

Sandra Regina de Alencar Silva

Proposta de Perfil de Aplicação  
Dublin Core para  
publicações em quadrinhos

São Paulo  
2011

DEPARTAMENTO DE BIBLIOTECONOMIA E DOCUMENTAÇÃO  
ESCOLA DE COMUNICAÇÕES E ARTES  
UNIVERSIDADE DE SÃO PAULO

Sandra Regina de Alencar Silva

Proposta de Perfil de Aplicação Dublin Core  
para publicações em quadrinhos

Monografia apresentada para  
obtenção do grau de Bacharel  
em Biblioteconomia pela  
Universidade de São Paulo.

Orientador: Professor Dr.  
Marcos Luiz Mucheroni

São Paulo  
2011

## TERMO DE APROVAÇÃO

Nome do Autor: Sandra Regina de Alencar Silva

Título do TCC: Proposta de Perfil de Aplicação Dublin Core para publicações em quadrinhos

Presidente da Banca: Prof. Dr. Marcos Mucheroni

Banca Examinadora:

Prof. Dr. Marcos Mucheroni

Prof. Dr. Fernando Modesto

Prof. Dr. Paulo Ramos

Instituição: Universidade de São Paulo

Instituição: Universidade de São Paulo

Instituição: Universidade Federal de São Paulo

Aprovada em: 08/12/2011

*Para meus pais.*

Agradeço,

Ao professor Marcos Luiz Mucheroni, pela orientação presente e pela motivação.

Ao professor Waldomiro Vergueiro, pela oportunidade de estágio no acervo do Observatório de Histórias em Quadrinhos, cuja experiência proporcionada foi essencial no desenvolvimento deste trabalho.

Ao colega Luciano Ramalho que, sem se dar conta, sugeriu o tema deste trabalho.

*Mudam-se os meios, sofisticam-se os instrumentos e surgem nomes novos para designar coisas velhas. Entretanto, a essência das coisas permanece. Os tempos digitais acenam para a concretização de sonhos: da grande biblioteca universal, do conhecimento sem fronteiras, da extensão infinita dos limites da memória, do processamento eletrônico de dados complexos, coisas que embora profetizadas, há anos, décadas, séculos, somente agora encontram as condições para se tornarem realidade.*

Lídia Alvarenga

*O profissional da informação deve mudar e adaptar-se aos novos ambientes de informação eletrônica, ele ou ela deve aprender sobre novas tecnologias e conscientizar-se de suas forças e fraquezas. Bibliotecários não deveriam sentir-se ameaçados pelos computadores e desenvolvimentos tecnológicos, mas ir em direção às novas tecnologias e assumir um papel fundamental dentro das organizações.*

J. Medhurst

## **RESUMO**

Este trabalho revisita conceitos relacionados a metadados e interoperabilidade e apresenta um estudo do padrão de metadados Dublin Core.

Analisa características das publicações em quadrinhos e traça um modelo de registro bibliográfico para essas publicações, por meio da elaboração de cada um dos componentes especificados no Framework Singapura para Perfis de Aplicação Dublin Core.

Sugere-se implementação de um Sistema de Informação Federado, utilizando-se da estrutura de coleta de dados disponibilizada pela Iniciativa de Arquivos Abertos – OAI.

Palavras-chave: Publicações em quadrinhos; Perfis de Aplicação; Dublin Core; Open Archives Initiative

## **ABSTRACT**

This monography revisits concepts related to metadata and interoperability and presents a study of the Dublin Core metadata standard.

It analyzes characteristics of comics outlining a model for a bibliographic record for this kind of publication through the elaboration of each of the components specified in the Singapore Framework for Dublin Core Application Profiles.

An implementation of a Federated Information System is suggested, using the structure for metadata harvesting provided by the Open Archives Initiative.

Keywords: Comics; Application Profiles; Dublin Core; Open Archives Initiative

## Índice de Figuras

Figura 1 – Obtenção de perfis de aplicação por meio de derivação e combinação de elementos de dados de padrões de metadados pré-existentes.....	23
Figura 2 – A declaração de dois <i>namespaces</i> distintos é expressa no documento XML representado nesta figura, que também fornece outras informações sobre a sintaxe da linguagem de marcação.....	26
Figura 3 – Diagrama das possíveis relações “tem alcance” e “tem domínio” de uma propriedade.....	41
Figura 4 – Modelo de Recurso DCMI, delineado em Powell (2007a) .....	47
Figura 5 – Modelo de Conjunto Descritivo DCMI em Powell (2007a).....	49
Figura 6 – Modelo de Vocabulário DCMI em Powell (2007a).....	51
Figura 7 – Framework Singapura delineado por Nilsson, Baker e Johnston (2008).....	54
Figura 8 – Relações entre esquemas de sistemas federados de bases de dados “com esquemas faltando em alguns níveis” (destaque nosso) propostos de Sheth e Larson (1990).....	62
Figura 9 – Plano de funcionamento do protocolo OAI-PMH .....	65
Figura 10 – Quadro comparativo dos protocolos Z39.50 e OAI-PMH, adaptado de Marcondes e Sayão (2002).....	67
Figura 11 – Ficha catalográfica do acervo da Biblioteca da Universidade Estadual de Michigan.....	69
Figura 12 – Registro MARC de revista em quadrinho da coleção da Universidade Estadual Bowling Green.....	70
Figura 13 – Registro de publicação seriada em ComicBookDC .....	73
Figura 14 – Registro de item de série em .....	73
Figura 15 – Cartão para inventario das edições de um título de revista em quadrinhos.....	74
Figura 16 – Modelo de Domínio PAPuQ.....	84
Figura 17 – Interface de usuário de registro pertencente ao catálogo <i>online</i> da Biblioteca Nacional .....	101
Figura 18 – Registro pertencente ao catálogo <i>online</i> da Biblioteca Nacional em formato MARC .....	102
Figura 19 – Plano de transformações de esquemas de metadados da implementação de base de dados proposta .....	103
Figura 20 – Campos do esquema de metadados Pré-PAPuQ.....	114



## Índice de Quadros

Quadro 1 – Exemplos de coincidências em nomes e definição de elementos nos padrões Dublin Core e IEEE LOM.....	27
Quadro 2 – Relação dos elementos cerne do padrão Dublin Core .....	30
Quadro 3 – Termos do padrão Dublin Core Qualificado mantidos pelo DCMI .....	34
Quadro 4 – Termos do Vocabulário de Tipos do DCMI ( <i>DCMIType</i> ).....	38
Quadro 5 – <i>Namespaces</i> sob responsabilidade da Iniciativa de Metadados Dublin Core (POWELL et al., 2007b) .....	43
Quadro 6 – Semântica das entidades do Modelo de Domínio PAPuQ.....	85
Quadro 7 – Semântica das propriedades PAPuQ e equivalência com elementos Dublin Core.....	88
Quadro 8 – Equivalência entre propriedades PAPuQ e elementos Dublin Core com suas restrições estruturais.....	97
Quadro 9 – Mapeamento de registro em formato MARC para PAPuQ .....	101
Quadro 10 – Sucessivos mapeamentos de metadados para padronização dos valores de registros .....	103
Quadro 11 – Semântica das propriedades do esquema de metadados Pré-PAPuQ.....	106
Quadro 12 – Equivalência entre designações de responsabilidade do Pré-PAPuQ e do <i>MARC Code List for Relators</i> .....	108
Quadro 13 – Sintaxe para formas de contato Pré-PAPuQ .....	113

## Sumário

Introdução.....	2
Objetivo geral .....	6
Objetivos específicos .....	6
Metodologia.....	6
1. Metadados e Interoperabilidade.....	7
1.1 Metadados.....	7
1.1.1 Esquemas de metadados .....	11
1.1.2 Sintaxe e semântica de metadados .....	14
1.2 Padrões e interoperabilidade .....	17
1.2.1 Perfis de aplicação .....	21
2. Estudo do padrão Dublin Core .....	29
2.1 Origens do Padrão Dublin Core .....	29
2.2 Iniciativa de Metadados Dublin Core.....	31
2.3 Termos Dublin Core.....	32
2.4 Atributos de termos Dublin Core .....	42
2.5 Namespaces Dublin Core .....	43
2.6 Modelo Abstrato Dublin Core .....	43
2.6.1 Modelo de Recurso.....	47
2.6.2 Modelo de Conjunto Descritivo .....	48
2.6.3 Modelo de Vocabulário .....	50
2.7 O Perfil de Aplicação Dublin Core .....	52
2.7.1 Framework Singapura.....	53
2.8 Níveis de interoperabilidade do Dublin Core .....	57
2.8.1 Nível 1: Compartilhamento de definição de termos.....	58
2.8.2 Nível 2: Interoperabilidade semântica formal.....	58
2.8.3 Nível 3: Conjunto Descritivo de interoperabilidade sintática .....	59
2.8.4 Nível 4: Interoperabilidade de Perfil de Conjunto Descritivo .....	59
3. Proposta de um Perfil de Aplicação para descrição de publicações em quadrinhos .....	60
3.1 Sistemas de Informação Federados.....	61
3.2 A Iniciativa de Arquivos Abertos e o Protocolo OAI-PMH.....	64
3.3 Catalogação de publicações em quadrinhos .....	68
3.3.1 Publicações em quadrinhos: características .....	75
3.4 PAPuQ: Perfil de Aplicação para Publicações em Quadrinhos.....	78
3.4.1 Requerimentos Funcionais .....	79
3.4.2 Modelo de Domínio .....	83
3.4.3 Perfil do Conjunto Descritivo .....	85
3.4.4 Diretrizes de uso.....	104
3.4.5 Orientação para Sintaxe e Formatos de Dados.....	115
Conclusão .....	117
Anexos.....	124

## Introdução

A difusão do conhecimento, a necessidade de seleção por relevância, interesse e qualidade dos documentos na Web fez emergir as necessidades de descrição semântica de conteúdos, os quais só podem ser feitos por meio de um trabalho colaborativo.

A tendência de cooperação e concentração de esforços entre iniciativas relacionadas a metadados foi exposta na quinta edição do *Metadata Watch Report* (METADATA WATCH REPORT, 2001), uma publicação trimestral do projeto SCHEMAS, que tem por objetivo prover um panorama mundial do progresso realizado em criação e desenvolvimento e aplicação de metadados.

O relatório, do ano de 2001, explana sobre uma mudança gradual na forma como essas iniciativas trabalham: para seus observadores, esforços de padronização que tendiam a se basear em requerimentos bastante específicos visando atender às necessidades de um domínio ou grupo de organizações cooperantes, e provendo serviços a uma audiência específica, passaram a enxergar que outras comunidades com objetivos semelhantes também poderiam ser beneficiadas por suas soluções.

Quando fronteiras se esvaecem, desenvolvedores de soluções em metadados estão percebendo que outros também estão ocupados desenhando soluções em metadados em domínios que se sobrepõem aos deles. (METADATA WATCH REPORT, 2001)

Não obstante museus, bibliotecas, arquivos, editoras, agências de informação governamental, de dados estatísticos e dados geoespaciais etc. lidem com muito dos mesmos dados, comumente os fazem de forma autônoma, não relacionada.

Dessa maneira, caracteriza-se uma competição entre padrões em vigência, que, sendo adotados de forma dispersa por comunidades afins, configuram as chamadas “ilhas de interoperabilidade”, onde a troca de informação<sup>1</sup> entre os utilizadores de um padrão é

---

<sup>1</sup> É com a mesma ressalva de Araújo (1995) de que, na área de Comunicação e Ciência da Informação, existe uma confusão terminológica entre o objeto trabalhado (dados, documentos, textos, mensagens) e o possível efeito sobre seu usuário (a informação em si), que o termo informação é usado. Essa ausência de rigor terminológico pode ser percebida em expressões como "tratamento da informação", "sistema de informação" e

relativamente fácil, mas mais ou menos complexa entre os utilizadores de outros padrões - quando essa troca não é impossibilitada.

Tal situação claramente contraria o conceito de interoperabilidade e a observação de que desenvolvedores de esquemas e padrões metadados estão alterando suas posturas é tida como uma boa notícia.

Eles estão começando a perceber que se trata de um desperdício de (escassos) recursos humanos duplicar esforços para resolver os mesmos problemas e especialmente que é improdutivo tentar ganhar a competição e se tornar O padrão de metadados. (METADATA WATCH REPORT, 2001)

Uma dessas iniciativas de metadados que tem obtido sucesso é a Iniciativa de Metadados Dublin Core, ou DCMI (*Dublin Core Metadata Initiative*), cuja inicial e principal contribuição refere-se ao estabelecimento de quinze propriedades para descrição de recursos informacionais de qualquer natureza.

