **UnderstandingMetadata**.

2004. Disponível em: <<http://www.niso.org/publications/press/UnderstandingMetadata.pdf>>. Acesso em: 05 set. 2017.

NETWORK WORKING GROUP R. MOATS. **URN Syntax**. 1997. Disponível em:

<<https://www.ietf.org/rfc/rfc2141.txt>>. Acesso em: 30 maio 2018.

OLIVEIRA, Eliane Braga de. **O conceito de memória na Ciência da Informação no**

Brasil: uma análise da produção científica dos programas de pós-graduação. 2010. 196 p.

Tese (Doutorado) - Curso de Ciência da Informação, Universidade de Brasília, Brasília, 2010.

Disponível

em:<http://repositorio.unb.br/bitstream/10482/7466/1/2010_ElianeBragaOliveira.pdf>.
Acesso em: 22 ago. 2017.

OLIVEIRA, Silvio Luiz de. **Tratado de metodologia científica**. 2. ed. São Paulo: Editora Pioneira, 2000.

PATRÍCIO, Helena Simões. Evolução do modelo de informação da Biblioteca Nacional Digital. In: 12º CONGRESSO NACIONAL BAD, 12., 2015, Portugal. **12º Congresso Nacional BAD**. Portugal: Bad, 2015. p. 1 - 13. Disponível em:
<http://purl.pt/resources/apresentacoes/2015_1_Evolucao_modelo_informacao_BND_hpatricio.pdf>. Acesso em: 03 jun. 2018.

PIMENTA, Ricardo M. **O futuro do passado**: desafios entre a informação e a memória na sociedade digital. In: ALBAGLI, Sarita (Org.). *Fronteiras da Ciência da Informação*. Brasília: IBICT, 2013. p.146-171. Disponível em: < <http://ridi.ibict.br/handle/123456789/451>>. Acesso em: 05 set. 2017.

PRATER, Scott. **Fedora Object XML (FOXML)**. 2014. Disponível em:
<<https://wiki.duraspace.org/pages/viewpage.action?pageId=66585857>>. Acesso em: 30 maio 2018.

PRODANOV, Cleber Cristiano; FREITAS, Ernani Cesar de. **Metodologia do trabalho científico**: métodos e técnicas da pesquisa e do trabalho acadêmico. 2. ed. Novo Hamburgo: Feevale, 2013. 276 p. Disponível em: < http://www.feevale.br/Comum/midias/8807f05a-14d0-4d5b-b1ad-1538f3aef538/E-book_Metodologia_do_Trabalho_Cientifico.pdf >. Acesso em: 14 set. 2017.

PRESERVATION METADATA: IMPLEMENTATION STRATEGIES. **PREMIS, preservation metadata maintenance activity**. 2018. Disponível em:
<<http://www.loc.gov/standards/premis/index.html>>. Acesso em: 28 maio 2018.

REDE MEMORIAL DE PERNAMBUCO. **Origem**. Disponível em:
<<http://redememorialpernambuco.blogspot.com.br/p/origem-da-rede.html>>. Acesso em: 27 maio 2018.

ROCHA, Rafael Port da. **Metadados**: esquemas. 2018. Disponível em:
<http://www.ufrgs.br/snote/wiki/doc.php?u=/ensino/notasaula/metadados_-_esquemas>.
Acesso em: 02 jun. 2018.

RODRIGUES, Daniel. *et al.* **Descobrimo o OAIS**. 2013. Disponível em:
<<http://preservacaodigital5.blogspot.com.br/2013/01/descobrimo-o-oais.html>>. Acesso em: 24 mar. 2018.

RODRIGUES, Nelson de Almeida. Introdução ao METS: Preservação e Intercâmbio de Objetos Digitais. **Encontro Bibli**: Revista eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, v. 13, n. 26, p.172-187, 2º sem. 2008. Disponível em:
<<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2008v13n26p172>>. Acesso em: 05 jul. 2018.

RUDDOCK, Bethan. **Behind the Copac record 2**: MODS and de-duplication. 2010. Disponível em: <<https://blog.copac.jisc.ac.uk/tag/mods/>>. Acesso em: 05 jul. 2018.

SARACEVIC, Tefko. Interdisciplinarity nature of Information Science. Tradução de: Durval de Lara Filho. **Ciência da Informação**, Brasília, 1995 v.24, n.1, p.36-41. Disponível em: <<http://www.uff.br/ppgci/editais/saracevicnatureza.pdf>>. Acesso em: 25 set. 2017.

SARAMAGO, Maria de Lurdes. Metadados para preservação digital e aplicação do modelo OAIS. **Actas do Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas**, Lisboa, n. 8, p.1-6, maio 2004. Disponível em: <<https://www.bad.pt/publicacoes/index.php/congressosbad/article/view/640/637>>. Acesso em: 04 set. 2018.

