

ENTENDENDO O PREMIS

Library of Congress Network Development and MARC Standards Office

Autora: Priscilla Caplan

Revisão: PREMIS Editorial Committee

Publicado originalmente em 2009; revisado em 2017

Copyright © 2009 The Library of Congress, exceto nos Estados Unidos da América

Os créditos devem ser atribuídos quando forem utilizados trechos desta publicação.

Tradução em 2021: Amarílis Montagnolli Gomes Corrêa (Mestre em Ciência da Informação pela Escola de Comunicações e Artes da USP, bibliotecária na Universidade de São Paulo) e Laerte Pereira da Silva Júnior (Doutor em Informação e Comunicação em Plataformas Digitais pela Universidade do Porto e Universidade de Aveiro, Analista de Tecnologia da Informação no Centro de Ciências Humanas, Letras e Artes da Universidade Federal da Paraíba)

Revisão (português): Acácia Montagnolli

Sumário

ENTENDENDO O PREMIS.....	2
1. O CONTEXTO DO PREMIS.....	2
1.1. O que são metadados de preservação?.....	2
1.2. O que é PREMIS?	3
1.3. O que tem no Dicionário de Dados PREMIS?	3
1.4. Como o PREMIS deve ser usado?.....	5
1.5. Você deveria estar usando o PREMIS?	5
2. CONCEITOS-CHAVE DO PREMIS	6
2.1. Unidades Semânticas.....	6
2.2. Contêineres e Subunidades.....	7
2.3. Contêineres de Extensão.....	7
3. O MODELO DE DADOS PREMIS.....	8
3.1. Entidade Objeto.....	8
3.2. Eventos.....	11
3.3. Agentes	11
3.4. Direitos	12
4. O DICIONÁRIO DE DADOS	13
4.1. Amostra de entrada do Dicionário de Dados para uma unidade semântica simples.....	13
4.2. Amostra de entrada do Dicionário de Dados para uma unidade contêiner	15
5. PREMIS EM USO.....	16
5.1. PREMIS em XML.....	16
5.2. Conformidade com o PREMIS.....	16
6. PARA MAIS INFORMAÇÕES	18
Apêndice A: Exemplos.....	19
A.1. Exemplo de Objeto.....	19
A.2. Exemplo de Evento	22
A.3. Exemplo de Agente.....	23
Apêndice B: Glossário	24

ENTENDENDO O PREMIS

Este guia é uma breve apresentação do padrão de metadados de preservação PREMIS. Ele não fornecerá informações suficientes para a implementação do PREMIS, mas dará uma ideia geral do que é o PREMIS. Para muitos leitores, isso será suficiente. Para aqueles que precisam desbravar as mais de 250 páginas do *Dicionário de Dados PREMIS para Metadados de Preservação*,¹ este guia pode servir como uma introdução que tornará o *Dicionário* mais assimilável.

I. O CONTEXTO DO PREMIS

I.1. O que são metadados de preservação?

Se você trabalha em uma biblioteca, arquivo ou museu, é bem provável que saiba algo sobre metadados e descrição de recursos. Os metadados são categorizados conforme sua finalidade: metadados descritivos auxiliam na descoberta e na identificação de recursos, metadados administrativos auxiliam no gerenciamento e rastreamento, e metadados estruturais indicam como objetos digitais complexos são combinados para que possam ser devidamente renderizados. Da mesma maneira, os *metadados de preservação*² dão apoio às atividades com propósito de garantir a usabilidade de um recurso digital a longo prazo.

O Dicionário de Dados PREMIS define metadados de preservação como “as informações que um repositório usa para amparar o processo de preservação digital”. A seguir estão alguns exemplos de atividades de preservação e de como os metadados dão esse apoio.

- Um recurso deve ser armazenado de maneira segura para evitar que seja modificado por descuido (ou por malícia). Uma soma de verificação armazenada como metadado pode ser usada para comprovar se um arquivo sofreu alguma alteração em um determinado intervalo de tempo.
- Arquivos devem ser armazenados em mídias que possam ser lidas por computadores atuais. Se as mídias estiverem danificadas ou obsoletas (como os disquetes de 8 polegadas usados nos anos 1970), pode ser difícil ou impossível restaurar os dados. Os metadados podem dar apoio ao gerenciamento das mídias de armazenamento registrando seu tipo e idade e as datas em que os arquivos foram refrescados pela última vez.
- Com o passar do tempo, mesmo os formatos de arquivo mais populares podem se tornar obsoletos, impedindo que os sistemas correntes possam renderizá-los. Para garantir que os recursos permaneçam utilizáveis, os gestores devem adotar *estratégias de preservação* como a migração de formatos antigos para equivalentes mais novos ou a emulação do antigo ambiente de renderização com hardware e software mais recentes. Ambas estratégias, *migração* e *emulação*, requerem metadados sobre os formatos originais dos arquivos e os ambientes de hardware e software que os suportam.

¹ O título original é *PREMIS Data Dictionary for Preservation Metadata*. [Esta e as demais notas são do tradutor.]

² No texto original o recurso gráfico itálico foi usado para destacar os termos que constam no glossário (em geral, sempre na primeira ocorrência), termos específicos do PREMIS, como os tipos de objetos e de entidades, entre outros. Nesta tradução, optou-se por manter esse mesmo uso, além de aplicar o itálico para destacar título de publicações e termos sem tradução para o português.

- Estratégias de preservação podem modificar os recursos originais (migração) ou a maneira como eles são renderizados (emulação). Isso pode colocar em xeque a autenticidade do recurso. Os metadados podem amparar a autenticidade documentando a *proveniência digital* do recurso – sua cadeia de custódia e o histórico de modificações autorizadas.

1.2. O que é PREMIS?

PREMIS provém de **PRE**servation **Meta**data: **Im**plementation **Str**ategies (Metadados de preservação: estratégias de implementação). É também o nome do grupo de trabalho internacional que foi patrocinado pela OCLC e pelo RLG³ entre 2003 e 2005. Esse grupo de trabalho produziu o relatório intitulado *PREMIS Data Dictionary for Preservation Metadata* (Dicionário de Dados PREMIS para Metadados de Preservação),⁴ que inclui o dicionário de dados e uma explanação sobre metadados de preservação. Em seguida, a Biblioteca do Congresso dos Estados Unidos publicou um conjunto de esquemas para representar em XML os elementos de metadados do Dicionário de Dados PREMIS. Uma segunda edição atualizada do Dicionário e um novo esquema de apoio foram publicados em março de 2008, seguidos de pequenas revisões, versões 2.1 e 2.2. Em junho de 2015, foi publicada uma revisão mais extensa, a versão 3.0.

Existe uma Atividade de Manutenção ativa do PREMIS (*PREMIS Maintenance Activity*), patrocinada pela Biblioteca do Congresso dos Estados Unidos, que inclui um site com todas as informações oficiais e extraoficiais sobre o PREMIS, uma lista de discussão, uma *wiki* para os implementadores do PREMIS e um comitê editorial responsável pelas revisões do Dicionário de Dados e do esquema XML. A Atividade de Manutenção busca, também, promover a conscientização sobre o PREMIS, patrocinar tutoriais e reuniões de implementação sobre o uso do PREMIS e encarregar estudos e publicações relacionadas ao PREMIS, como este guia.

Em geral, quando as pessoas usam o termo “PREMIS” elas se referem ao Dicionário de Dados. Porventura, podem se referir ao esquema XML, ao grupo de trabalho ou à iniciativa como um todo, incluindo a Atividade de Manutenção.

Dicionário de Dados PREMIS: www.loc.gov/premis/v3/premis-3-0-final.pdf

Site PREMIS: www.loc.gov/premis

Lista de discussão do Grupo de Implementadores do PREMIS: pig@listserv.loc.gov

para se inscrever envie e-mail para listserv@listserv.loc.gov com a mensagem:
subscribe pig [Seu Nome]

1.3. O que tem no Dicionário de Dados PREMIS?

O Dicionário de Dados PREMIS define um conjunto principal de elementos de metadados, ou melhor, unidades semânticas (falaremos sobre isso adiante),⁵ que os repositórios devem considerar para desempenhar suas funções de preservação. Essas funções podem variar de um repositório para outro, mas em geral incluem ações para assegurar que os objetos digitais permaneçam viáveis e renderizáveis (ou seja, que possam ser lidos a partir de mídias e exibidos, executados ou interpretados por um

³ OCLC – Online Computer Library Center e RLG – Research Library Group.

⁴ Daqui em diante será usado o título do relatório traduzido em português.

⁵ Na seção 2.1. Unidades semânticas.

software) e incluem ações para garantir que os objetos não sejam modificados por descuido e que as alterações legítimas sejam documentadas.