Naturalmente, para corresponder a tamanha abrangência descritiva, esse conjunto de propriedades caracteriza-se pela generalidade, que em muitos casos não responde bem à função de identificar um recurso informacional. Assim, não é isenta de críticas a conformação desse conjunto de elementos – e mesmo a possibilidade de prover a interoperabilidade entre recursos informacionais de qualquer natureza pode parecer presunçosa, senão de discutível funcionalidade em termos práticos, quando a busca por informação especializada torna-se imperativa na sociedade.

No entanto, são inquestionáveis os benefícios que mesmo um grau genérico de padronização descritiva oferece em ambientes saturados pela produção desmedida de informação, cujo exemplo mais evidente é o ambiente Web. A adoção de práticas e ferramentas uniformizantes de caracterização de produtos informacionais, promovida pelo uso de metadados, mostra-se, portanto, bem-vinda, embora a satisfação da necessidade por informação especializada continue sendo um ponto nevrálgico que os sistemas de informação devem suprir.

---

"recuperação da informação", porém nesses casos o termo é sustentado a fim de manter correspondência com a literatura da área. Em outras ocasiões, no entanto, busca-se minimizar o problema com a qualificação do termo, como por exemplo utilizando as expressões "recurso informacional" ou "fonte de informação".

No âmbito da Iniciativa de Metadados Dublin Core, essa questão foi considerada na proposição de uma estrutura que provesse uma forma de organizar dados equilibrando necessidades locais (especializadas) com padrões de escopo mais genérico, garantindo a interoperabilidade entre diferentes aplicações desses dados. O *Framework* Singapura, como é chamada essa estrutura, corresponde a uma metodologia para criação de Perfis de Aplicação Dublin Core, ou DCAPs (*Dublin Core Application Profiles*).

Neste trabalho, a revisão de literatura no que concerne a metadados, interoperabilidade e afins, de uma forma mais genérica, e o estudo da documentação do padrão Dublin Core, especialmente o que se refere ao *Framework* Singapura, visam obter informações sobre funções e características de seus perfis de aplicação e sobre a maneira pela qual podem beneficiar Sistemas de Informação, considerando o papel do profissional da informação no contexto do desenvolvimento e uso de esquemas de metadados.

Esse Perfil de Aplicação para Publicações em Quadrinhos, batizado PAPuQ, faz uso dos quinze elementos cerne do padrão Dublin Core, e tenciona realizar a descrição bibliográfica desse tipo de publicação, compondo, assim, um modelo de catálogo bibliográfico. Os registros de metadados que compusessem instâncias desse perfil, de acordo com seu objetivo, permitiriam identificar exemplares, perfazendo, assim, uma espécie de controle bibliográfico, uma vez que, “em sentido amplo, o controle bibliográfico pressupõe um domínio completo sobre os materiais que registram o conhecimento objetivando sua identificação, localização e obtenção. (CAMPELLO; MAGALHÃES, 1997:7)

Neste trabalho sugerimos, ainda, uma forma de implementação desse Perfil de Aplicação, que pelo uso do protocolo para coleta de metadados desenvolvido pela Iniciativa de Arquivos Abertos, ou OAI (*Open Archives Initiative*), possibilitaria agregar registros de exemplares de publicações em quadrinhos, provenientes de bases de dados – disponibilizadas em rede – de indivíduos ou organizações, em um repositório único.

A forma relativamente simples pela qual se dá tal agregação é possibilitada por acontecer em ambiente de rede, na Internet, como já adotada por agências nacionais:

Aliada ao fenômeno da globalização, que dilui as fronteiras comerciais dos países, uma rede mundial de informações como a Internet tem potencial para modificar a forma como o controle bibliográfico vem sendo operacionalizado até o momento. Essas modificações dizem respeito, principalmente, às formas de acesso às informações bibliográficas. (CAMPELLO; MAGALHÃES, 1997:9)

A viabilidade de gerar bases de dados a partir do reuso de metadados disponíveis na Internet é ressaltada por Marcondes e Sayão, tendo em vista o alto do tratamento técnico de catalogação e indexação que produto similar despenderia em modelo tradicional de formação de bases de dados – ou seja, centralizado – trazendo à tona ainda que o “software para tornar provedores de dados e de serviços compatíveis com o OAI PMH está disponível gratuitamente na Internet, no site do Open Archives Initiative” (MARCONDES, SAYÃO, 2002:52).

No mundo dos apreciadores de quadrinhos, a tendência cooperação e concentração de esforços para reunir informações sobre esses produtos editoriais são evidenciadas pela existência de bases de dados na Web, construídas em conjunto, que em graus variáveis de aprofundamento catalográfico, discorrem sobre os mais diversos aspectos dessas publicações, desde sua descrição física até a inserção de resumos e resenhas, sem falar em biografias de personagens e de seus criadores.

Embora a caracterização detalhada de uma obra apresente muitas vantagens, ela também apresenta custos elevados, se não afetando diretamente as finanças de uma organização – como no caso dessas bases de dados baseadas na cooperação e no voluntariado –, em termos de tempo e produtividade.

O Perfil de Aplicação proposto neste trabalho, em conjunto com sua implementação sugerida, não visa se colocar como alternativa a esse modelo de trabalho voluntário, tão bem estabelecido pelos consumidores de quadrinhos, mas servir-lhe de complemento, pela caracterização concisa e objetiva de exemplares já descritos ou disponíveis em algum outro lugar, indicando essa localização, onde se poderia obter maiores informações sobre o item descrito como instância do PAPuQ.

## **Objetivo geral**

Propor um Perfil de Aplicação para publicações em quadrinhos conforme especificações do *Framework Singapura* da Iniciativa de Metadados Dublin Core, ressaltando o papel de práticas biblioteconômicas no processo de elaboração dessa ferramenta para descrição de recursos em meio digital.

## **Objetivos específicos**

Esquematizar a aplicação dos elementos do padrão Dublin Core Simple na descrição de publicações em quadrinhos, visando favorecer a interoperabilidade dos registros bibliográficos por meio do uso de metadados amplamente adotados, em diversos domínios do conhecimento, e sugerir a implementação de um Sistema de Informação Federado, utilizando-se do serviço de coleta de metadados da Iniciativa de Arquivos Abertos – OAI.

## **Metodologia**

Revisão de literatura e da documentação do padrão de metadados Dublin Core e do protocolo OAI-PMH para coleta de metadados em bases de dados distribuídas, a fim de apoiar a proposta do Perfil de Aplicação que é objeto deste trabalho; e a análise empírica de características bibliográficas de publicações em quadrinhos, para embasamento de um modelo para sua catalogação.

# 1. Metadados e Interoperabilidade

## 1.1 Metadados

Um relatório técnico sobre linguagens de programação, datado de 1968, parece ser a referência mais antiga de uso do termo *metadados* disponível em linha<sup>2</sup>, o que torna patente a ligação da origem do termo com a Ciência da Computação – área que define *dado* como um “elemento de informação, ou representação de fatos ou de instruções, em forma apropriada para armazenamento, processamento ou transmissão por meios automáticos” (FERREIRA, 1999). No entanto, o conceito e as funções atribuídas aos metadados já vinham sendo desenvolvidos e aplicados bem antes do desenvolvimento da informática.

O termo é formado pelo prefixo grego *meta-* que, dentre outros significados, exprime a noção de reflexão sobre si, e pelo vocábulo *dado*, que se origina do latim *datum* e significa “algo que é dado”, uma notificação fatural sobre algo concreto ou abstrato. Assim, é ordinariamente conceituado como “dado sobre dado” ou “informação sobre informação”. Para Gail Hodge (2001:3), “metadados são informações estruturadas que descrevem, explanam, localizam, ou facilitam recuperar, usar ou gerenciar um recurso de informação”.

Nos campos da Biblioteconomia e da Documentação, teorias e técnicas deram origem a instrumentos com os mesmos objetivos, principalmente no que concerne à organização de informação bibliográfica, podendo mesmo ser metadado considerado “sinônimo de ponto de acesso” (ALVARENGA, 2001). Dessa maneira, a associação do termo com as práticas descritivas dessas instituições é inevitável, mormente quando se dão por meios automatizados ou semi-automatizados.

---

<sup>2</sup> O website Bibsonomy apresenta resumo de livro de 1968 intitulado *Extension of programming language concepts*, de Philip Rutherford Bagley, contendo o termo *metadata* (cf. <http://www.bibsonomy.org/bibtex/251f4903f493abc74788b96ceac4c3ab9/voj> Acesso em: maio de 2011) O website Foldoc: free on-line dictionary of computing, por sua vez, ao definir o termo, faz menção ao trabalho de Solntseff and Yezerski *A survey of extensible programming languages (Annual Review in Automatic Programming, 1974, p. 267-307)* conteria a “a noção de ‘metadados’ introduzida por Bagley” (the notion of ‘metadata’ introduced by Bagley) (cf. <http://foldoc.org/metadata> Acesso em: outubro de 2011).



(...) no ambiente de biblioteca, metadado é comumente usado para qualquer esquema formal de descrição de recursos, aplicado a qualquer tipo de objeto, digital ou não-digital. A catalogação tradicional feita por bibliotecas é uma forma de gerar metadados e MARC 21 e o conjunto de regras que acompanha seu uso, como as do AACR2 são padrões de metadados. (HODGE, 2001:3)

Como exemplificado, o formato MARC<sup>3</sup> (*Machine-Readable Cataloguing*) é um desses instrumentos da biblioteconomia, que possibilita a automatização de tarefas relacionadas administração de registros por meio da descrição de diversos tipos de documentos – de monografias a *softwares*, de imagens a partituras musicais. O AACR2 (*Anglo-american Cataloguing Rules - 2nd edition*) configura-se como um conjunto de normas para os conteúdos dessas descrições.

Assim, a principal característica que pode ser considerada para diferenciar metadados de outros tipos de ferramentas não-digitais de descrição de recursos reside em seu processamento por máquina. Metadados podem ser independentes ou acoplados aos recursos – sendo acoplados principalmente nos casos em que esses recursos são originados em meio digital, como, por exemplo, páginas da Web ou artigos de periódicos em linha.

Alvarenga (2001) relembra que a representação de descrições e registros bibliográficos tradicionais da biblioteconomia dá-se de forma independente da fonte primária, o que não necessariamente acontece com os metadados de documentos e objetos originados em meio virtual. Mas que, em ambos os casos, as descrições não se limitam a dados temáticos dos textos<sup>4</sup>, incluindo também outros tipos de pontos de acesso, como os que se referem às características físicas dos documentos e a seu contexto de produção. Corrobora Hodge, para quem metadados e interoperabilidade estão relacionados:

Uma razão importante para criar metadados descritivos é facilitar a descoberta de informação relevante. Somando-se a isso, os metadados podem ajudar a organizar recursos eletrônicos, facilitando a interoperabilidade e a integração de recursos provenientes de sistemas de informação tecnicamente obsoletos, e auxiliar na identificação digital, arquivo e preservação. (HODGE, 2001:3)

---

<sup>3</sup> O formato do registro MARC foi desenvolvido pela Library of Congress e a British Library para a comunicação de descrições bibliográficas em formato legível por computador, de modo que os registros pudessem ser reformatados, para atender a qualquer objetivo imaginável. (ROWLEY, 2002: 117)

<sup>4</sup> Para os casos em que os metadados referem-se ao conteúdo informacional de um recurso, o neologismo inglês *metacontent* foi cunhado para fins de eliminação de ambiguidade

Senso e Piñero (2003) fazem uma análise de várias concepções de metadados, e numa tentativa de síntese, chegam à conclusão de que

"Metadado é toda a informação descritiva sobre o contexto, qualidade, condição ou características de um recurso, dado ou objeto que tem a finalidade de facilitar sua recuperação, autenticação, avaliação, preservação ou interoperatividade" (SENSO; PIÑERO, 2003:99).

A NISO<sup>5</sup> (National Information Standards Organization) estabelece três tipos principais de metadados: descritivos, estruturais e administrativos. (NISO, c2004:1)

- *Metadados descritivos*: descrevem um recurso com propósito de identificação e descoberta, podendo incluir elementos como título, resumo, autor e palavras-chave.
- *Metadados estruturais*: instruem como partes de objetos compostos são agrupados; por exemplo, como páginas de um livro digitalizadas separadamente são ordenadas para formar um único objeto.
- *Metadados administrativos*: provêem informações que ajudam a administrar um recurso; por exemplo, como e quando o recurso foi criado, o tipo de arquivo e outras informações técnicas, como questões referentes à acessibilidade.

Existem ainda subconjuntos de metadados administrativos que algumas vezes são considerados como tipos à parte de metadados: os *metadados de gerenciamento de direitos*, que tratam de direitos de propriedade intelectual e os *metadados de preservação*, os quais contém dados necessários para preservar um recurso.