SAYÃO, Luís Fernando. **Metadados para preservação digital: Aplicação do modelo OAIS**. Conselho Nacional de Energia Nuclear – Centro de Informações Nucleares – MCT/CNEN/CIN, 2007. 67 slides, color.

_____. Preservação digital no contexto das bibliotecas digitais: uma breve introdução. In: (ORG), Carlos H. Marcondes. *et al.* **Bibliotecas Digitais: saberes e práticas**. 2. ed. Brasília: Edufba, 2005. p. 127-150. Disponível em: <https://www.researchgate.net/publication/275031138_Preservacao_digital_no_contexto_das_bibliotecas_digitais_uma_breve_introducao>. Acesso em: 25 set. 2017.

_____. Uma outra face dos metadados: informações para a gestão da preservação digital. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 15, n. 30, p.1-31, 2010. Universidade Federal de Santa Catarina (UFSC). Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2010v15n30p1>>. Acesso em: 24 out. 2017.

SUMMER INSTITUTE OF LINGUISTICS (SIL). **ISO 639-3**. 2015. Disponível em: <<http://www-01.sil.org/iso639-3/default.asp>>. Acesso em: 27 dez. 2017.

SILVA, Edna Lúcia da; MENEZES, Estera Muszkat. **Metodologia da Pesquisa e Elaboração de Dissertação**. 4. ed. Florianópolis: UFSC, 2005.

SOUZA, Arthur Heleno Lima Rodrigues de. *et al.* O modelo de referência OAIS e a preservação digital distribuída. **Ciência da Informação**, Brasília, v. 41, n. 1, p.65-73, jan./abr. 2012. Disponível em: <http://basessibi.c3sl.ufpr.br/brapci/_repositorio/2014/11/pdf_20ae7f8b97_0031823.pdf>. Acesso em: 24 maio 2018.

SOUZA, Marcia Izabel Fugisawa; VENDRUSCULO, Laurimar Gonçalves; MELO, Geane Cristina. **Metadados para a descrição de recursos de informação eletrônica: utilização do padrão Dublin Core**. *Ci. Inf.* [online]. 2000, vol.29, n.1, pp.93-102. Disponível em: <http://www.scielo.br/scielo.php?pid=S0100-19652000000100010&script=sci_abstract&tlng=pt>. Acesso em: 24 out. 2017.

STANFORD UNIVERSITY. **What Is LOCKSS?** 2018. Disponível em: <<https://www.lockss.org/about/what-is-lockss/>>. Acesso em: 05 jul. 2018.

THOMAZ, Kátia de Padua; SOARES, Antônio José. A preservação digital e o modelo de referência Open Archival Information System (OAIS). **Datagramazero**, Rio de Janeiro, v. 5,

n. 1, p. A01, fev. 2004. Disponível em:

<<http://www.brapci.inf.br/index.php/article/view/0000007749/6f3e125671cfe45b7fddce5de6cf37f8>>. Acesso em: 24 maio 2018.

THE LIBRARY OF CONGRESS. **METS**. 2017. Disponível em:

<<http://www.loc.gov/standards/mets>>. Acesso em: 24 out. 2017.

_____. **Standards: ISO 639-2**. 2017. Disponível em: <http://www.loc.gov/standards/iso639-2/php/code_list.php>. Acesso em: 11 fev. 2018.

TRIBUNAL DE JUSTIÇA DE PERNAMBUCO. **Memorial Digital**. 2018. Disponível em:

<<http://www.tjpe.jus.br/web/memorial-da-justica/digital>>. Acesso em: 27 maio 2018.

VIEIRA, Itala Maduell. A memória em Maurice Halbwachs, Pierre Nora e Michel Pollak.

In: Encontro Regional Sudeste de História Oral, 11, 2015, Niterói. **Anais**. Disponível em:

<http://www.sudeste2015.historiaoral.org.br/resources/anais/9/1429129701_ARQUIVO_Memoria_Itala_Maduell.pdf>. Acesso em: 11 out. 2017.

WORLD WIDE WEB CONSORTIUM. **XML ESSENTIALS**. 2015c. Disponível em:

<<https://www.w3.org/standards/xml/core>>. Acesso em: 18 maio 2018.

WORLD WIDE WEB CONSORTIUM BRASIL. **Curso HTML5**. São Paulo: W3c Brasil, 2011c. Disponível em: <<http://www.w3c.br/Cursos/CursoHTML5>>. Acesso em: 18 maio 2018.

_____. **Sobre o W3C**. 2011c. Disponível em: <<http://www.w3c.br/Sobre/>>. Acesso em: 1 maio 2018.

WORLD WIDE WEB CONSORTIUM SCHOOLS. **Tag HTML**. 2018c. Disponível em:

<https://www.w3schools.com/tags/tag_body.asp>. Acesso em: 18 maio 2018.