O Dicionário de Dados não tem por objetivo definir todos os possíveis elementos de metadados de preservação, apenas aqueles mais comuns para a maioria dos repositórios. Várias categorias de metadados estão fora do escopo, entre elas:

- Metadados de formatos específicos, ou seja, metadados que pertencem a apenas um formato de arquivo ou a uma classe de formatos, como áudio, vídeo ou gráficos vetoriais.
- Metadados específicos de implementação e regras de negócio, ou seja, metadados que descrevem as políticas ou práticas de um repositório em particular, tais como aquelas que preconizam o acesso aos materiais.
- Metadados descritivos. Embora a descrição dos recursos seja obviamente relevante para a preservação, muitos padrões independentes podem ser usados para essa finalidade, como MARC, MODS e Dublin Core.
- Informações detalhadas sobre mídias ou hardware. Mais uma vez, embora isso seja obviamente relevante para a preservação, a definição desses metadados compete a outras comunidades.
- Informações detalhadas sobre agentes (pessoas, organizações ou softwares) além das que são necessárias para identificação.
- Informações extensas sobre direitos e permissões. O foco do PREMIS está naquelas que são essenciais às funções de preservação.

Se você pensar em todos os metadados que uma organização precisa para gerenciar um repositório de preservação, o PREMIS pode ser entendido como o definidor de um subconjunto central. Ele não se dedica à recuperação e ao acesso nem se propõe à definição detalhada de metadados de um formato específico. Ele define apenas os metadados frequentemente necessários para se executar as funções de preservação em todos os materiais.

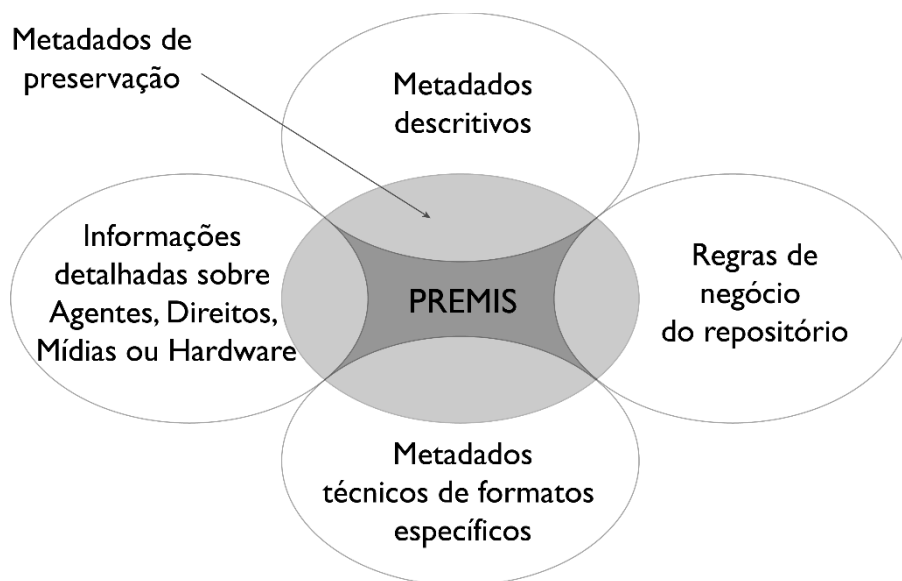


Figura 1: PREMIS como um subconjunto de todos os Metadados de Preservação

A Figura 1 mostra todos os metadados de preservação como o círculo no centro do diagrama. Ele inclui alguns metadados descritivos, algumas regras de negócio, alguns metadados técnicos detalhados e algumas informações detalhadas sobre agentes, direitos, mídias e hardware. O PREMIS é o núcleo dos metadados de preservação, que exclui todos os demais tipos.

I.4. Como o PREMIS deve ser usado?

O Dicionário de Dados PREMIS define o que um repositório de preservação precisa saber. É importante notar que o foco está no sistema do repositório e seu gerenciamento, e não nos autores de conteúdo digital, nas pessoas que escaneiam ou convertem de alguma outra maneira conteúdo analógico para formato digital ou na equipe que avalia e licencia recursos eletrônicos comerciais. Os principais usos do PREMIS são para o projeto e a avaliação do repositório e para o intercâmbio de pacotes de informações arquivadas.

Aqueles que projetam e/ou desenvolvem softwares para repositórios de preservação devem usar o PREMIS como um guia das informações que devem ser obtidas e registradas pelo sistema ou, ao menos, conhecidas para o gerenciamento do repositório.

Aqueles que estão planejando implementar um repositório de preservação devem usar o PREMIS como uma lista de parâmetros para avaliar os softwares disponíveis a essa finalidade. Os sistemas que suportam o Dicionário de Dados PREMIS serão mais capazes de preservar recursos informacionais a longo prazo.

Eventualmente, um repositório em funcionamento precisará exportar pacotes de informações arquivadas para ingestão em outro repositório. Por exemplo, uma organização de custódia poderá migrar de um sistema de repositório para outro, um cliente poderá trocar de serviço terceirizado de armazenamento ou uma instituição poderá usar os serviços de preservação do repositório digital confiável de outra instituição. O PREMIS fornece um conjunto comum de elementos de dados que podem ser entendidos por ambos repositórios, o que exporta e o que importa informações, sobretudo se for usado o esquema PREMIS XML.

I.5. Você deveria estar usando o PREMIS?

Depende. A maioria dos funcionários de bibliotecas, arquivos, museus e outras organizações de patrimônio cultural não tem nenhum envolvimento direto com preservação digital. Nesse caso, basta você saber o que é o PREMIS: um dicionário de dados para metadados de preservação. Se o seu trabalho implicar alguma responsabilidade sobre qualquer aspecto da preservação digital, você provavelmente achará útil entender o PREMIS. Se estiver envolvido com a avaliação ou implementação de um repositório institucional ou sistema de preservação, você deverá entender bem o PREMIS. Considere utilizar os materiais de capacitação disponíveis no site do PREMIS ou participar dos tutoriais quando forem oferecidos.

Se você trabalha em projetos de digitalização, pode estar se perguntando se deveria criar metadados PREMIS para uso posterior. Muitos dos elementos PREMIS estão projetados para serem fornecidos automaticamente pelo sistema do repositório de preservação (claro que isso não significa que todos os sistemas atualmente disponíveis executam essa função). Entretanto, há algumas informações que você deve registrar, se possível:

Inibidores. *Inibidores* são definidos como qualquer característica de um objeto cujo propósito é inibir acesso, uso ou migração. Inibidores incluem senha de proteção e criptografia. O software do repositório pode não conseguir identificar um inibidor porque este impede que o objeto seja analisado, por isso, se você sabe que um arquivo tem inibidores, é importante documentá-los. O PREMIS define unidades semânticas para tipo de inibidor, alvo (as ações inibidas) e chave (senha ou outro mecanismo para driblar o inibidor).

Proveniência. *Proveniência digital* é o registro da cadeia de custódia e do histórico de modificações de um objeto digital. Se a instituição criou o objeto, as circunstâncias dessa criação são obviamente uma parte importante da proveniência. O nome e a versão do programa de criação, e a data de criação, em geral, podem ser extraídos do cabeçalho do arquivo. No entanto, como nem sempre é possível, recomenda-se registrar essas informações. O PREMIS permite que o histórico de alterações seja registrado como informação de Evento,⁶ que está descrito na seção 3.2. Eventos. Um vocabulário controlado para tipos de evento está disponível em: id.loc.gov/vocabulary/preservation/eventType.

Muitos dos tipos de Evento PREMIS estão projetados para descrever ações que acontecem depois que algo é submetido à ingestão em um repositório, embora alguns dos tipos sejam usados para eventos anteriores a essa etapa, como captura e aquisição.

Propriedades significativas. *Propriedades significativas* são as características de um objeto que devem ser mantidas pelas ações de preservação. Por exemplo, se você tem um documento, nele são importantes apenas as palavras e as imagens ou também as fontes, o plano de fundo, a formatação e outras características visuais? Embora propriedades significativas seja um dos conceitos menos compreendidos, é um dos mais fundamentais em preservação digital. Portanto, toda instituição que cria ou coleciona materiais digitais para uma comunidade de usuários deve identificar cuidadosamente quais características desses materiais são importantes para essa comunidade e, então, tentar registrá-las como propriedades significativas.

Direitos. Informações sobre direitos não são exclusivas da preservação, mas é muito importante para esse processo saber o que se pode fazer com um objeto. Toda informação disponível sobre direitos deve ser registrada, por exemplo, copyright, licenças e permissões especiais.