*Objeto de informação* – ou simplesmente *objeto*, pode ser entendido como qualquer entidade conceitual identificável, como um texto, uma imagem ou uma composição

---

<sup>5</sup> NISO é uma organização sem fins lucrativos de origem estadunidense, acreditada pela ANSI (*American National Standards Institute*), que tem como objetivo identificar, desenvolver, manter e publicar padrões técnicos para administração da informação em ambiente digital. *Website* da organização: <http://www.niso.org>

musical. *Recurso* ou *recurso informacional* é um termo usado para designar uma unidade ou peça de informação – geralmente usada como sinônimo de documento<sup>6</sup>. É importante notar, no entanto, que a noção dessa unidade pode variar conforme meios, finalidades e mesmo contexto histórico: na “catalogação tradicional”, refere-se normalmente ao registro de um item bibliográfico “inteiro”, mas em meio virtual existe a tendência para dividi-lo em partes menores, considerando recurso, assim, cada uma das declarações usadas em sua descrição. Essa tendência se fortalece conforme a capacidade de processamento do computador se sofisticava, e maiores possibilidades de integração e combinação de dados provenientes de fontes diversas se instauram. Isso posto, metadados são usados para descrever recursos em qualquer nível de agregação, como uma coleção, um item dessa coleção ou as partes desse item (por exemplo, tabelas ou imagens contidas em um artigo científico).

Outra característica marcante do tratamento da informação em meio virtual está na possibilidade de descentralização da descrição dos recursos, o que pode ocorrer em diferentes etapas de seu “ciclo de vida”. Destarte, diferentes tipos de metadados podem ser adicionados por uma variedade de agentes, como exemplificado pela NISO:

No estágio de criação, metadados sobre autores, contribuidores, fonte e audiência pretendida podem ser providas pelos seus próprios criadores. No estágio da organização, metadados sobre assuntos, histórico de publicação e tipos de licença podem ser registrados por catalogadores ou indexadores. No estágio de acesso e uso, informação de avaliação, como resenhas e comentários poderiam ser incluídas pelo usuário. (NISO, 2007)

De fato, a Tecnologia da Informação (TI) procura responder à necessidade de novos paradigmas no tratamento da informação, proporcionando a automatização e semi-automatização de tarefas como gerenciamento, organização, indexação e recuperação de grandes quantidades de fontes de informação; além de promover novas práticas e otimizar as antigas, favorecendo a realização dessas tarefas à distância ou de forma coletiva, bem como o reaproveitamento de trabalhos já realizados, que também contribuem para racionalização de processos em sistemas que objetivam facilitar o acesso à informação.

---

<sup>6</sup> No contexto do Modelo Abstrato Dublin Core, como veremos adiante, a noção de “recurso” é mais ampla, usada não só para expressar a idéia de documento, mas também a de objeto, quando considera que qualquer tipo de valor, constante de um par propriedade-valor, é também um recurso.

### **1.1.1 Esquemas de metadados**

Considerando que o valor potencial da informação aumenta com seu uso, compartilhamento e combinação com outras informações, conforme leis propostas por Moody e Walsh (1999)<sup>7</sup>, torna-se fundamental o papel dos metadados em sua estruturação, para o alcance desses fins. Esse aspecto econômico é bastante relevante, sendo pautado pela racionalização na geração de bens e serviços (comerciais ou não-comerciais) e na incessante busca por eficiência operacional, o que inclui a organização dos recursos e sua utilização da forma mais produtiva.

Comunidades de informação estão cientes que quanto mais um objeto informacional é estruturado, mais essa estrutura pode ser explorada para pesquisa, manipulação e interrelação com outros objetos informacionais. (GILLILAND, 2008)

Costa sintetiza que múltiplas abordagens para o termo comunidade, em Ciência da Informação, de uma maneira geral se referem a “agrupamentos específicos de pares dentro do universo do conhecimento” (COSTA, 2000:87). E em consonância com o Aldo Barreto (1999), para quem “a crescente produção de informação precisa ser reunida e armazenada de forma eficiente, obedecendo a critérios de produtividade na estocagem, ou seja, o maior número de documentos deve ser colocado em menor espaço possível dentro de limites de eficácia e custo”, pondera que “pesquisadores em qualquer área do conhecimento precisam comunicar seus resultados por meio de um sistema de comunicação e no mais breve espaço de tempo possível.” (COSTA, 2000:87)

Para Araújo (1995) a transmissão da informação pressupõe um processo de comunicação, objetivo de todo Sistema de Informação (SI); e aponta os Sistemas de Recuperação da Informação (SRI) como um tipo de especial interesse para a área:

---

<sup>7</sup> Moody e Walsh propuseram sete "leis de comportamento" que poderiam ser atribuídas à informação (um bem estratégico para organizações) com a finalidade de calcular seu valor econômico, sendo elas: 1) *A informação é infinitamente compartilhável*; 2) *O valor da informação aumenta com seu uso*; 3) *A informação é deteriorável*; 4) *O valor da informação aumenta com a precisão*; 5) *O valor da informação aumenta quando combinada com outra informação*; 6) *Mais informação não é necessariamente melhor*; 7) *A informação não é exaurível (quanto mais você usa, mais você tem)*.

Sistemas de informação são aqueles que objetivam a realização de processos de comunicação. Sistemas humanos de processamento da informação, sistemas eletrônicos de processamento de dados e sistemas de recuperação da informação constituem exemplos de mecanismos “especificamente planejados para possibilitar a recuperação da informação”. Dessa forma, sistemas de recuperação da informação são tipos de sistemas de comunicação que, entre outras funções, visam dar acesso às informações neles registradas. (ARAÚJO, 1995)

O sucesso de um Sistema de Informação, assim, pode ser medido pela sua capacidade em atender às necessidades da comunidade que atende, observando os preceitos de natureza econômica. Para tanto, é essencial que se mantenha o foco em seu contexto de atuação – a área de interesse de seu público – estabelecendo-se limites no escopo dos objetos com os e sobre os quais o sistema atua. Além disso, as atividades de gerenciamento, organização, preservação com vistas à recuperação e transmissão dos estoques de informação do sistema devem ser apoiadas por um conjunto de metadados que atendam a essas funcionalidades, responsáveis por prover a estruturação dos dados que compõem esses estoques.

Dessa maneira, é necessário realizar uma seleção de metadados que atendam as funcionalidades desse sistema, os quais também devem conformar um sistema próprio<sup>8</sup>, em que elementos de metadados estejam relacionados uns com os outros de modo a atingir os objetivos a que se propõem. Por exemplo, esses metadados devem responder a necessidades de especificação de dados exigidos por um *software* para realização de processos, ou a definição de campos para descrição de um determinado recurso de informação.

Em outras palavras, faz-se necessária a adoção de um *esquema de metadados*. Dentre outras definições para a palavra esquema, o dicionário Aurélio (FERREIRA, 1999) apresenta as seguintes:

#### ESQUEMA

Figura que representa, não a forma dos objetos, mas as suas relações e funções.

---

<sup>8</sup> Segundo Araújo (1995) sistemas podem ser conceituados como um conjunto de partes interrelacionadas, interagindo para atingir determinado(s) objetivo(s).

*Lóg.* Conjunto de relações que, para fins operatórios, é suficiente para caracterizar um sistema, sem, contudo, esgotar-lhe a complexidade.

Assim,

Esquema de metadados são conjuntos de elementos de dados que, por meio das relações estabelecidas entre eles, formam um sistema para descrição de objetos, voltadas para uma área de conhecimento ou de aplicação particulares. A definição ou significado dos elementos em si é conhecido como semântica do esquema; os valores atribuídos aos elementos de metadados são o conteúdo. (NISO, c2004)

Características peculiares de esquemas de metadados individuais variam com os tipos e graus de restrições que regem seu uso, como critérios usados para determinação do conteúdo (por exemplo, como definir a autoria de um documento), regras de representação de conteúdo (por exemplo, a definição de que conjunto de caracteres é permitido ou a conformidade com normas como a AACR2) e a sintaxe em que os metadados devem ser codificados (por exemplo, a linguagem de marcação ou de programação a ser usada). Alguns esquemas de metadados também estipulam normas para o formato ou origem dos valores (i.e. informações sobre a essência do recurso) a serem atribuídos aos elementos de dados (categorizadores da essência do recurso), como a imposição de um vocabulário controlado para essa finalidade.

Baseados em Nagamori e Sugimoto (2007), podemos separar os componentes de um esquema de metadados da seguinte maneira:

- Vocabulário de propriedades: um conjunto de termos definidos para expressar as propriedades do recurso. Por exemplo: título, criador, formato, etc.
- Vocabulário de valores: um conjunto de termos que expressam tipos de valores das propriedades e/ou quais são usados como valores de propriedades. Por exemplo: LCSH (*Library of Congress Subject Headings*) e DDC (*Decimal Dewey Classification*).
- Estrutura descritiva neutra: regras que definem restrições estruturais que, de uma forma geral, podem ser expressas via qualquer linguagem

de descrição de esquemas. Por exemplo: quais propriedades são obrigatórias e quais são opcionais, em que ordem elas devem aparecer no registro, etc.

- Estrutura descritiva dependente: a definição da linguagem de descrição para uma implementação específica, que por sua vez carrega as regras inerentes à própria linguagem. Por exemplo: XML, HTML e RDF/XML.

As duas primeiras categorias definem o nome e o significado dos termos, os quais dão a base semântica do esquema. (...) As duas últimas categorias definem as características sintáticas de forma abstrata e concreta, respectivamente. (NAGAMORI; SUGIMOTO, 2007)

### **1.1.2 Sintaxe e semântica de metadados**

Semântica se refere ao significado do metadado; sintaxe se refere a seu formato. Embora ambas as dimensões sejam consideradas para fins de compartilhamento de metadados, manter essas representações de forma independente é uma estratégia que pode beneficiar a utilidade de esquemas em longo prazo.

As rápidas mudanças da primeira década da Web ilustram bem isso. Testemunhamos várias versões de HTML, a emergência do XML e o desenvolvimento de tecnologias derivativas que incluem XML Schemas e RDF Schemas. A falta de estabilidade na estrutura no campo da marcação enfatiza a necessidade de manter independência entre a semântica dos elementos de metadados e sua representação sintática. (DUVAL et al., 2002)

A codificação de páginas HTML (*Hipertext Markup Language*) é a forma mais popular de inclusão de metadados na Web, os quais são integrados aos próprios recursos informacionais que estruturam. Embora não requeira mais que o protocolo HTTP (*Hiper Text Transfer Protocol*) para funcionar, garantindo a simplicidade do sistema, a desvantagem reside nas limitações para estruturação da informação inerentes à própria sintaxe, à qual a semântica está atrelada.

A linguagem XML (*Extensible Markup Language*) surge como uma alternativa que provê maior flexibilidade a essa estruturação, uma vez que sua sintaxe é totalmente

dissociada de sua semântica. *Tags*<sup>9</sup> são usadas para rotular, categorizar e organizar informação, descrevendo um documento ou a estrutura de um documento, e cabe ao criador de um documento XML especificar essas marcas e atribuir a elas significado. Constructos chamados *XML namespaces* tornam mais fácil o reaproveitamento de esquemas de metadados (como se verá mais adiante neste trabalho).

RDF (*Resource Description Framework*) é uma linguagem criada para representar informação conceitual na Web, de forma a possibilitar o processamento automatizado da informação. O RDF não se limita a uma sintaxe específica, e pode ser representada inclusive graficamente. Para fins de troca de dados, no entanto, sintaxes textuais ainda são necessárias. A linguagem XML/RDF, por exemplo, faz uso da sintaxe e funcionalidades do XML na formulação de descrições RDF.

A implementação de um esquema de metadados se dá pela associação entre declarações sobre um recurso a uma codificação específica. Para que essas declarações façam o sentido desejado, o significado dos metadados deve ser explicitado.

Para que um dado cru torne-se metadado, uma especificação de metadados deve, além da especificação de sintaxe, também definir uma interpretação da sintaxe em termos de informação sobre a coisa. Isso significa essencialmente que o padrão deve definir um mapeamento da sintaxe concreta para alguma forma de significado do metadado. (NILSSON, 2010)

Assim, para Nilsson, tal mapeamento transforma uma estrutura de dados sem significado em uma descrição, definindo-se a semântica de metadados.

Semântica é o ramo da Linguística que estuda o significado das palavras. Em consonância com as teorias linguísticas modernas, a dualidade de expressão e conteúdo em um signo é reforçada. Conforme Hjelmslev (1961), sentido seria o fator comum a todas as línguas, cuja execução (operada na dualidade do signo) é diferente em cada uma delas: um pensamento não delimitado, uma “nebulosa”, segundo Saussure, que será ordenado e articulado em cada língua de diferentes maneiras.

---

<sup>9</sup> O conteúdo do documento em si, como texto, imagem e outros dados é marcado pela tag (marca) apropriada, no início e ao final de cada peça de informação, de acordo com as regras sintaxe dessa linguagem. Em um esquema de metadados criado em XML, tags são utilizadas para compor elementos.



Analogamente, semântica de metadados pode ser compreendida como o significado atribuído a elementos de metadados, compartilhado entre aqueles que deles façam uso.

A questão da semântica dos metadados tem ganhado corpo nas pesquisas nos últimos anos, principalmente no que se refere às teorias relacionadas à Web Semântica<sup>10</sup>, que visam aprimorar meios de manipulação de dados de forma a automatizar a interpretação do conteúdo de documentos, por meio de inferências lógicas.

Nessa interpretação automatizada do conteúdo<sup>11</sup>, no entanto, a semântica ainda é claramente limitada pela codificação – ou *binding* – do esquema de metadados utilizado, representando a semântica, nesse caso, nas palavras de Paepcke et al. (1998) “a sintaxe de sua execução”.

A concepção de semântica de metadados adotada neste trabalho é mais ampla, abarcando a associação entre expressão e conteúdo que é convencionalizada socialmente, dentro de uma comunidade uso, limitada pelas do contexto de um domínio do conhecimento - cujas fronteiras são igualmente convencionais.

Nilsson (2010) assinala a existência de três tipos de semântica de metadados:

**Semântica Informal:** Definição de significado dirigido à interpretação humana, geralmente expresso em texto nas especificações de metadados.

**Semântica processável por máquina:** Expressão de significado por meio de um formato passível de ser processado por meio de uma linguagem de programação. O uso desse tipo de codificação abre caminho para a interpretação automatizada de expressões de metadados.

---

<sup>10</sup> Segundo seu idealizador, Tim Berners-Lee, Web Semântica é uma extensão da Web atual, que visa favorecer o processamento do conteúdo informacional de seus recursos por agentes inteligentes de software, desenvolvidos por meio de técnicas de Inteligência Artificial. Para tanto, “computadores devem acessar coleções estruturadas de informação e conjuntos de regras de inferência que eles podem usar para conduzir raciocínio automatizado” (BERNERS-LEE; HENDLER; LASSILA, 2001)

<sup>11</sup> Termos derivados do verbo inglês *to parse* (analisar) são comumente usados na literatura para se referir a interpretação automatizada de dados por softwares.

**Semântica formal:** Expressão do significado em termos de um modelo matemático formal. Por meio desse modelo, combinam-se ontologias e formatos “compreendidos” por máquina, possibilitando a racionalização automatizada entre padrões semanticamente sobrepostos.

## **1.2 Padrões e interoperabilidade**

A configuração de um esquema de metadados, por responder às necessidades de uma área de conhecimento ou de aplicação particulares, defronta-se com questões ligadas a possibilidade de troca de dados entre sistemas de informação.

Segundo Duval (2001), essas questões tratam de fronteiras contextuais de natureza técnica, linguística e social. Para ele, maior complexidade reside nos problemas linguísticos – como quando metadados devem ser traduzidos, e nos sociais – como quando forma e conteúdo de metadados são marcados por convenções e métodos próprios de uma comunidade, determinados sob influência regional ou cultural. Os aspectos técnicos, como os que se referem à compatibilidade de *hardware* e *software* entre bases de dados, seriam mais simples de serem solucionados.

Alves e Souza (2007) consideram esse aspecto técnico na definição de interoperabilidade, com segue:

A capacidade de bases de dados trocarem e compartilharem documentos, consultas e serviços, usando diferentes plataformas de hardware e software, estrutura de dados e interfaces, é chamada de interoperabilidade. Através dessa troca e compartilhamento são realizadas interações entre sistemas. (ALVES; SOUZA, 2007)

Patel et al. (2005) apresentam uma visão mais ampla do termo, definindo interoperabilidade como qualquer forma de comunicação entre sistemas, ou a habilidade de um sistema de fazer uso de dados de uma fonte previamente desconhecida. Para ele, interoperabilidade em geral concerne à capacidade de diferentes sistemas de informação se comunicarem.

Para Nilsson (2010), “a administração e troca de metadados é uma atividade central em muitos sistemas que gerenciam objetos digitais e não-digitais, tais como sistemas de gerenciamento de conteúdo, repositórios de objetos de aprendizagem e bibliotecas”. Assim, apresenta sua definição para interoperabilidade focando nesse aspecto dos Sistemas de Informação. Segundo ele, interoperabilidade de metadados é

(...) a habilidade de dois ou mais sistemas ou componentes troca dados descritivos sobre coisas, e interpretar os dados descritivos que foram trocados de maneira que haja consistência com a interpretação do criador dos dados. (NILSSON, 2010)

Quando esquemas de metadados são desenvolvidos e mantidos por instituições padronizadoras como a ISO<sup>12</sup> e a W3C<sup>13</sup> ou por organizações que tomaram para si essa responsabilidade, como a *Dublin Core Metadata Initiative* (que responde pelo padrão Dublin Core<sup>14</sup>), esses esquemas são conhecidos como “padrões de metadados”. Exemplos de padrões de metadados são TEI<sup>15</sup>, METS<sup>16</sup>, MARC<sup>17</sup>, MODS<sup>18</sup>, GILS<sup>19</sup> etc.

Duval (2001) nos alerta que “(...) a noção de interoperabilidade não é binária, onde sistemas poderiam ser ou interoperáveis ou não interoperáveis. Ao invés disso, existe um grau mais alto ou mais baixo de interoperabilidade.”

Ouksel e Sheth (1999), retomando estudo prévio de Sheth, identificam três tipos de *heterogeneidade da informação* que correspondem a três tipos de potencial interoperabilidade:

---

<sup>12</sup> ISO (*International Organization for Standardization*) é uma organização não-governamental que desenvolve e publica padrões internacionais. (cf. <http://www.iso.org/iso/>)

<sup>13</sup> O Consórcio *World Wide Web* (W3C) foi fundado em 1994, por Tim Bernes-Lee, com o objetivo de desenvolver padrões para a Web.

<sup>14</sup> Dublin Core é um padrão de metadados genérico, apropriado para descrição de uma ampla gama de recursos tais como textos, imagens e vídeo, de semântica relativamente fácil de ser apreendida por usuários leigos, e sobre o qual se tratará com mais profundidade adiante neste trabalho.

<sup>15</sup> *Text Encoding Initiative* é um padrão para representação de textos em formato digital, voltado para a área de humanas.

<sup>16</sup> *Metadata Encoding & Transmission Standard*, baseado em XML, tem como função codificar metadados descritivos, administrativos e estruturais de objetos de bibliotecas digitais.

<sup>17</sup> *Machine Readable Cataloguing* é usado para representar informação bibliográfica idealizado para facilitar troca de registros entre bibliotecas.

<sup>18</sup> *Metadata Object Description Schema*: um esquema de elementos bibliográficos que pode ser usado para vários propósitos, principalmente em aplicações referentes a bibliotecas.

<sup>19</sup> *Global Information Locator Service*. Trata-se de um padrão voltado para instituições governamentais, companhias e outras organizações que visa a descoberta de recursos.

- Estrutural: variedade em modelos de dados, estrutura de dados e esquemas
- Sintática: diferenças em codificação ou formato
- Semântica: inconsistências em terminologia e significados

Segundo Alves (2005), a interoperabilidade estrutural está relacionada com características das estruturas dos elementos descritivos que compõem um padrão ou formato de metadados.

A interoperabilidade estrutural, para Cássia Barreto (1999:86),

(...) define cada elemento componente de um padrão de metadados, descreve os seus tipos, a escala de valores possíveis para esses elementos e os mecanismos utilizados para se agrupar (ou relacionar) esses elementos de modo a que possam ser processados de forma automática.

A interoperabilidade sintática é promovida por meio dos códigos utilizados para o intercâmbio de dados. "A sintaxe provê uma linguagem comum para representar a estrutura do metadado". (BARRETO, 1999:87)

Sem a interoperabilidade sintática, dados e informações não podem ser manipulados propriamente no que se refere a seus formatos, codificações, propriedade, valores, tipos, etc., não podendo ser mesclados ou trocados. (PATEL et al., 2005)

A interoperabilidade semântica refere-se ao significado dos elementos que descrevem o recurso, levando-se em consideração a forma como se associam uns com os outros, i.e. o contexto no qual estão inseridos. Esse tipo de interoperabilidade, segundo Marcondes e Sayão (2002:50) requer o uso generalizado de instrumentos comuns de descrição temática, tendo como reflexo imediato a melhoria na qualidade da recuperação das informações e a otimização da consulta a sistemas interligados. Alves (2005:78) defende que o uso de vocabulários específicos, ontologias e/ou padrões de metadados são essenciais para assegurar esse tipo de interoperabilidade.

Em certa medida, o intuito de cada um dos três tipos de interoperabilidade se sobrepõe, e é em conjunto que atuam, configuram-se como importantes princípios para o funcionamento otimizado de sistemas e serviços de informação em rede.

A interoperabilidade está na ordem do dia no que diz respeito a todos os tipos de sistema de informação digital, mas ganha proeminência com a Web. Segundo Oliveira (2005:35), a busca da interoperabilidade é um objetivo da área da Ciência da Computação, principalmente depois da criação da Web, configurando-se como um ambiente informacional global, com múltiplas plataformas e sistemas operacionais, sendo seu conteúdo caracterizado pela heterogeneidade, que vai desde ensino e cultura a entretenimento e comércio eletrônico, e ilustra:

A própria Web é um exemplo de interoperabilidade, embora em uma escala limitada: o protocolo http permite a qualquer usuário navegar por páginas disponibilizadas em computadores os mais diferentes, com os mais diversos sistemas operacionais, criando um ambiente interoperável. Um organismo internacional, o W3C, se encarrega de propor padrões de TI que tornem a Web cada vez mais interoperável. (OLIVEIRA, 2005:35)

O ambiente Web possibilita a reaproveitamento de esquemas e padrões por meio de ferramentas como o protocolo HTTP e URIs (*Uniform Resource Identifier*) e a linguagem de marcação XML – ou, mais precisamente, XML Schema, uma linguagem que, sendo ela mesma baseada em XML, provê um meio de definir a estrutura, conteúdo e semântica de documentos XML.

A combinação de tais ferramentas propicia esse reuso por meio de perfis de aplicação, favorecendo-se dos princípios de modularidade e extensibilidade de esquemas de metadados.

*Modularidade* é princípio que permite a transposição de partes de um esquema para que seja usado por outro, ajustando-os para a formação de um todo homogêneo e funcional. Essa é uma prática que permite a projetistas de bases de dados e outras aplicações que utilizam metadados combinarem “peças” pré-fabricadas de esquemas que já foram empregadas, “testadas e aprovadas” por experiências bem sucedidas de projetos similares.

*Extensibilidade* é a capacidade de um esquema de receber aquelas partes (módulos) em sua arquitetura, aproveitando-se desse incremento para fornecer possibilidades descritivas inexistentes no esquema “puro”. Quando tais módulos são usados para prover maior precisão descritiva relativas a metadados pré-existentes, procede-se ao “refinamento” dos metadados do esquema receptor. Uma boa prática, no entanto, que visa manter a interoperabilidade com aplicações que não estejam aptas para fazer uso desses refinamentos modulares imprevistos, é a de prepará-las para ignorar a extensão, voltando-se para o conjunto de metadados original.

### **1.2.1 Perfis de aplicação**

Ainda que padrões e outros esquemas de metadados sejam desenvolvidos de acordo com as necessidades previstas para uma comunidade usuária, não raramente uma porção dessa comunidade encontra inconveniências para aplicá-los da forma como foram concebidos, devido a características desse subgrupo. Os perfis de aplicação surgem como uma resposta a essa questão pragmática: casos assim podem ser solucionados pela “customização” do esquema adotado. Ao efetuar as alterações necessárias no esquema (que de uma forma geral serve bem à comunidade, tendo sido, portanto, definido como modelo), atende-se a requerimentos de aplicações locais sem deixar de manter um grau de interoperabilidade com o conjunto de usuários a que pertence num nível mais amplo.

O termo perfil é largamente usado para se referir a um documento que descreve como padrões ou especificações são empregados para apoiar os requerimentos de uma aplicação particular, função, comunidade ou contexto. Na comunidade de metadados, o termo perfil de aplicação tem sido aplicado para descrever a costura de padrões para aplicações específicas. (NILSSON; BAKER; JOHNSTON, 2008)

A crescente interdisciplinaridade de domínios do conhecimento e as convergências tecnológicas figuram fenômenos que também podem se beneficiar da funcionalidade dos perfis de aplicação. Assim como é plausível que subgrupos de comunidades usuárias decidam, por exemplo, simplificar um esquema de metadados pela subtração de elementos considerados supérfluos, ou resolvam refinar (restringindo o escopo) ou acrescentar elementos a seu esquema padrão, respondendo a questões pontuais, elementos de metadados referentes a uma disciplina podem ser incorporados ao modelo de metadados que estrutura a informação

de outra disciplina, gerando registros que suportem a pesquisa de conteúdo em que relações entre elas sejam constituídas. Da mesma maneira, a fim de alcançar seu objetivo, aplicações que propiciam a convergência de tecnologias podem fazer uso da mesma idéia básica por trás dos perfis de aplicação.