2. CONCEITOS-CHAVE DO PREMIS

2.1. Unidades Semânticas

O Dicionário de Dados PREMIS define *unidades semânticas*, e não elementos de metadados. A distinção é sutil mas existe. Uma unidade semântica é um fragmento de informação. Um elemento de metadados é uma maneira definida de se representar essa informação em um registro de metadados, esquema ou base de dados. O PREMIS não especifica como os metadados devem ser representados em um sistema; ele define apenas o que o sistema precisa saber e o que deve conseguir exportar para outros sistemas. Sendo assim, um purista em PREMIS pensa em unidades semânticas. Para os demais, é suficiente entendê-las como elementos de metadados.

⁶ O uso de letra maiúscula para termos específicos do Dicionário de Dados seguirá o padrão usado no *Understanding PREMIS* (2017).

Os nomes das unidades semânticas PREMIS são escritas em estilo *camelCase*. Ou seja, as palavras não são separadas por espaços, mas escritas juntas e com cada letra inicial maiúscula. Exemplos: `objectIdentifier`,⁷ `relatedEventIdentification`.

2.2. Contêineres e Subunidades

Algumas unidades semânticas são definidas como *contêineres*,⁸ o que significa que elas não têm valor próprio mas existem para agrupar unidades semânticas relacionadas. Por exemplo, sempre que se registra um identificador no PREMIS você tem que dizer que tipo de identificador ele é (DOI, ISBN, identificação atribuída pelo sistema local, etc.). O contêiner `objectIdentifier` é usado para reunir as duas subunidades `objectIdentifierType` e `objectIdentifierValue`.

Os contêineres estabelecem uma estrutura hierárquica no Dicionário de Dados que se reflete na numeração das unidades semânticas:

1.1 <code>objectIdentifier</code> (M, R)
1.1.1 <code>objectIdentifierType</code> (M, NR)
1.1.2 <code>objectIdentifierValue</code> (M, NR)

Esse trecho do Dicionário de Dados mostra que o elemento `objectIdentifier`, obrigatório (M)⁹ e repetível (R), não tem valor próprio, mas serve como contêiner para os elementos componentes `objectIdentifierType` e `objectIdentifierValue`. Como os elementos tipo e valor não são repetíveis (NR) no contêiner, você teria que repetir a estrutura inteira do contêiner para registrar dois identificadores diferentes.

2.3. Contêineres de Extensão

Um *contêiner de extensão* é um tipo especial de contêiner que não tem subunidades definidas. Ele foi concebido para viabilizar o registro de metadados não PREMIS. Desse modo, o PREMIS pode ser expandido para incluir metadados de fora de seu escopo, ou que por alguma razão não foram incluídos no Dicionário de Dados. Contêineres de extensão têm “Extensão” na última parte de seus nomes.

Por exemplo, metadados técnicos específicos de um formato não estão incluídos no PREMIS porque são considerados fora do escopo, mas são informações muito importantes para a preservação digital. O contêiner de extensão `objectCharacteristicsExtension` possibilita registrar os metadados técnicos definidos por outros esquemas, como o dicionário de dados Z39.87 para imagens bitmap (seu esquema XML é denominado NISO MIX - Esquema em XML de Metadados para Imagens).¹⁰

Se você tem familiaridade com XML, notará que o Dicionário de Dados PREMIS foi desenhado para ser compatível com essa linguagem. As unidades semânticas PREMIS podem ser implementadas como elementos XML; as unidades contêiner são elementos que aceitam apenas outros elementos como conteúdo, e as unidades de extensão são contêineres para elementos definidos por esquemas externos.

⁷ O nome das unidades semânticas foi mantido em inglês, pois em eventual implementação do PREMIS por um repositório, assim serão identificadas.

⁸ A entrada no glossário é: Unidades contêiner.

⁹ M de *mandatory*, em inglês.

¹⁰ O título em inglês é *NISO Metadata for Images in XML Schema – NISO MIX*.

Há mais informação sobre PREMIS e XML na seção 5.1. PREMIS em XML. Esforços posteriores produziram a Ontologia PREMIS OWL¹¹ para ser usada em aplicações de dados interligados (*linked data*). A primeira versão da ontologia era compatível com o PREMIS versão 2.2 e agora com a versão 3.0. Informações sobre a Ontologia PREMIS OWL estão disponíveis em: www.loc.gov/standards/premis/ontology

3. O MODELO DE DADOS PREMIS

Um dos princípios fundamentais do PREMIS é descrever com clareza. Nas versões 1 e 2, o PREMIS definia cinco tipos de *Entidades* sobre as quais você podia falar: Entidade Intelectual, Objeto, Agente, Evento e Direitos (um diagrama do modelo de dados correspondente às versões 1 e 2 está disponível no site do PREMIS). Na versão 3, lançada em junho de 2015, o Modelo de Dados PREMIS foi revisado para transformar *Entidade Intelectual* em outra categoria de Objeto.

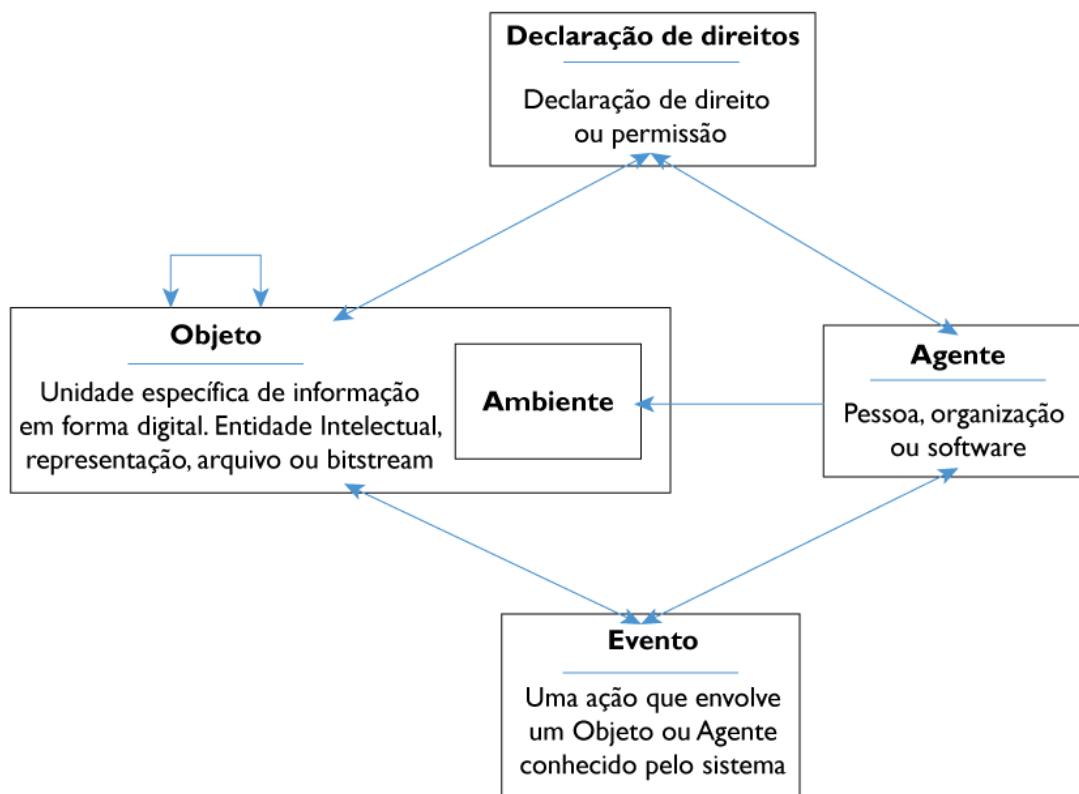


Figura 2: Modelo de Dados PREMIS, versão 3

3.1. Entidade Objeto

Objetos são o que realmente se armazena e gerencia em um repositório de preservação. A maior parte do PREMIS está dedicada a descrever objetos digitais. As informações que podem ser registradas são:

- um identificador único para o Objeto (tipo e valor);
- informação de fixidez, como uma soma de verificação e o algoritmo usado para gerá-la;
- o tamanho do Objeto;

¹¹ O título em inglês é *PREMIS OWL Ontology*.

- o formato do Objeto, que pode ser especificado diretamente ou através da ligação com um registro externo de formatos;
- o nome original do Objeto;
- informação sobre sua criação;
- informação sobre inibidores;
- informação sobre suas propriedades significativas;
- informação sobre seu ambiente (veja na p. 10);
- informação sobre onde e em qual mídia está armazenado;
- informação sobre assinatura digital;
- relações com outros Objetos e outros tipos de Entidades.

O PREMIS define quatro tipos diferentes de Objetos e exige que os implementadores façam uma distinção entre eles. Os Objetos são Bitstreams, Arquivos, Representações e Entidades Intelectuais.

Objeto Arquivo é exatamente o que parece, um arquivo de computador como um PDF ou JPEG.