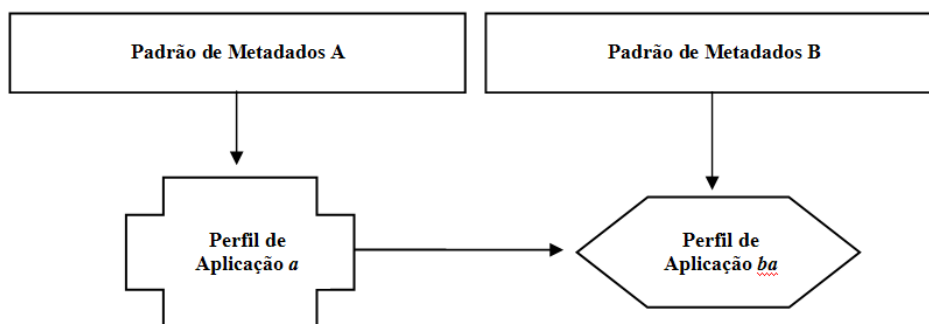
Em suma, podemos afirmar que um perfil de aplicação é um tipo de esquema de metadados composto por elementos de dados retirados de um ou mais padrões ou esquemas de metadados, elaborado com o objetivo de atender aos requisitos de um subgrupo em uma comunidade usuária, a qual em um nível mais generalizado é atendida pelo padrão de metadados alvo das modificações.

Os perfis de aplicação usualmente consistem de elementos de metadados extraídos de um ou mais esquemas de metadados, combinados por implementadores em um esquema composto, e otimizados para uma aplicação local particular. (HEERY; PATEL, 2000)

Em seu estudo sobre metodologias de interoperabilidade e padronização de metadados, Chan e Zeng (2006) classificam o processo em que um único esquema de metadados dá origem a subconjuntos desse esquema (geralmente de escopo mais limitado) como “método de derivação”, ressaltando que nos casos em que um conjunto de metadados dá origem a outro, o novo esquema é dependente do esquema-fonte. Segundo os autores, essa derivação pode acontecer por meio de adaptação, modificação, expansão, adaptação parcial, tradução, etc. E ilustram:

Em uma coleção de bases de dados digitais onde diferentes componentes têm diferentes necessidades e diferentes requerimentos quanto aos detalhes descritivos, um esquema existente complexo como o formato MARC pode ser usado como “fonte” ou “modelo” do qual novos e mais simples esquemas podem ser derivados. (CHAN; ZENG, 2006)

Uma representação gráfica da formação de perfis de aplicação por meio do reuso de elementos de dados existentes pode ser observada na figura 1:



**Figura 1** – Obtenção de perfis de aplicação por meio de derivação e combinação de elementos de dados de padrões de metadados pré-existentis.

Na figura 1 observamos a constituição de dois esquemas de metadados: o *perfil de aplicação a* composto a partir de elementos de dados do Padrão de Metadados A; e o *perfil de aplicação ba*, composto por elementos de dados do Padrão de Metadados B e também por elementos de dados existentes no *perfil de aplicação a*.

Essa relação entre esquemas fonte e produto evidencia uma espécie de conflito envolvendo criadores de padrões de metadados e implementadores de metadados - embora muitas vezes esses profissionais desempenhem ambos os papéis. Isso acontece porque, enquanto os criadores de padrões de metadados estão mais preocupados em sistematizar as relações entre os elementos de dados, esclarecer seus vínculos e avaliar resultados e aplicações, implementadores tendem a utilizar padrões de metadados de forma prática, adaptando-os para seus propósitos.

A intenção de ambos é descrever recursos para manipulá-los de alguma forma. Criadores de padrões buscam o acordo numa abordagem comum para assegurar interação entre sistemas e economia de escala. Enquanto implementadores, apesar de poderem almejar, em parte, o uso de padrões, também querem descrever aspectos específicos de um recurso de uma maneira “especial”. (HEERY; PATEL, 2000)

A liberdade dos implementadores, no entanto, encontra barreiras: primeiramente, diferentes padrões de metadados apresentam níveis de flexibilidade variáveis: alguns padrões podem ser bastante prescritivos e deixar relativamente poucas opções para alterações e adaptações, enquanto oferecem uma gama de opções, que demandam um considerável grau de seleção e “costura” para sua implementação (NILSSON; BAKER; JOHNSTON, 2008). Além disso, implementadores de metadados não podem criar novos elementos, que poderiam mais



diretamente atender às demandas de seus serviços, estando assim limitados aos elementos de dados apropriadamente declarados em seus respectivos *namespaces*: “coleção de todos os nomes de elementos, tipos e atributos de um esquema específico”. (DYKES, 2005:186).

Dessa maneira, havendo real necessidade de elementos de dados novos, eles devem ser declarados, ou seja, devem constituir um *namespace* que os referencie. Segundo determinação do Consórcio *World Wide Web* (W3C) – comunidade internacional responsável pelo desenvolvimento de padrões para Web – *namespaces* devem ser declarados por meio de nomes de domínios (“rótulos” utilizados para identificar campos de autonomia administrativa, autoridade ou controle na Internet), expressos em linguagem XML, garantindo a cada elemento do esquema uma identificação única.

No âmbito das especificações do W3C para a linguagem XML e o esquema RDF, *namespaces* são nomes de domínio associados a elementos que, juntamente com o nome que individualiza um elemento, produzem uma URL que identifica univocamente um elemento. (HEERY; PATEL, 2000)

## **Namespaces**

Em Ciência da Computação, *namespace* é um contêiner abstrato criado para agrupar identificadores (nomes) de forma unívoca dentro de um determinado contexto, permitindo a desambiguação de identificadores homônimos que existam em diferentes contextos. Linguagens de programação que suportam *namespaces* especificam regras próprias para determinar o vínculo entre identificadores e esses *namespaces*.

No sistema operacional de um computador, por exemplo, o diretório funciona como um *namespace*, cujos arquivos contidos são identificados sem equívocos por meio de seus “endereços”: a descrição de todo o caminho a ser percorrido por uma aplicação (ou uma pessoa) para se chegar até esses arquivos. O protocolo http, da Web, é outro exemplo de *namespace*, configurando-a como um ambiente em que URLs apontados após a expressão *http://* são garantidamente identificáveis por navegadores de forma singular.

Em XML, *namespace* é uma coleção de nomes usados para designar os elementos e atributos que compõem as marcas de um documento específico, provendo a esses nomes

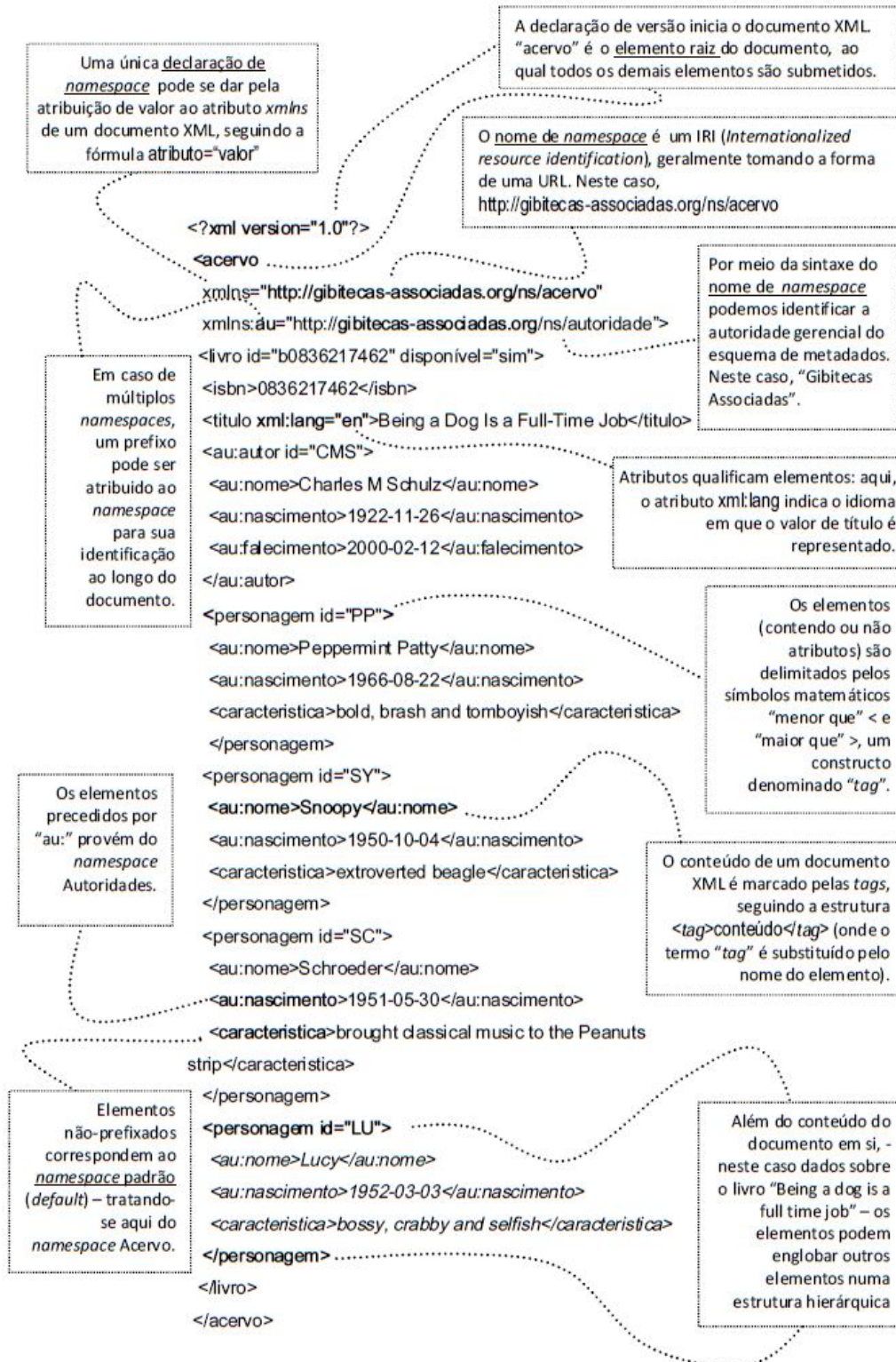
(i.e., seus elementos e atributos) uma identificação única em comparação a nomes de outros *namespaces* - que estejam em conformidade com as diretrizes para *namespaces* especificadas pelo W3C (2006). Também segundo norma do W3C, um *namespace* é identificado por uma IRI, que pode tomar a forma de uma URL (*Uniforme Resource Locator*) ou URN (*Uniforme Resource Name*), ainda que, nos termos do W3C,

(...) um namespace não precisa estar incorporado num registro de autoridade real, nem os elementos precisam ser identificados por uma URL definida por um endereço da web real. No entanto, essa é considerada uma boa prática para assegurar um ambiente de metadados bem gerenciado. (HEERY; PATEL, 2000)

Esquemas de metadados criados como documentos baseados em XML possibilitam o reuso de seus elementos por meio da declaração de seus respectivos *namespaces*, por parte de aplicações interessadas na adoção desses elementos para organização da informação com as quais lidam.

De acordo com a sintaxe do XML, em uma declaração de *namespace* é definida na *tag* de abertura de um elemento, estendendo a seus sub-elementos a abrangência do *namespace*. O IRI que identifica o namespace é inserido entre aspas ( “ ” ) como valor do atributo *namespace* (expresso como “xmlns”). Essa atribuição é dada de forma direta ou indireta – neste último caso, fazendo uso de um prefixo.

No documento XML representado na figura 2, elementos de metadados provenientes de dois *namespaces* são combinados para descrição de um item bibliográfico constante do acervo da organização fictícia Gibitecas Associadas. “Acervo”, que é o *namespace* padrão e, portanto, domínio dos elementos não-prefixados, enquanto os elementos precedidos pelo prefixo “au” correspondem àqueles provenientes do *namespace* “Autoridades”.



**Figura 2** – A declaração de dois *namespaces* distintos é expressa no documento XML representado nesta figura, que também fornece outras informações sobre a sintaxe da linguagem de marcação.

A identificação única dos elementos possibilitada pela adoção de *namespaces* permite também diferenciar, em esquemas de metadados distintos, elementos que apresentem o mesmo nome, assim como atribuir nomes diferentes a um mesmo elemento, o que é útil na tradução de perfis de aplicação.

Por exemplo, nos padrões DCMES e LOM<sup>20</sup> os elementos correspondentes a título não dizem respeito a mesmo tipo de informação, o que se pode constatar nas descrições fornecidas nos respectivos esquemas, na quadro 1. Em DCMES, esse elemento tem uma abrangência mais ampla, referindo-se ao nome dado a um recurso; em LOM, destina-se a nomear um objeto de aprendizagem<sup>21</sup>:

	<b>Elemento</b>	<b>Descrição</b>	<b>Namespace</b>
<b>DCMES</b>	Title	A name given to the resource.	<a href="http://purl.org/dc/elements/1.1/">http://purl.org/dc/elements/1.1/</a>
<b>LOM</b>	Title	Name given to this learning object.	<a href="http://ltsc.ieee.org/xsd/LOM/">http://ltsc.ieee.org/xsd/LOM/</a>

**Quadro 1** – Exemplos de coincidências em nomes e definição de elementos nos padrões Dublin Core e IEEE LOM.

No caso de traduções para outros idiomas de um perfil de aplicação, é importante salientar que a mudança ocorre apenas na designação do elemento, definida para o leitor humano. O padrão Dublin Core especifica como propriedades de um metadado “nome” e “rótulo”, segundo o qual o nome de um elemento é aquele a ser utilizado numa declaração de *namespace* (e portanto necessário na análise realizada por *software*) e o rótulo de um elemento é destinado para a interface humana da aplicação.

A restrição ao uso apenas de elementos de dados declarados em *namespaces* é vantajosa, pois favorece a pesquisa prévia e cuidadosa de padrões de metadados existentes por parte do implementador, evitando a criação desnecessária de novos esquemas. Tal risco foi identificado por Rocha, que adverte:

<sup>20</sup> A aplicação em formato para ambientes de educação *on-line* (*e-learning*) pode usar o formato LOM (*Learning Object Metadata*) que é uma ampliação de conteúdos (normalmente apenas textos) para Objetos (que inclui outros elementos).

<sup>21</sup> Em LOM, objeto de aprendizagem (*learning object*) define-se como qualquer entidade, digital ou não digital, que pode ser usada para aprendizagem, educação ou treinamento, conforme o documento *Draft Standard for Learning Object Metadata*, disponível em [http://ltsc.ieee.org/wg12/files/LOM\\_1484\\_12\\_1\\_v1\\_Final\\_Draft.pdf](http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf).

Na *web* comunidades especializadas em diversas áreas do conhecimento têm definido esquemas próprios de metadados para padronizar a descrição dos seus recursos. (...) O surgimento de uma grande diversidade de esquemas de metadados, entretanto, pode levar a uma sobreposição de esquemas, com esquemas sendo definidos para as mesmas finalidades. (ROCHA, 2004:114-115)

A duplicidade de elementos com funções/semânticas similares origina-se de iniciativas independentes de estabelecimento de esquemas e pode afetar negativamente a interoperabilidade entre sistemas que, de outra maneira, poderiam compartilhar informações de forma automatizada, com o uso comum de seus dados pelas mesmas aplicações.

## 2. Estudo do padrão Dublin Core

### 2.1 Origens do Padrão Dublin Core

Durante a 2ª Conferência Anual do Consórcio *World Wide Web*, em Chicago, uma “conversa de corredor” se deu “centrada na discussão da necessidade de uma infraestrutura para possibilitar a descoberta de recursos na nascente web” (HARPER, 2010:21), o que levaria a um primeiro encontro destinado a discutir um formato de metadados básico para recursos da web, e ao nascimento do *Dublin Core Metadata Initiative* (DCMI), contando com a participação de profissionais da informação, da computação e afins.

Um conjunto de elementos de metadados foi proposto, simples o suficiente para ser empregado pelos próprios produtores de recursos informacionais, ainda que não tivessem experiência com esse tipo de atividade, e genérica o bastante para que pudesse ser largamente adotado, fixando um padrão.

Entre 1995 e 2001, o DCMI promoveu uma série de *workshops* e reuniões para discutir essa necessidade e para desenvolver um padrão extensível e amplamente aplicável. A necessidade percebida era muito específica e focada numa descrição simples com propósito de descoberta. Em 1999, o conjunto de quinze elementos foi finalizado e publicado como uma RFC. O Conjunto de Elementos de Metadados Dublin Core (DCMES) tornou-se um padrão nacional em 2001 (ANSI/NISO Z39.85) e um padrão internacional em 2003 (ISO 15386). (HARPER, 2010:21)

O primeiro desses encontros foi realizado em Dublin (no Estado norte-americano de Ohio), de onde se origina o primeiro nome que designa esse padrão de metadados. “Core”, que pode ser traduzido como “cerne”, faz menção à característica basilar desses elementos.

O resultado dessa busca, finalmente apresentado em 1999, consistiu os quinze elementos de metadados apresentado no quadro seguinte.

<b>CONJUNTO DE ELEMENTOS DE DADOS DUBLIN CORE</b>	
Título ( <i>Title</i> )	Um nome dado ao recurso
Criador ( <i>Creator</i> )	Uma entidade principal responsável pela elaboração do conteúdo do recurso
Contribuidor ( <i>Contributor</i> )	Uma entidade responsável pela contribuição ao conteúdo do recurso
Assunto ( <i>Subject</i> )	Assunto referente ao conteúdo do recurso
Publicador ( <i>Publisher</i> )	A instituição responsável pela difusão do recurso
Data ( <i>Date</i> )	Data associada com um evento no ciclo de vida do recurso
Formato ( <i>Format</i> )	Manifestação física ou digital do recurso
Linguagem ( <i>Language</i> )	Idioma do conteúdo intelectual do recurso
Cobertura ( <i>Coverage</i> )	A extensão ou cobertura espaço-temporal do conteúdo do recurso
Direitos ( <i>Rights</i> )	Informação sobre os direitos relacionados ao recurso e a seu uso
Relação ( <i>Relation</i> )	Uma referência a outro recurso que se relaciona com o recurso
Fonte ( <i>Source</i> )	Uma referência para outro recurso do qual o recurso é derivado
Tipo ( <i>Type</i> )	A natureza ou gênero do conteúdo do recurso
Descrição ( <i>Description</i> )	Uma descrição sobre o conteúdo do recurso
Identificação ( <i>Identifier</i> )	Identificação não ambígua do recurso dentro de um dado contexto

**Quadro 2** – Relação dos elementos cerne do padrão Dublin Core

No padrão Dublin Core, todos os elementos são opcionais e repetíveis, o que significa dizer que não existem restrições acerca de uma quantidade mínima ou máxima referente a qualquer dos elementos no registro de um item por eles descrito. As descrições são estruturadas por *pares propriedade-valor*, ou seja, para cada elemento (propriedade) usado em uma descrição, existe um valor correspondente.

Um valor, nesse contexto, é a representação de “uma entidade física, digital ou conceitual”, ou ainda um “literal”, conforme terminologia do Modelo Abstrato Dublin Core, visto mais adiante, que descreve um recurso de informação. Uma entidade física, por exemplo, pode ser uma pessoa, uma região geográfica ou um artefato – essas entidades poderiam ser descritas utilizando-se os elementos Criador (ou Contribuidor), Cobertura e Título (ou Tipo), respectivamente; uma data ou um assunto são entidades do tipo conceitual; um registro sonoro em formato MP3 pode ser considerado uma entidade digital; a grafia do nome de uma pessoa, um literal. Dessa maneira é possível concluir que um valor, por si mesmo, é também um recurso de informação.

## 2.2 Iniciativa de Metadados Dublin Core

A Iniciativa de Metadados Dublin Core ou DCMI (*Dublin Core Metadata Initiative*) é a organização que responde pelo desenvolvimento, manutenção e promoção do padrão Dublin Core, assim como administra toda a estrutura que envolve o padrão, cada vez mais complexa devido à utilização de seus elementos por parte de uma comunidade internacional crescente de usuários e desenvolvedores, representantes de variados propósitos e modelos de negócios.

Além da busca pelo atendimento a diversas áreas produtoras de conhecimento, a aspiração pela internacionalidade do padrão Dublin Core demanda o cuidado por manter a semântica de seus elementos o mais abrangente, de forma a encontrar termos equivalentes em significado nas diversas linguagens a que venha a ser traduzidos.

A semântica do DC tem sido estabelecida por um grupo internacional e multidisciplinar de profissionais da Biblioteconomia, Ciência da Computação, codificadores de texto, comunidades de museus e outros campos de escolaridade e prática relacionados. (HILLMANN, 2005)

O DCMI é uma organização sem fins lucrativos, e a gratuidade tem sido uma estratégia para disseminação do padrão Dublin Core. A ausência de taxas para obtenção das informações produzidas pela organização potencializa a adoção do padrão em larga escala.



Para atender a um público diversificado, faz parte da política do DCMI manter sua linha de trabalho neutra e multidisciplinar, encorajando o uso do padrão por setores públicos e privados, favorecendo a interoperabilidade entre diferentes nichos de informação. O desenvolvimento de padrões e dos documentos que os regem são assim discutidos de forma aberta, por grupos e indivíduos interessados nos resultados desses trabalhos, em listas públicas na Internet e em encontros promovidos periodicamente.

Um fator que contribuiu para a larga disseminação do padrão foi sua adoção pelo projeto *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH). Refletindo a solidificação do trabalho e dos resultados, o padrão Dublin Core foi ratificado pelo IETF RFC 5013, ANSI/ISSO z39.85:2007 e ISO 15836:2009.

Interações com outros criadores e implementadores de padrões de metadados e sob a influência de comunidades relacionadas à Web e a organização automatizada da informação, como as comunidades envolvidas com RDF (*Resource Discovery Framework*), Web Semântica e *Linked-data*<sup>22</sup>, contribuíram para a diversificação das linhas de trabalho e extensão dos objetivos da organização.

O DCMI cresceu muito além do conjunto dos 15 elementos carregando seu nome. Hoje a Iniciativa provê um *framework* e um modelo, bem como um conjunto de princípios para projetar metadados. Ele é também uma diversificada comunidade ligada pelo interesse comum no desenvolvimento dos fundamentos metadados ricos e interoperáveis. O valor real da proposta do Dublin Core está em seu comprometimento com a interoperabilidade, assim como na aplicabilidade das diretrizes e recomendações da organização para esquemas de metadados, conjunto de elementos e sintaxe de implementação. (HARPER, 2010:22)

## 2.3 Termos Dublin Core

Harper (2010:22) afirma que, para algumas comunidades de uso – como a formada pelas bibliotecas –, a proposta inicial dos quinze elementos do padrão Dublin Core foi

---

<sup>22</sup> O termo *Linked Data* foi cunhado por Tim Berners-Lee em 2006, para designar a técnica que combina o uso do protocolo http, de URIs na identificação de metadados e da linguagem RDF para interconexão hipertextual de documentos no escopo da Web Semântica.

considerada simplista, visto que a implementação desse conjunto de elementos não se mostrava suficiente em aplicações que requeriam uma complexidade maior na descrição de seus recursos.

Ainda que a proposta inicial do DCMI se pautasse por um modelo genérico de classificação, a fim de possibilitar sua adoção da forma mais ampla possível, “discussões paralelas dentro e fora da organização” levaram ao desenvolvimento do que foi chamado Dublin Core Qualificado (em oposição a Dublin Core Simples, representando os quinze elementos cerne), “o qual apareceu no *website* do DCMI em julho de 2000”. (HARPER, 2010:23)

O Dublin Core Qualificado acrescenta três *elementos* ao conjunto inicialmente proposto e um grupo de *elementos de refinamento*, assim como uma seleção de *esquemas de codificação* recomendados com a função de prover valores padronizados ao conteúdo dos elementos ou “propriedades” – designação que os elementos de metadados passaram a receber na documentação do DCMI.

Nessa proposta renovada de metadados, a alteração na terminologia, a elaboração de um modelo de dados para representação das relações entre os metadados Dublin Core (em contraposição à preocupação inicial marcada pela mera definição de formato de dados) e a inclusão da categoria Classes ao conjunto de refletem influência direta dos trabalhos desenvolvidos pela comunidade RDF e a busca por uma compatibilização dos modelos de dados de ambas as iniciativas.