Objetos Bitstream são subconjuntos de arquivos. Um Objeto Bitstream é definido como dados (bits) em um arquivo que: a) tem propriedades comuns para os fins de preservação; b) não é independente sem adição de um cabeçalho de arquivo ou outra estrutura. Por exemplo, se você tiver um arquivo em formato AVI (áudio-vídeo intercalados), poderá querer distinguir o bitstream de áudio do bitstream de vídeo e descrever cada um como Objetos Bitstream distintos.

Objeto Representação é o conjunto de todos os Objetos Arquivo necessários para renderizar uma Entidade Intelectual. Por exemplo, digamos que você queira preservar uma página web, talvez a página inicial do site da sua instituição a partir de uma data. Provavelmente, a página inicial que você vê no seu navegador está composta por muitos arquivos diferentes – um ou mais arquivos HTML, algumas imagens em formato GIF ou JPEG, talvez áudios ou animações em formato Flash. É provável, também, que tenha sido usada uma folha de estilos para criar a exibição. São necessários todos esses arquivos juntos para visualizar a página inicial em um navegador; portanto, se um repositório pretende preservar essa página como tal, renderizável, ele tem que conhecer todos esses arquivos e saber como combiná-los. O Objeto Representação permite ao repositório não apenas identificar o conjunto de arquivos relacionados, como também descrever as características da totalidade (por exemplo, a página web como um todo) que podem ser diferentes das características de cada parte em si.

Objeto Entidade Intelectual é definido como um conjunto de conteúdos que é considerado uma única unidade intelectual para fins de gerenciamento e descrição, por exemplo, um livro, mapa, fotografia ou base de dados. O PREMIS, em geral, não define metadados descritivos para as Entidades Intelectuais, porque existem muitos padrões para essa finalidade. Nas versões 1 e 2, Entidades Intelectuais só podiam ser referenciadas com um identificador, porque eram consideradas apenas conceituais. Mas na versão 3, uma Entidade Intelectual pode ser descrita com metadados descritivos não PREMIS **ou** com metadados de preservação como um Objeto PREMIS. Na maioria dos casos, as unidades semânticas usadas para Entidades Intelectuais são as mesmas que para Representações.

O PREMIS determina que um Objeto em um sistema de preservação deve estar associado à Entidade Intelectual conceitual representada por ele, incluindo um identificador da Entidade Intelectual nos metadados do Objeto. Por exemplo, se estamos preservando uma cópia do livro *Buddhism: The Ebook*:

an *Online Introduction* devemos usar o ISBN como link para a descrição da Entidade Intelectual na descrição PREMIS do livro.

Algumas unidades semânticas definidas no Dicionário de Dados PREMIS são aplicáveis a todos os quatro tipos de Objeto, enquanto outras são aplicáveis apenas a um ou dois tipos.

Várias unidades semânticas são definidas para registrar o *ambiente* de um Objeto, ou seja, o hardware e o software necessários para renderizá-lo e suas dependências em relação a outros Objetos. Nas versões 1 e 2, essas informações eram parte da descrição do Objeto. Na versão 3, o Ambiente, como um tipo especial de Objeto Entidade Intelectual, pode estar vinculado aos Objetos Arquivos, Representações e Bitstreams que o utilizam. Por exemplo, um arquivo PDF pode ser exibido por várias versões de Adobe Acrobat e Adobe Reader, assim como por outros programas de código aberto ou comerciais. Cada um deles, por sua vez, está baseado em diversos sistemas operacionais e requerem certas especificações mínimas de hardware (velocidade de processador, memória e disco). Como o Adobe Reader não é um programa autônomo, mas um plug-in para navegadores da internet, ele depende também de certas versões de determinados navegadores, por exemplo, o Reader 9 para Mac OS precisa do navegador Safari versão 2.0.4 ou posterior. Informações sobre o Ambiente são cruciais para algumas estratégias de preservação, mas são complicadas e demandam tempo para serem determinadas e podem ser melhor registradas em centrais como o PRONOM, mantido pelo Arquivo Nacional do Reino Unido. O PREMIS permite que os repositórios se conectem a informações armazenadas em registros externos quando essa for a alternativa preferível a armazená-las localmente.

Para ilustrar as relações entre Objetos, considere de novo a página web mencionada anteriormente. A página web é uma Entidade Intelectual que pode ser expressa em diversas Representações. Uma Representação, como foi detalhada no exemplo, consiste em Objetos Arquivo específicos (página HTML, imagens, folha de estilos, etc.). No entanto, a página web também pode ser arquivada como um único arquivo em formato Web ARChive (WARC) (para mais detalhes, consulte www.digitalpreservation.gov/formats/fdd/fdd000236.shtml). Esta é uma Representação diferente da mesma Entidade Intelectual, contendo apenas um arquivo, como se pode observar:

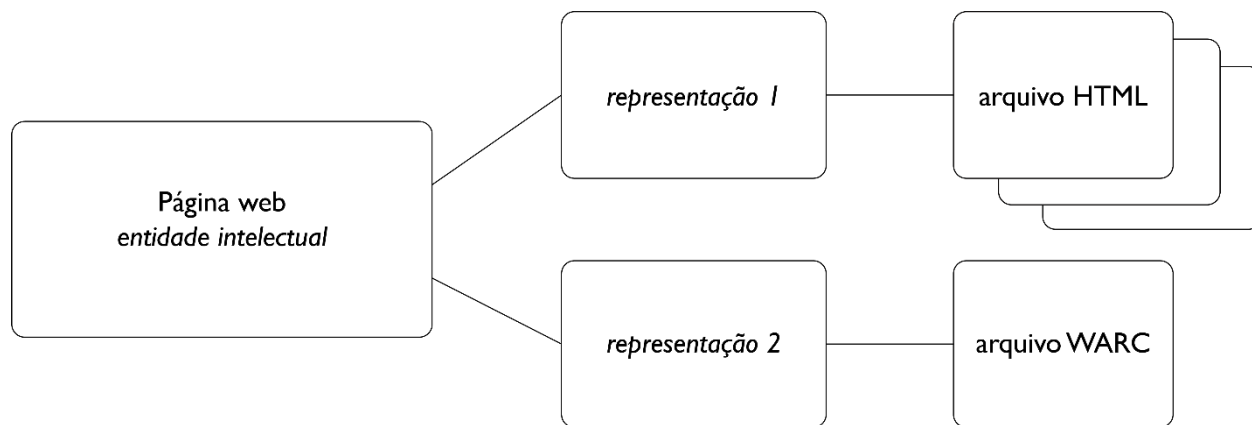


Figura 3: Exemplo de relação do Objeto PREMIS

3.2. Eventos

A entidade *Evento* agrega informações sobre as ações que afetam os Objetos no repositório. Um registro preciso e confiável dos Eventos é essencial para manter a proveniência digital de um Objeto, que por sua vez é imprescindível para demonstrar a autenticidade deste.

As informações que podem ser registradas sobre Eventos são:

- um identificador único para o Evento (tipo e valor);
- o tipo de Evento (criação, ingestão, migração, etc.);
- a data e hora em que o Evento ocorreu;
- uma descrição detalhada do Evento;
- um resultado codificado do Evento;
- uma descrição mais detalhada do resultado;
- os Agentes envolvidos no Evento e suas funções;
- os objetos envolvidos no Evento e suas funções.

Cada sistema de repositório deve decidir quais Eventos registrar como parte permanente da história de um Objeto. O PREMIS recomenda que as ações que modificam um Objeto sejam sempre registradas e, para isso, fornece um vocabulário controlado dos tipos importantes de Evento para incentivar os repositórios a registrarem esses Eventos de maneira consistente.

Este vocabulário controlado está disponível em: <http://id.loc.gov/vocabulary/preservation/eventType>.

3.3. Agentes

Agentes são os atores que têm funções nos Eventos, nas Declarações de direitos (ver seção 3.4. Direitos) e nos *Objetos Ambiente*. Os agentes podem ser pessoas, organizações, softwares ou hardwares. O PREMIS define apenas um mínimo de unidades semânticas necessárias para identificar os Agentes, uma vez que existem vários padrões externos que podem ser adotados para informações mais detalhadas. Um repositório pode escolher usar um padrão específico para registrar informações adicionais sobre os Agentes ou pode usar o identificador de Agente para apontar às informações registradas externamente.

O Dicionário de Dados inclui:

- um identificador único para o Agente (tipo e valor);
- o nome do Agente;
- a designação do tipo de Agente (pessoa, organização, software, hardware);
- a versão do Agente (em caso de software ou hardware);
- uma anotação geral sobre o Agente;
- os eventos associados ao Agente;
- a Declaração de direitos relacionados ao Agente;
- os *Objetos Ambiente* relacionados ao Agente.