Todos os componentes que fazem parte dessas inovações e que são mantidos pelo DCMI são chamados “termos” na terminologia da organização, constituindo o *DCMI Metadata Terms* (DCMI Terms). A corrente configuração desses componentes foi aprovada em janeiro de 2008 pelo comitê *DCMI Usage Board* (DCMI USAGE BOARD, 2010a), e o resultado pode ser conferido no quadro 3:

---

## Termos do Dublin Core

---

<i>Elementos Dublin Core (versão 1.1)</i>	contributor, coverage, creator, date, description, format, identifier, language, publisher, relation, rights, source, subject, title, type
<i>Versão estendida dos elementos, com a inclusão de três novas propriedades (em negrito)</i>	<b>audience</b> , contributor, coverage, creator, date, description, format, identifier, language, <b>provenance</b> , publisher, relation, rights, <b>rightsHolder</b> , source, subject, title, type
<i>Qualificadores</i>	abstract, accessRights, accrualMethod, accrualPeriodicity, accrualPolicy, alternative, available, bibliographicCitation, conformsTo, created, dateAccepted, dateCopyrighted, dateSubmitted, educationLevel, extent, hasFormat, hasPart, hasVersion, instructionalMethod, isFormatOf, isPartOf, isReferencedBy, isReplacedBy, isRequiredBy, issued, isVersionOf, license, mediator, medium, modified, references, replaces, requires, spatial, tableOfContents, temporal, valid
<i>Esquemas de codificação vocabular</i>	<ul style="list-style-type: none"> <li>• DCMIType (DCMI Type Vocabulary)</li> <li>• DDC (Dewey Decimal Classification)</li> <li>• IMT (Internet Media Types)</li> <li>• LCC (Library of Congress Classification)</li> <li>• LCSH (Library of Congress Subject Heading)</li> <li>• MESH (Medical Subject Headings)</li> <li>• NLM (National Library of Medicine Classification)</li> <li>• TGN (Getty Thesaurus of Geographic Names)</li> <li>• UDC (Universal Decimal Classification)</li> </ul>
<i>Esquemas de codificação sintática (datatypes)</i>	<ul style="list-style-type: none"> <li>• Box (DCMI Box Encoding Scheme)</li> <li>• ISO3166 (Padrão ISO de códigos para nomes de países)</li> <li>• ISO639-2 (Padrão ISO de códigos com dois caracteres para idiomas)</li> <li>• ISO639-3 (Padrão ISO de códigos com três caracteres para idiomas)</li> <li>• Period (DCMI Period Encoding Scheme)</li> <li>• Point (DCMI Point Encoding Scheme)</li> <li>• RFC 1766, RFC3066, RFC4646 e RFC5646 (tags RFC para identificação de idiomas)</li> <li>• URI (Uniform Resource Identifier)</li> <li>• W3C-DTF (W3C Date and Time Formats Specification)</li> </ul>
<i>Classes</i>	Agent, AgentClass, BibliographicResource, FileFormat, Frequency, Jurisdiction, LicenseDocument, LinguisticSystem, Location, LocationPeriodOrJurisdiction, MediaType, MediaTypeOrExtent, MethodOfAccrual, MethodOfInstruction, PeriodOfTime, PhysicalMedium, PhysicalResource, Policy, ProvenanceStatement, RightsStatement, SizeOrDuration, Standard

**Quadro 3** – Termos do padrão Dublin Core Qualificado mantidos pelo DCMI

## Propriedades e qualificadores

Propriedades (ou elementos) representam as características dos recursos escolhidas como relevantes para sua descrição. “Em um registro baseado em metadados de um sistema, as propriedades são usualmente chamadas ‘campos’”. (DCMI Media Wiki, 2011)

O Dublin Core Qualificado acrescenta as propriedades *audience*, *provenance* e *rightsHolder* (audiência, proveniência e detentor dos direitos, respectivamente) ao conjunto dos quinze elementos inicialmente propostos pelo DCMI, e assim como esses elementos previamente estabelecidos, são passíveis de serem qualificados por propriedades de refinamento (ou qualificadores).

Os qualificadores restringem o significado das propriedades, tornando-o mais específico. A propriedade *data*, por exemplo, pode ser refinada pelos qualificadores *dateAccepted*, *dateCopyrighted* ou *valid*, delimitando o escopo do valor da propriedade usada na descrição do recurso à data de aceitação (no caso de uma dissertação, por exemplo), à data em que seus direitos autorais foram instituídos ou a sua data de validade, respectivamente.

Apesar da designação dos rótulos de termos refinadores como *available*, *alternative*, *valid* e *modified* conotarem o uso em conjunto com as propriedades por eles refinados, os Princípios Gramaticais e Modelo Abstrato Dublin Core ratificam seu emprego de modo independente, sendo esses qualificadores comumente re-rotulados como “Date Created” e “Alternative Title” ao invés de simplesmente “Created” e “Alternative”, (BAKER, 2008), conforme designação original dos rótulos.

Algumas representações sintáticas de registros que utilizam propriedades Dublin Core, no entanto, combinam propriedades e qualificadores usando pontuação para conectá-los, como por exemplo com expressões como *date.valid* ou *title#alternative*, ao suporem a leitura desses registros por mecanismos de aplicações capazes de processar os dados dessa maneira – o que também conferiria a possibilidade de decomposição da forma aglutinada.

Essa decomposição, que no padrão Dublin Core é conhecida como princípio *Dumb-down*, serviria ao caso em que a qualificação da propriedade não fosse benéfica para fins de interoperabilidade, ou seja, ao caso em que essa qualificação causasse incompatibilidade na estrutura dos registros trocados entre bases de dados.

É importante frisar, no entanto, que a correta interpretação dos dados com o uso do princípio *Dumb-down* deve ser possibilitada pelas aplicações que processam os dados, não estando relacionada *à priori* com a sintaxe dos registros. É por essa razão que diferentes formas de expressar essa sintaxe foram desenvolvidas por empreendimentos distintos, inclusive aqueles que optaram por utilizar os termos de qualificação de forma independente, rotulando-os de maneira a tornar seu significado claro para o um suposto leitor humano.

Isso traz à tona a questão de que, se por um lado a pormenorização na descrição de recursos possibilita resultados mais objetivos ao recuperá-los, por outro lado pode dificultar a interoperabilidade entre sistemas de informação com níveis de detalhamento variados na descrição de seus recursos. Mesmo nos casos em que esse problema possa ser contornado por implementações que tirem proveito do princípio *Dumb-down*, isso não deixará de causar a perda de especificidade da informação que seja transferida de um sistema em que os dados sejam qualificados, para outro cujas propriedades correspondentes apresentem um sentido mais amplo.

#### Esquemas de codificação vocabular

Esquemas de codificação vocabular identificam vocabulários controlados, como classificações, cabeçalhos de assunto, tesouros e taxonomias, cujos termos são usados como valores que se associam às propriedades na descrição de um recurso. Um termo escolhido com este objetivo auxilia na interpretação semântica dos dados que descrevem um recurso.

A vantagem do uso de uma lista predefinida de termos dessa natureza consiste no maior grau de interoperabilidade semântica que propicia – tanto maior quanto mais amplamente o vocabulário for conhecido e adotado entre sistemas.

Chances são de que um vocabulário controlado seja mais conhecido se mantido por uma agência autorizada, ou seja, uma instituição que possui o reconhecimento social de sua expertise em relação ao universo de discurso dos termos constantes no vocabulário, e capacidade gerencial e operacional que garanta a continuidade e regularidade desse serviço. A disponibilização do vocabulário para domínio público é outra medida que favorece seu uso em larga escala.

Em termos técnicos, “um valor expresso usando um esquema de codificação será então um *token* selecionado de um vocabulário controlado” (HILLMANN, 2005). Mesmo que uma aplicação não esteja apta a realizar o processamento do vocábulo por meio de um código que o referencie (*token*), ainda será útil para interpretação humana, e nesse caso as vantagens obteníveis pelo uso vocabulários controlados mencionadas anteriormente continuam válidas.

Exemplos de esquemas de codificação vocabular recomendados pelo DCMI são o TGN (*Getty Thesaurus of Geographic Names*), que provê uma nomenclatura de locais geográficos e o LCSH (*Library Congress Subject Headings*), desenvolvido e mantido pela Biblioteca do Congresso Norte-americano, cujos termos são usados para refinar o elemento “subject” (assunto) do padrão Dublin Core.

Assim como o TGN e o LCSH, os esquemas de codificação vocabular constantes na relação de Termos Dublin Core são desenvolvidos e mantidos por instituições fora do âmbito do DCMI, com exceção do DCMIType.

O DCMIType (*DCMI Type Vocabulary*), ou Vocabulário de Tipos do DCMI, provê uma lista de termos interdisciplinares que podem ser usadas como valores do elemento Dublin Core *type* (tipo), com a função de atribuir um gênero ao recurso descrito.

Os termos que compõem o DCMIType (DCMI USAGE BOARD, 2010b) estão relacionados no quadro 4.

<b>Vocabulário de Tipos do DCMI</b>	
Coleção ( <i>Collection</i> )	Objeto físico ( <i>Physical Object</i> )
Conjunto de Dados ( <i>Dataset</i> )	Serviço ( <i>Service</i> )
Evento ( <i>Event</i> )	Software ( <i>Software</i> )
Imagem ( <i>Image</i> )	Som ( <i>Sound</i> )
Recurso Interativo ( <i>Interactive Resource</i> )	Imagem estática ( <i>Still Image</i> )
Imagem móvel ( <i>Moving Image</i> )	Texto ( <i>Text</i> )

**Quadro 4** – Termos do Vocabulário de Tipos do DCMI (*DCMIType*)

### Esquemas de codificação sintática

Esquemas de Codificação Sintática (ECS) apresentam regras que especificam como um valor deve ser estruturado, ou seja, o ECS define a sintaxe dos valores que se associam às propriedades na descrição de um recurso.

Esquemas de codificação sintática, ou *datatypes*, apresentam conjuntos de valores constituídos de cadeias de caracteres dispostos de acordo com uma notação formal. Trata-se de uma representação de valor mais distanciada da linguagem natural, expressando seu conteúdo informativo de forma estruturada.

A opção pelo uso de esquemas de codificação sintática na atribuição de valores a um determinado elemento de metadados garante os mesmos benefícios relacionados a aplicação de vocabulários controlados, ou seja, o grau de interoperabilidade que proporciona é tanto maior quanto mais adotado seja o formato.

Porém, ao contrário dos Esquemas de Codificação Vocabular, a semântica para o entendimento humano não é o foco, mas sim a sintaxe para o entendimento da máquina; ou seja, a estruturação do dado num formato padrão visa promover seu processamento automatizado: a limitação na forma de representar uma informação por meio dessa estrutura simplifica a correta interpretação da informação pelo software que realiza sua leitura, uma vez que irregularidades típicas da linguagem natural são evitadas.

Assim como no caso dos vocabulários controlados, agências como a ISO e a RFC<sup>23</sup> podem tomar a responsabilidade pela padronização de códigos, formatos e procedimentos. No entanto, nos casos referentes às tecnologias de comunicação (inclusive), o uso espontâneo pode ser fator que mais pesa para efeito de suas consolidações, em contraposição ao estabelecimento formal desses tipos de padrão.

(...) todas as especificações de sucesso são padrões de fato, desde que existe pequena influência que seus desenvolvedores tem sobre os adotantes. Os padrões de direito podem falhar e falham em ganhar popularidade. Assim, é muito mais proveitoso considerar essas tecnologias por seus méritos que por sua origem. (CROCKER, 1993)

Um exemplo de Esquemas de Codificação Sintática é o *W3C Date time*. Trata-se de um subconjunto dos formatos para data e horário determinados pela norma ISO 8601, definido com o objetivo de simplificar e promover seu uso na Web.

O *W3C Date time* oferece 6 níveis de granularidade, como segue:

- Ano: AAAA  
Ex. 1997
- Ano e mês: AAAA-MM  
Ex. 1997-07
- Data completa: AAAA-MM-DD  
Ex.: 1997-07-16
- Data completa, horas e minutos: AAAA-MM-DDThh:mmTZD  
Ex. 1997-07-16T19:20+01:00
- Data completa, horas, minutos e segundos: AAAA-MM-DDThh:mm:ssTZD  
Ex. 1997-07-16T19:20:30+01:00
- Data completa, horas, minutos, segundos e fração decimal de um segundo:  
AAAA-MM-DDThh:mm:ss.sTZD. Ex. 1997-07-16T19:20:30.45+01:00

Onde:

AAAA equivale a quatro dígitos representando o ano;  
MM equivale a dois dígitos para representação do mês, (de 01 a 12, em que o mês 01 equivale a janeiro)  
DD equivale a dois dígitos representando o dia do mês

---

<sup>23</sup> RFC (Request for Comments) é um documento publicado pelo IETF (Internet Engineering Task Force) que descreve métodos, formatos e outros tipos de solução referentes a forma de tratar a informação na internet, a fim de possibilitar a discussão acerca dessas tendências favorecendo a definição de padrões.