Sempre que um Agente é atribuído a um Evento ou a uma Declaração de direitos, sua função também deve ser registrada. Qualquer Agente pode assumir várias funções. Por exemplo, eu poderia ser o autor e o detentor dos direitos de uma obra, o autor mas não o detentor dos direitos de uma segunda obra, e o depositante de uma terceira. No modelo PREMIS, um repositório atribuiria um identificador único

para mim e o vincularia a qualquer registro de Evento ou Declaração de direitos no qual eu estivesse envolvido, descrevendo também minha função em cada contexto.

3.4. Direitos

A maioria das estratégias de preservação implica a produção de cópias idênticas e versões derivadas de objetos digitais, ações que podem ser limitadas por legislação de proteção aos direitos de autor e de propriedade. A entidade *Direitos* agrega informações sobre direitos e permissões referentes aos objetos em um repositório de preservação, para que este possa executar o que for necessário para preservá-los. Cada Declaração de direitos PREMIS afirma duas coisas: ações que o repositório tem o direito de realizar e a base para a alegação do direito.

Por exemplo, um repositório pode deter uma versão digitalizada de um livro publicado em 1848 que está em domínio público.¹² O repositório pode fazer qualquer coisa com sua versão digital baseado nas permissões legais existentes para obras como essa. Outro repositório tem um objeto copiado de um CD publicado, cuja licença *shrink-wrap* permite fazer cópias para backup, mas restringe o acesso e o uso.

As informações que podem ser registradas em uma Declaração de direitos são:

- um identificador único para a Declaração de direitos (tipo e valor);
- se a base para alegação do direito é copyright, licença, estatuto ou outro (por exemplo, política institucional);
- informações mais detalhadas sobre a situação do copyright, termos de licença ou estatuto, conforme cada caso;
- as ações que a Declaração de direitos permite;
- quaisquer restrições aplicáveis às ações;
- o termo de garantia ou de restrição, ou o período de tempo ao qual a declaração se aplica;
- os Objetos aos quais a declaração se aplica;
- os Agentes envolvidos na Declaração de direitos e suas funções.

A maioria das informações é projetada para ser *acionável*, ou seja, registrada de maneira controlada para poder ser executada por um programa de computador. A Declaração de direitos PREMIS é uma afirmação de direitos, não um registro de informações a partir das quais os direitos podem ser determinados. Isso significa que o PREMIS não define informações detalhadas sobre autores, data e local de publicação e notificações sobre copyright que estão definidas, por exemplo, nas especificações copyrightMD da California Digital Library (<https://cdlib.org/groups/rights-management-group-copyrightmd/>).¹³ O objetivo do esquema copyrightMD é auxiliar os humanos a fazerem determinações de direitos de maneira contínua, enquanto o propósito da entidade Direitos PREMIS é fornecer informações acionáveis para os sistemas de repositório de preservação.

¹² O texto original considera a legislação dos Estados Unidos. Os repositórios devem considerar a legislação local e tratados internacionais sobre direitos autorais para planejar suas ações de preservação.

¹³ Link atualizado. CALIFORNIA DIGITAL LIBRARY. **CopyrightMD**, May 2006. Acesso em: 20 dez. 2020.

4. O DICIONÁRIO DE DADOS

4.1. Amostra de entrada do Dicionário de Dados para uma unidade semântica simples

A tabela I mostra a entrada do Dicionário de Dados para a unidade semântica *size*, que é um componente ou subunidade do contêiner chamado *objectCharacteristics*. *Size* não tem subunidades. A entrada do Dicionário de Dados inclui uma definição do elemento e uma justificativa para incluí-lo entre os metadados PREMIS fundamentais, assim como exemplos e notas sobre como o valor deve ser obtido e usado. Tudo isso é para auxiliar os implementadores a usarem o elemento adequadamente.

Na tabela, as linhas **Categoria do Objeto** e **Aplicabilidade** são usadas juntas para mostrar se a unidade semântica é adequada para descrever Entidades Intelectuais, Representações, Arquivos e/ou Bitstreams. A unidade semântica *size* está apresentada como aplicável apenas para Arquivos e Bitstreams. Por fim, apresenta-se um conjunto de regras para o seu uso: Restrição de dados, Repetibilidade e Obrigação.

A **Restrição de dados** especifica as limitações dos valores que uma unidade semântica pode assumir. No caso de *size*, o valor deve ser um número inteiro. Outra restrição de dados comum é que o valor seja extraído de um vocabulário controlado. Se existir um vocabulário controlado recomendado, ele será mencionado na restrição de dados e alguns termos do vocabulário estarão indicados como exemplos no Dicionário de Dados (na tabela da unidade semântica correspondente). Os vocabulários controlados indicados no PREMIS podem ser consultados em <http://id.loc.gov/preservationdescriptions/>.¹⁴

Outros vocabulários controlados podem ser usados e, nesse caso, o nome do vocabulário deve ser registrado. Não existem unidades semânticas definidas no Dicionário de Dados para o nome dos vocabulários, mas o esquema PREMIS XML proporciona espaço para essa informação e, se forem identificados por um URI,¹⁵ em geral corresponderá ao nome do vocabulário controlado.

A **Repetibilidade** indica se a unidade semântica pode ser usada mais de uma vez.

A **Obrigação** indica se um valor para a unidade semântica é obrigatório ou opcional. Isso pode gerar certa confusão, porque o PREMIS afirma claramente que não exige que um repositório armazene nenhuma informação específica. Uma unidade semântica que é obrigatória não precisa ser registrada e armazenada no repositório. Entretanto, o repositório tem que conseguir gerar o valor da unidade semântica quando necessário, como no intercâmbio com outro repositório. Por exemplo, na improvável situação de um repositório não ter nada além de imagens TIFF 6.0, ele não precisaria registrar a informação sobre o formato de cada um dos objetos. Mesmo assim, o repositório saberia que seus Objetos Arquivo são imagens TIFF 6.0 e poderia fornecer essa informação caso solicitado (ver seção 5.2. Conformidade com o PREMIS). Algumas unidades semânticas são obrigatórias de acordo com o contêiner. Se um contêiner não obrigatório estiver ausente, a unidade semântica obviamente não será usada; por outro lado, se o contêiner não obrigatório estiver presente, a unidade semântica deverá ser usada.

¹⁴ Trecho adaptado da tradução italiana de *Understanding PREMIS* (2017).

¹⁵ Uniform Resource Identifier.

Tabela 1: Trecho do Dicionário de Dados para a unidade semântica size

Unidade semântica	I.5.3 size		
Componentes semânticos	Nenhum		
Definição	Tamanho em bytes do arquivo ou bitstream armazenado no repositório.		
Justificativa	O tamanho é útil para garantir que a quantidade correta do número de bytes do armazenamento foi recuperada e que um sistema tem espaço suficiente para mover ou processar arquivos. Essa informação pode ser usada também para calcular o custo de armazenamento.		
Restrição de dados	Número inteiro		
Categoria de objeto	Entidade Intelectual / Representação	Arquivo	Bitstream
Aplicabilidade	Não aplicável	Aplicável	Aplicável
Exemplos		2038937	2038937
Repetibilidade		Não repetível	Não repetível
Obrigaç�o		Opcional	Opcional
Notas de cria�o/manuten�o	Obtido automaticamente pelo reposit�rio.		
Notas de uso	Definir esta unidade sem�ntica como tamanho em bytes descarta a necessidade de registrar uma unidade de medida. No entanto, para fins de interc�mbio de dados, a unidade de medida deve ser declarada ou entendida por ambas as partes envolvidas.		

4.2. Amostra de entrada do Dicionário de Dados para uma unidade contêiner

A tabela 2 mostra o início da entrada do Dicionário de Dados para objectCharacteristics, a unidade contêiner para size. Você pode afirmar que a unidade é um contêiner porque tem componentes semânticos e a restrição de dados é “contêiner”. Perceba que os componentes semânticos incluídos podem ser unitários, como size, ou outros contêineres, como format.

Tabela 2: Trecho do Dicionário de Dados para a unidade semântica objectCharacteristics

Unidade semântica	1.5 objectCharacteristics		
Componentes semânticos	1.5.1 compositionLevel 1.5.2 fixity 1.5.3 size 1.5.4 format 1.5.5 creatingApplication 1.5.6 inhibitors 1.5.7 objectCharacteristicsExtension		
Definição	Propriedades técnicas de um arquivo ou bitstream que são aplicáveis a todos os formatos ou à maioria deles.		
Justificativa	Existem algumas propriedades técnicas importantes que se aplicam a objetos de qualquer formato. A definição detalhada das propriedades de um formato específico está fora do escopo deste Dicionário de Dados, embora tais propriedades possam ser registradas em objectCharacteristicsExtension.		
Restrição de dados	Contêiner		
Categoria de objeto	Entidade Intelectual / Representação	Arquivo	Bitstream
Aplicabilidade	Não aplicável	Aplicável	Aplicável
Repetibilidade		Repetível	Repetível
Obrigaçã		Obrigatório	Obrigatório
Notas de uso	<p>As unidades semânticas contidas em objectCharacteristics devem ser tratadas como informações pertinentes a um único objeto em uma única unidade compositionLevel. As características do objeto podem ser repetidas quando um objeto foi criado a partir de duas ou mais codificações, como compressão e criptografia. Nesse caso, cada repetição de objectCharacteristics terá uma unidade semântica compositionLevel incrementalmente maior.</p> <p>Quando se aplica criptografia, o bloco objectCharacteristics deve incluir a unidade semântica inhibitors.</p> <p>Um bitstream incorporado a um arquivo pode ter características de objeto diferentes das características do arquivo. Quando essas características forem relevantes para a preservação elas devem ser registradas.</p>		

5. PREMIS EM USO

5.1. PREMIS em XML

Embora não seja uma exigência, existe uma expectativa de que o PREMIS seja representado em XML quando usado para intercâmbio. A Atividade de Manutenção do PREMIS fornece um esquema XML, que corresponde ao Dicionário de Dados, para a descrição direta de Objetos, Eventos, Agentes e Direitos. A figura 4 apresenta uma amostra de metadados PREMIS transcritos usando o esquema PREMIS XML.

Na prática, a maioria dos sistemas de preservação já usa formatos XML para importar e exportar dados. Muitos usam o METS (*Metadata Encoding and Transmission Standard*),¹⁶ outro padrão mantido pela Biblioteca do Congresso dos Estados Unidos, como um contêiner XML para reunir diferentes tipos de metadados. É possível usar o PREMIS dentro do METS, mas isso não é totalmente simples por duas razões. Primeira, o METS desmembra as informações em seções de metadados técnicos, de direitos ou de proveniência. O esquema PREMIS, de acordo com o Dicionário de Dados, tem seções para Objetos, Direitos, Eventos e Agentes. Existe certa correspondência entre as duas estruturas, mas não é perfeita, especialmente para informações sobre Agente. A segunda razão é que o PREMIS e o METS têm algumas sobreposições entre si, por exemplo, cada um deles define uma etiqueta para armazenar somas de verificação. Se os dois padrões são usados juntos, você tem que decidir se registrará os elementos sobrepostos nas seções PREMIS, METS ou em ambas.

Claro que se cada implementação tomasse suas próprias decisões, poderia haver uma grande variação na forma de representação dos dados, impedindo a interoperabilidade. Portanto, os implementadores estabeleceram boas práticas para o uso combinado do PREMIS com o METS. As diretrizes para intercâmbio de dados, usando esses dois padrões, estão disponíveis no site da Atividade de Manutenção do PREMIS: <http://www.loc.gov/standards/premis/guidelines2017-premismets.pdf>

```
<event>
  <eventIdentifier>
    <eventIdentifierType>DAITSS</eventIdentifierType>
    <eventIdentifierValue>10012</eventIdentifierValue>
  </eventIdentifier>
  <eventType>Format Validation</eventType>
  <eventDateTime>2008-05-06T10:40:22-04:00</eventDateTime>
  <eventOutcomeInformation>
    <eventOutcome>Invalid</eventOutcome>
    <eventOutcomeDetail>
      <eventOutcomeDetailNote>ill-formed DateTime value</eventOutcomeDetailNote>
    </eventOutcomeDetail>
  </eventOutcomeInformation>
</event>
```

Figura 4: Trecho do PREMIS em XML

5.2. Conformidade com o PREMIS

O Dicionário de Dados PREMIS tem uma seção sobre o que significa para um repositório estar em conformidade com o PREMIS. A declaração de conformidade, que o PREMIS Editorial Committee

¹⁶ Padrão de Codificação e Transmissão de Metadados.

publicou pela primeira vez em outubro de 2010 e revisou em 2015, especifica princípios de uso, níveis de conformidade e diretrizes para implementação. O documento está disponível em: <http://www.loc.gov/standards/premis/premis-conformance-20150429.pdf>.

Princípios de uso

- 1) Se o repositório implementa (armazena ou exporta) um elemento de dados que se supõe ser uma unidade semântica PREMIS, o elemento de dados deve ter a mesma definição, as mesmas restrições de dados e a mesma aplicabilidade que a unidade semântica definida no PREMIS. Se um elemento de metadados compartilha a definição de uma unidade semântica PREMIS, mas não o nome, o repositório deve estabelecer a equivalência entre o seu elemento de metadados e a unidade semântica PREMIS correspondente.
- 2) Se o repositório implementa uma unidade semântica PREMIS, ele pode optar por adotar repetibilidade e obrigação mais rigorosas, porém não mais liberais do que determinadas no PREMIS. Ou seja, uma unidade semântica repetível pode ser implementada como não repetível, mas não vice-versa, e um elemento obrigatório não pode ser considerado opcional.
- 3) Uma implementação deve incluir as unidades semânticas obrigatórias para qualquer Entidade do Modelo de Dados (Objeto, Evento, Agente ou Direitos) adotada pelo repositório. Note que um componente semântico obrigatório será incluído apenas se o contêiner principal for implementado.
- 4) Se o repositório exporta informações para uso por outro repositório, ele deve fornecer valores para todas as unidades semânticas que são obrigatórias no Dicionário de Dados. Entretanto, existe certa flexibilidade quanto a isso, porque não se exige que os repositórios mantenham unidades semânticas obrigatórias para as Entidades que eles não adotam. Por exemplo, um repositório é livre para adotar ou não a entidade PREMIS Agente, mas caso o faça, a unidade semântica agentIdentifier passa a ser obrigatória em seu sistema. Da mesma maneira, um repositório que não gerencia Objetos Bitstream não tem que fornecer o identificador obrigatório.

Níveis de conformidade

Os níveis apontam três maneiras de avaliar a conformidade de um repositório com o PREMIS:

- **Nível 1.** Ser capaz de mapear os metadados de preservação em relação ao PREMIS;
- **Nível 2.** Ser capaz de exportar os metadados de preservação como PREMIS;
- **Nível 3.** Adotar o PREMIS como esquema interno de maneira que não seja necessário nenhum mapeamento ou conversão adicionais.

Esses níveis estão divididos em categorias: implementação apenas da Entidade Objeto ou implementação da Entidade Objeto em adição a uma ou mais Entidades. Em outras palavras, um repositório deve, ao menos, implementar informações sobre o Objeto, mas não precisa adotar todos os tipos de Entidade definidas no Modelo de Dados PREMIS. Também não é exigido armazenar os metadados internamente usando os nomes das unidades semânticas PREMIS ou usando valores que sigam as restrições de dados PREMIS. Ou seja, não importa como um repositório “reconhece” um valor PREMIS – seja armazenando com o mesmo nome ou nome diferente, mapeando a partir de outro valor, direcionando para um registro externo, inferindo, padronizando, seja usando quaisquer outros meios. O repositório estará em conformidade contanto que forneça um bom valor PREMIS quando necessário.

Por outro lado, quanto mais unidades semânticas um repositório adotar, mais ele será valorizado por usar PREMIS. O Dicionário de Dados PREMIS foi desenvolvido para identificar as informações mais importantes que a maioria dos repositórios precisará para preservar conteúdo digital a longo prazo. Um repositório de preservação responsável deve analisar cuidadosamente o PREMIS e apresentar uma boa justificativa para a não implementação de qualquer parte do Dicionário de Dados.

6. PARA MAIS INFORMAÇÕES

O site da Atividade de Manutenção do PREMIS (www.loc.gov/standards/premis/) tem informações para todos, incluindo links para o grupo de implementadores (*PREMIS Implementers Group - PIG*), informações sobre o evento *PREMIS Implementation Fair*, tutoriais, esquemas, ferramentas e notícias. No site existe também uma seção de recursos (*PREMIS Resources*) que apresenta bibliografia sobre o PREMIS e temas relacionados (www.loc.gov/standards/premis/bibliography.html). Alguns dos recursos mais úteis para os interessados em geral estão listados a seguir.

Sobre metadados de preservação em geral:

“Preservation Metadata” 2nd edition (PDF:986KB/36pp.)

Brian Lavoie (OCLC) and Richard Gartner (Oxford)

Published by the Digital Preservation Coalition as DPC Technology Watch Report No. 13-03:

May 2013.

dx.doi.org/10.7207/twr13-03

Sobre implementação do PREMIS:

Digital Preservation Metadata for Practitioners: Implementing PREMIS.

Angela Dappert, Rebecca Squire Guenther, Sébastien Peyrard, Editors. Springer, 2016.

dx.doi.org/10.1007/978-3-319-43763-7

Sobre mudanças na versão 3.0 (Webinar):

“Digital Preservation Metadata and Improvements to PREMIS in Version 3.0: A DCMI/ASIST Joint Webinar Presented by Angela Dappert” (Wednesday, May 27, 2015)

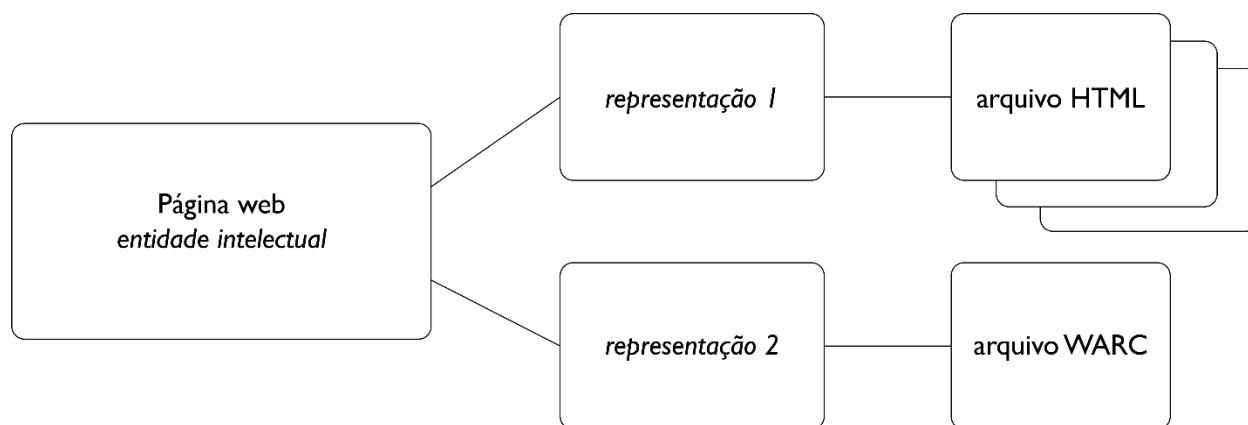
www.loc.gov/standards/premis/v3/tutorial.html

Apêndice A: Exemplos

A.1. Exemplo de Objeto

Este exemplo descreve como um site poderia ser modelado em um sistema de preservação hipotético. Ele apresenta um diagrama de alto nível, alguns detalhes sobre como as unidades semânticas podem ser preenchidas e, por fim, trechos em XML para mostrar como as unidades semânticas poderiam ser implementadas no repositório. Os exemplos não mostram todas as unidades semânticas que são relevantes para cada nível do Objeto.

Neste exemplo, o site se chama *PREMIS Website*, essa é a Entidade Intelectual que deve ser preservada. Existem duas Representações desse site, uma é a matriz de preservação e a outra a de acesso.



Entidade Intelectual

Neste exemplo, o nível Entidade Intelectual estabelece uma propriedade significativa do site. Isso determina o comportamento de suas duas Representações. É bem provável que quaisquer propriedades significativas seriam muito mais granulares, mas o exemplo serve para demonstrar como elas podem ser usadas no nível Entidade Intelectual.

Unidade Semântica PREMIS	Entidade Intelectual I
I.1 objectIdentifier	
I.1.1 objectIdentifierType	IEPID
I.1.2 objectIdentifierValue	17415491
I.4 significantProperties	
I.4.1 significantPropertiesType	behavior
I.4.2 significantPropertiesValue	External links open in new browser window

Representação

O nível de preservação descreve a política institucional para este Objeto e ele pode ser determinado nos níveis Entidade Intelectual, Representação ou Arquivo. Neste exemplo, ele foi definido no nível Representação, conforme a decisão da instituição de que cada Representação seja gerenciada de maneira distinta. A instituição deste exemplo decidiu que a Representação composta pelo arquivo WARC (Representação 2) é a matriz de preservação, a que será preservada permanentemente e a partir da qual serão geradas as matrizes de acesso. A Representação 1 é a matriz de acesso,

disponibilizada para a consulta do público. Sua importância não é permanente, ela será mantida enquanto o mecanismo de acesso corrente estiver em operação.

À **Representação 1** foi atribuído o valor *bit preservation* (preservação de bit) para o nível de preservação. De acordo com a política institucional isso exige apenas que sejam feitas cópias de backup, sem necessidade de análises de risco ou procedimentos de emulação ou migração.

Por outro lado, a **Representação 2** foi classificada com o nível de *logical preservation* (preservação funcional) com desempenho máximo. Segundo a instituição, essa Representação será submetida a análises de risco relevante e, quando necessário, submetida à emulação ou migração para garantir sua preservação permanente.

Unidade Semântica PREMIS	Representação 1	Representação 2
I.1 objectIdentifier		
I.1.1 objectIdentifierType	REPPID	REPPID
I.1.2 objectIdentifierValue	17415492	17415495
I.3 preservationLevel		
I.3.1 preservationLevelType	bit preservation	logical preservation
I.3.2 preservationLevelValue	low	full
I.3.3 preservationLevelRole	capability	intention
I.3.4 preservationLevelRationale	institutional policy	institutional policy
I.3.5 preservationLevelDateAssigned	2015-02-23	2015-02-23
I.13 relationship		
I.13.1 relationshipType	derivation	derivation
I.13.2 relationshipSubType	has source	is source of
I.13.3 relatedObjectIdentifier		
I.13.3.1 relatedObjectIdentifierType	REPPID	REPPID
I.13.3.2 relatedObjectIdentifierValue	17415495	17415492

Um trecho de *preservationLevel* descrito em XML:

```
<premis:preservationLevel>
  <premis:preservationLevelType>logical preservation</premis:preservationLevelType>
  <premis:preservationLevelValue>full</premis:preservationLevelValue>
  <premis:preservationLevelRole authority="preservationLevelRole"
    authorityURI="http://id.loc.gov/vocabulary/preservation/preservationLevelRole"
    valueURI="http://id.loc.gov/vocabulary/preservation/preservationLevelRole/int">intention
  </premis:preservationLevelRole>
  <premis:preservationLevelRationale>institutional policy</premis:preservationLevelRationale>
  <premis:preservationLevelDateAssigned>2015-02-23</premis:preservationLevelDateAssigned>
</premis:preservationLevel>
```

Arquivos

A Representação 1 inclui vários arquivos distintos, cada um é descrito como objeto Arquivo com seu próprio conjunto de elementos de metadados.

Unidade Semântica PREMIS	Arquivo 1	Arquivo 2
1.1 objectIdentifier		
1.1.1 objectIdentifierType	FILEPID	FILEPID
1.1.2 objectIdentifierValue	17415493	17415494
1.5 objectCharacteristics		
1.5.1 compositionLevel	0	0
1.5.2 fixity		
1.5.2.1 messageDigestAlgorithm	SHA256	SHA256
1.5.2.2 messageDigest	d2bed92b73c7090bb30a0b30016882e7069c437488e1513e9deaacbe29d38d92	074862dbfa0806ef5a26c3ff748d394e79728e9c957ff8c198ae13214c9cfec0
1.5.2.3 messageDigestOriginator	NRI	NRI
1.5.3 size	4859385	123451
1.5.4 format		
1.5.4.1 formatDesignation		
1.5.4.1.1 formatName	Extensible Hypertext Markup Language	JPEG File Interchange Format
1.5.4.1.2 formatVersion	1	1.02
1.5.4.2 formatRegistry		
1.5.4.2.1 formatRegistryName	PRONOM	PRONOM
1.5.4.2.2 formatRegistryKey	fmt/102	fmt/44
1.5.4.2.3 formatRegistryRole	specification	specification
1.13 relationship		
1.13.1 relationshipType	structural	structural
1.13.2 relationshipSubType	Is included in	Is included in
1.13.3 relatedObjectIdentifier		
1.13.3.1 relatedObjectIdentifierType	REPPID	REPID
1.13.3.2 relatedObjectIdentifierValue	17415492	17415492

Um trecho de objectCharacteristics do Arquivo 1 descrito em XML:

```

<premis:objectCharacteristics>
  <premis:compositionLevel>0</premis:compositionLevel>
  <premis:fixity>
    <premis:messageDigestAlgorithm>SHA256</premis:messageDigestAlgorithm>
    <premis:messageDigest>d2bed92b73c7090bb30a0b30016882e7069c437488e1513e9deaacbe29d38d92
      </premis:messageDigest>
    <premis:messageDigestOriginator>NRI</premis:messageDigestOriginator>
  </premis:fixity>
  <premis:size>4859385</premis:size>
  <premis:format>
    <premis:formatDesignation>
      <premis:formatName>Extensible Hypertext Markup Language</premis:formatName>
      <premis:formatVersion>1</premis:formatVersion>
    </premis:formatDesignation>
    <premis:formatRegistry>
      <premis:formatRegistryName>PRONOM</premis:formatRegistryName>
      <premis:formatRegistryKey>fmt/102</premis:formatRegistryKey>
      <premis:formatRegistryRole>specification</premis:formatRegistryRole>
    </premis:formatRegistry>
  </premis:format>
</premis:objectCharacteristics>

```

A.2. Exemplo de Evento

Este exemplo descreve um evento executado pelo repositório em um objeto. Uma função comum de um repositório é identificar o formato de arquivo de um objeto e validar se ele atende aos padrões correspondentes ao formato, observando conformidade ou exceções. O exemplo demonstra a validação de formato do Arquivo I do exemplo anterior.

Unidade Semântica PREMIS	Arquivo I
2.1 eventIdentifier	
2.1.1 eventIdentifierType	DPS
2.1.2 eventIdentifierValue	25
2.2 eventType	validation
2.3 eventDateTime	2013-07-01T17:23:25Z
2.4 eventDetailInformation	
2.4.1 eventDetail	Format identification performed on file
2.5 eventOutcomeInformation	
2.5.1 eventOutcome	success
2.5.2 eventOutcomeDetail	
2.5.2.1 eventOutcomeDetailNote	FORMAT_ID=fmt/102;IDENTIFICATION_METHOD=SIGNATURE;FILE_EXTENSION=html;DEPOSIT_ACTIVITY_ID=659202;PID=FL26274026;SIP_ID=444381;PRODUCER_ID=29408010705;TASK_ID=48;PROCESS_ID=;MF_ID=8
2.6 linkingAgentIdentifier	
2.6.1 linkingAgentIdentifierType	local
2.6.2 linkingAgentIdentifierValue	DROID6_63_1
2.6.3 linkingAgentIdentifierRole	implementer
2.7 linkingObjectIdentifier	
2.7.1 linkingObjectIdentifierType	FILEPID
2.7.2 linkingObjectIdentifierValue	17415493

Um trecho do Evento de identificação de formato descrito em XML:

```
<premis:event>
  <premis:eventIdentifier>
    <premis:eventIdentifierType>DPS</premis:eventIdentifierType>
    <premis:eventIdentifierValue>25</premis:eventIdentifierValue>
  </premis:eventIdentifier>
  <premis:eventType authority="event type"
    authorityURI="http://id.loc.gov/vocabulary/preservation/eventType"
    valueURI="http://id.loc.gov/vocabulary/preservation/eventType/val"
    validation</premis:eventType>
  <premis:eventDateTime>2013-07-01T17:23:25Z</premis:eventDateTime>
  <premis:eventDetailInformation>
    <premis:eventDetail>Format identification performed on file</premis:eventDetail>
  </premis:eventDetailInformation>
  <premis:eventOutcomeInformation>
    <premis:eventOutcome>success</premis:eventOutcome>
    <premis:eventOutcomeDetail>
      <premis:eventOutcomeDetailNote>
        FORMAT_ID=fmt/102;IDENTIFICATION_METHOD=SIGNATURE;FILE_EXTENSION=html;
        DEPOSIT_ACTIVITY_ID=659202;PID=FL26274026;SIP_ID=444381;
        PRODUCER_ID=29408010705;TASK_ID=48;PROCESS_ID=;MF_ID=8
      </premis:eventOutcomeDetailNote>
    </premis:eventOutcomeDetail>
  </premis:eventOutcomeInformation>
</premis:event>
```



```

    </premis:eventOutcomeDetail>
  </premis:eventOutcomeInformation>
  <premis:linkingAgentIdentifier>
    <premis:linkingAgentIdentifierType>LOCAL</premis:linkingAgentIdentifierType>
    <premis:linkingAgentIdentifierValue>DROID6_63_1</premis:linkingAgentIdentifierValue>
    <premis:linkingAgentRole authority="eventRelatedAgentRole"
      authorityURI="http://id.loc.gov/vocabulary/preservation/eventRelatedAgentRole"
      valueURI="http://id.loc.gov/vocabulary/preservation/eventRelatedAgentRole/imp">
      implementer</premis:linkingAgentRole>
  </premis:linkingAgentIdentifier>
  <premis:linkingObjectIdentifier>
    <premis:linkingObjectIdentifierType>FILEPID</premis:linkingObjectIdentifierType>
    <premis:linkingObjectIdentifierValue>17415493</premis:linkingObjectIdentifierValue>
  </premis:linkingObjectIdentifier>
</premis:event>

```

A.3. Exemplo de Agente

Este exemplo descreve o agente associado ao evento apresentado no tópico anterior. Neste caso, o Agente é o software que executou a ação no objeto.

Unidade Semântica PREMIS	
3.1 agentIdentifier	
3.1.1 agentIdentifierType	local
3.1.2 agentIdentifierValue	DROID6_63_1
3.2 agentName	REG_SA_DROID
3.3 agentType	software
3.4 agentVersion	Version 6.01
3.5 agentNote	Signature version Binary SF v.63/ Container SF v.1

Um trecho das informações sobre o Agente descrito em XML:

```

<premis:agent>
  <premis:agentIdentifier>
    <premis:agentIdentifierType>local</premis:agentIdentifierType>
    <premis:agentIdentifierValue>DROID6_63_1</premis:agentIdentifierValue>
  </premis:agentIdentifier>
  <premis:agentName>REG_SA_DROID</premis:agentName>
  <premis:agentType>software</premis:agentType>
  <premis:agentVersion>Version 6.01</premis:agentVersion>
  <premis:agentNote>Signature version Binary SF v.63/ Container SF v.1</premis:agentNote>
</premis:agent>

```

Apêndice B: Glossário

Este glossário reúne as definições presentes no texto deste guia. As definições podem ser menos formais do que as que constam no *Dicionário de Dados PREMIS para Metadados de Preservação*.

Acionável (Actionable): Qualidade de ser registrado de maneira controlada para poder ser acionado por um programa de computador.

Agente (Agent): Pessoa, organização ou programa de computador que desempenha papel relacionado a um Evento ou a uma Declaração de direitos.

Ambiente (Environment): Hardware, software e outros objetos necessários para renderizar um Objeto.

Contêiner de extensão (Extension container): Tipo especial de unidade contêiner PREMIS que não tem subunidades definidas, mas que foi projetado para permitir o registro de metadados não PREMIS.

Emulação (Emulation): Estratégia de preservação que envolve reproduzir um ambiente de renderização antigo em um hardware e/ou software mais novos.

Entidade Direitos (Rights entity): Entidade PREMIS que reúne informações sobre direitos e permissões referentes aos Objetos em um repositório de preservação.

Entidade Evento (Event Entity): Entidade PREMIS que reúne informações sobre as ações que afetam os Objetos em um repositório de preservação.

Entidade Intelectual (Intellectual Entity): Conjunto de conteúdos tratado como uma unidade para fins de gerenciamento e descrição; semelhante ao tratamento de uma “entidade bibliográfica” pela Biblioteconomia.

Estratégias de preservação (Preservation strategies): Técnicas aplicadas para garantir que recursos digitais permaneçam utilizáveis a longo prazo; duas estratégias usuais são migração e emulação.

Inibidores (Inhibitors): Atributos de um objeto digital que têm por objetivo restringir acesso, uso ou migração.

Metadados de preservação (Preservation metadata): Metadados que dão suporte às atividades destinadas a garantir a usabilidade de recursos digitais a longo prazo.

Migração (Migration): Estratégia de preservação que envolve produzir uma versão de um arquivo digital em um formato mais novo.

Objeto Ambiente (Environment Object): Objeto que é parte de um conjunto técnico de software, hardware e outros elementos necessários para interpretar corretamente as Representações, os Arquivos e os Bitstreams.

Objeto Arquivo (File Object): Tipo de Objeto PREMIS; um arquivo de computador como um PDF ou JPEG.

Objeto Bitstream (Bitstream Object): Tipo de Objeto PREMIS; dados internos de um arquivo com propriedades comuns para fins de preservação e que não são independentes.

Objeto Representação (Representation Object): Tipo de Objeto PREMIS; o conjunto de todos os Objetos Arquivo necessários para renderizar uma Entidade Intelectual.

Objetos (Objects): Itens digitais que são de fato armazenados e gerenciados em um repositório de preservação. O PREMIS define quatro tipos de Objetos: Arquivos, Bitstreams, Representações e Entidade Intelectuais.

Propriedades significativas (Significant properties): Características de um Objeto que devem ser mantidas pelas ações de preservação.

Proveniência digital (Digital provenance): Documentação da cadeia de custódia e do histórico de modificações de um recurso digital.

Unidades contêiner (Container units): Unidades semânticas que não têm valor próprio, mas existem para reunir subunidades relacionadas.

Unidades semânticas (Semantic units): Fragmentos de informação ou conhecimento.