*hh* equivale a dois dígitos representando as horas (de 00 a 23)  
*mm* equivale a dois dígitos representando os minutos (de 00 a 59)  
*ss* equivale a dois dígitos representando os segundos (de 00 a 59)  
*s* equivale a um dígito representando uma fração decimal de segundo  
*TZD* designa o fuso horário (*Z* ou *+hh:mm* or *-hh:mm*)

Dois métodos são utilizados para definir a data conforme o fuso horário: o primeiro, pela expressão do horário local, mais a indicação em horas e minutos do deslocamento de tempo em relação ao UTC (*Coordinated Universal Time*) – ou seja, em relação ao horário vigente no meridiano de Greenwich (Reino Unido), onde por convenção se inicia o cálculo de distâncias em longitudes do globo terrestre, permitindo se estabelecer os fusos horários; o segundo, pela expressão do horário vigente no meridiano de Greenwich, ou seja, o próprio UTC, indicado pelo uso da letra *Z* (*zone*) ao final do formato de data e horário.

1994-11-05T08:15:30-05:00

1994-11-05T13:15:30Z

Os dois exemplos acima representam a expressão esquemática do dia 5 de novembro de 1994, às 8 horas, 15 minutos e 30 segundos da manhã na costa leste da América do Norte, conforme os dois métodos apresentados, respectivamente.

### Classes

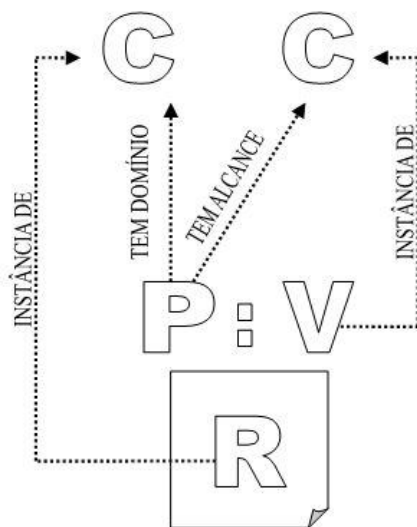
Classes são grupos de recursos com certas características em comum, sendo assim consideradas membros de um conceito (DCMI Media Wiki, 2011). Uma coleção de classes é geralmente definida para um propósito ou domínio.

Classes são organizadas hierarquicamente, podendo ser estendida por meio de subclasses. Assim, ainda que duas ou mais classes sejam criadas em momentos distintos, com propósitos distintos, se seus membros forem relacionados, elas podem eventualmente ser “conectadas”, via refinamentos de classe. No caso do RDF, essa extensibilidade permite que processadores de metadados percorram instâncias de classes para além daquelas previstas para uma determinada aplicação, performando “ações sobre metadados que eles não foram projetados para processar” (W3C, 1999) – desempenho que está ligada aos mecanismos de inferências lógicas possibilitadas por essa linguagem.

A noção de classe no âmbito do padrão Dublin Core foi admitida em 2008, com a inserção de termos Dublin Core do tipo *Classe*. Membros do vocabulário DCMIType são consideradas metadados do tipo *Classe* entre os termos DCMI. Classes inéditas referem-se ao domínio (*domain*) ou alcance (*range*) associados a alguma propriedade.

É dito que o conjunto de todos os recursos que são descritos por determinada propriedade, pertencem ao domínio associado a essa propriedade. Analogamente, o conjunto dos valores que são atribuídos a recursos descritos via determinada propriedade, pertencem ao alcance associado a essa propriedade.

A figura 3 abaixo expressa as relações entre uma propriedade (P) do recurso descrito (R), o valor (V) atribuído ao recurso por meio da propriedade, e as respectivas classes (C) às quais R e V pertencem devido às relações “tem domínio” e “tem alcance” manifestadas pela propriedade.



**Figura 3** – Diagrama das possíveis relações “tem alcance” e “tem domínio” de uma propriedade

Imputar domínios e alcances a propriedades é um artifício que possibilita o processamento automatizado de informações implícitas em definições da linguagem natural.

## 2.4 Atributos de termos Dublin Core

A documentação do padrão Dublin Core apresenta uma relação de atributos referente a cada um de seus termos. As características de um termo determinam a presença ou ausência de certas propriedades, mas de um modo geral todos os termos apresentam o seguinte conjunto mínimo de atributos:

<b>Nome</b>	Uma unidade designativa do termo.
<b>Rótulo</b>	Um rótulo atribuído ao termo para facilitar sua identificação por leitores humanos.
<b>URI</b>	<i>Uniform Resource Identifier</i> composto pelo nome do termo pós-fixado ao URI do <i>namespace</i> em que o termo é definido, gerando um identificador único do termo.
<b>Definição</b>	Declaração que representa o conceito e natureza essencial do termo.
<b>Tipo do termo</b>	O Tipo do termo conforme descrito no Modelo Abstrato Dublin Core.

Onde aplicável, os seguintes atributos provêm informação adicional sobre o termo:

<b>Comentário</b>	Informação adicional sobre o termo ou sua aplicação.
<b>Veja</b>	Documentação oficial relacionada ao termo.
<b>Refina</b>	Uma Propriedade cujo termo descrito é uma subpropriedade.
<b>Mais amplo que</b>	Uma Classe da qual o termo descrito é uma Superclasse.
<b>Mais estreito que</b>	Uma Classe da qual o termo descrito é uma Subclasse.
<b>Tem domínio</b>	Uma Classe da qual o recurso descrito é uma instância.
<b>Tem alcance</b>	Uma Classe da qual um valor descrito pelo termo é uma instância.
<b>Membro de</b>	Um conjunto enumerado de recursos (Esquema de Codificação Vocabular) do qual o termo é membro.
<b>Versão</b>	Designação dada ao conjunto características do termo, que podem variar sutilmente durante seu desenvolvimento.
<b>Propriedade equivalente</b>	Uma Propriedade à qual o termo descrito é equivalente.

Ao se pautar por esses atributos na definição de seus termos, um grau maior de consistência interna e em relação a outras comunidades de metadados é proporcionado.

## 2.5 Namespaces Dublin Core

Com a inclusão dos termos referentes ao Dublin Core Qualificado, a Iniciativa de Metadados Dublin Core passa a responder por novos *namespaces*, conforme visto no quadro 5. O *namespace* original, contendo os quinze elementos precursores, foi mantido como definido em 1999, com o objetivo de não afetar aplicações que foram nele baseadas.

Implementadores podem escolher livremente usar essas quinze propriedades tanto de sua variante legada dc: (ex., <http://purl.org/dc/elements/1.1/creator>) quanto à variante em dcterms: (ex., <http://purl.org/dc/terms/creator>) a depender dos requerimentos da aplicação. (DCMI, 2010)

Namespace	Conteúdo
<a href="http://purl.org/dc/elements/1.1/">http://purl.org/dc/elements/1.1/</a>	Conjunto de elementos Dublin Core (versão 1.1, referente aos 15 elementos cerne)
<a href="http://purl.org/dc/terms/">http://purl.org/dc/terms/</a>	Todas as propriedades DCMI, classes e esquemas de codificação (exceto os indicados abaixo)
<a href="http://purl.org/dc/dcmitype/">http://purl.org/dc/dcmitype/</a>	Classes do Vocabulário de Tipos do DCMI
<a href="http://purl.org/dc/dcam/">http://purl.org/dc/dcam/</a>	Termos usados no Modelo Abstrato do DCMI

**Quadro 5** – *Namespaces* sob responsabilidade da Iniciativa de Metadados Dublin Core (POWELL et al., 2007b)

O quarto dos *namespaces* referidos no quadro acima, refere-se aos termos criados para serem usados na modelagem de dados segundo o modelo abstrato proposto pelo DCMI.

## 2.6 Modelo Abstrato Dublin Core

Modelos são representações de algo. Eles servem para tornar algo conhecido e compreendido de acordo com um propósito particular e, portanto, focam somente aspectos

importantes para aquele propósito. Assim, modelos são sempre uma versão limitada da coisa representada.

Um modelo de dados define o significado e a estrutura dos dados que representa, conforme a finalidade que o modelo pretende apoiar. Para tanto, entidades, i.e. “categorias” ou “tipos” de objetos pertencentes ao escopo do modelo de dados são usados ao invés de dados em si.

A definição de entidades para um modelo de dados se dá por meio de uma abstração. Nesse sentido, abstração é a desconsideração de detalhes de algo, para ampliar a visão sobre um fenômeno ou situação, estabelecendo generalizações. Desse modo, podemos “observar similaridades que não eram aparentes ou mesmo não existiam anteriormente”. (HOBBERMAN, 2009:48)

Assim, um modelo de dados é um modelo abstrato que determina a semântica de entidades e as relações entre elas, dentro de um domínio de interesse. No estabelecimento dessa semântica, as fronteiras conceituais delimitadas pelo domínio de interesse do modelo de dados devem ser observadas, a fim de manter sua integridade e, por conseguinte, sua funcionalidade.

Chances são que abstrações na modelagem de dados de uma esfera de interesse guardem um significativo grau de semelhança com esferas de interesse afins. É com base nessa generalização conceitual que é possibilitada a interoperabilidade entre sistemas construídos originalmente por meio de formatos de dados diferentes.

Portanto, a relevância dos modelos abstratos no compartilhamento de dados se dá não apenas internamente, entre aplicações que se baseiam em um modelo de dados comum, mas também entre aquelas que utilizam diferentes especificações de metadados. Em sua tese sobre harmonização de metadados, Nilsson (2010) defende que a troca de informação entre sistemas criados a partir de diferentes padrões é tão mais propícia quanto os modelos abstratos desses padrões se assemelhem.

O favorecimento máximo de interoperabilidade por meio de um modelo de dados acontece quando estruturas de metadados são derivadas de um mesmo modelo conceitual, que

é “o nível mais importante de modelo para lidar com os problemas de compartilhamento e troca de dados”. (WEST, 2003:10) Dessa maneira, as entidades representadas pelo modelo são comuns aos perfis das aplicações nele baseadas, variando as instâncias dessas entidades – os dados compartilhados.

Modelos de dados apóiam dados e sistemas de computador provendo a definição e formato de dados. Se isso é feito consistentemente entre sistemas, então compatibilidade de dados pode ser alcançada. Se as mesmas estruturas de dados são usadas para armazenar e acessar dados, então diferentes aplicações podem compartilhar dados. (WEST, 2003:7)

O termo metamodelo pode ser usado para designar um modelo que descreve outros modelos. (ISO, 2003) Um metamodelo também pode ser usado para comparar esquemas de metadados preconcebidos com o uso de diferentes formatos de metadados, configurando-se como um modelo comum ao qual esses esquemas sejam “traduzidos”, preservando os conceitos apresentados pelos conjuntos de metadados comparados.

O Modelo Abstrado Dublin Core, ou DCAM (*Dublin Core Abstract Model*), publicado como uma recomendação pelo DCMI em março de 2005, foi idealizado para servir como um metamodelo. Ele especifica os componentes e constructos usados nos metadados Dublin Core (POWELL, 2007a), mostrando como esses componentes são combinados com a finalidade de estruturar a informação.

Por tratar-se de um modelo de dados destinado a um uso interdisciplinar e visando beneficiar o maior número de produtores de informação, o DCAM é independente de qualquer codificação de sintaxe.

Esse modelo de dados é voltado a desenvolvedores de perfis de aplicação que fazem uso de metadados Dublin Core, bem como de outros vocabulários compatíveis, a fim de trazer benefícios para essas realizações.

Tal modelo de informação nos permite obter um melhor entendimento dos tipos de descrição que estamos codificando e facilita o desenvolvimento de melhores mapeamentos e traduções entre sintaxes. (POWELL, 2007a)

O Modelo Abstrato Dublin Core tem sido construído sobre o trabalho realizado pelo W3C referente ao modelo de dados RDF (*Resource Description Framework*), refletindo uma tendência que diz respeito a novas possibilidades na organização e tratamento da informação favorecida por avanços tecnológicos na área da computação:

Nos últimos anos, alguns profissionais da informação – particularmente aqueles fora da comunidade bibliotecária – começaram a mudar seu conceito de metadados. Historicamente, registros – e não declarações sobre os recursos que eles agregam e empacotam – tem sido tratados como componentes centrais de metadados. (HARPER, 2010:22)

Embora essa visão dos metadados continue sendo essencial para fins de transporte dos registros entre sistemas, um novo paradigma de compartilhamento da informação em rede focaliza sua atenção em unidades informativas menores, “tirando proveito da infraestrutura aberta da web” (HARPER, 2010:24) ao expandir seu uso além de um conjunto dados definido para aplicações isoladas.

A documentação do Modelo Abstrato Dublin Core é formado por três modelos que expõem os componentes utilizados e suas relações em três níveis: o nível do recurso, da descrição de recursos e dos termos (ou vocabulário).

Os gráficos apresentados fazem uso do diagramas de classe UML. Neles, cada quadro representa um componente do Modelo Abstrato DCMI, e cada tipo de seta e as expressões de multiplicidade apresentam as seguintes denotações:

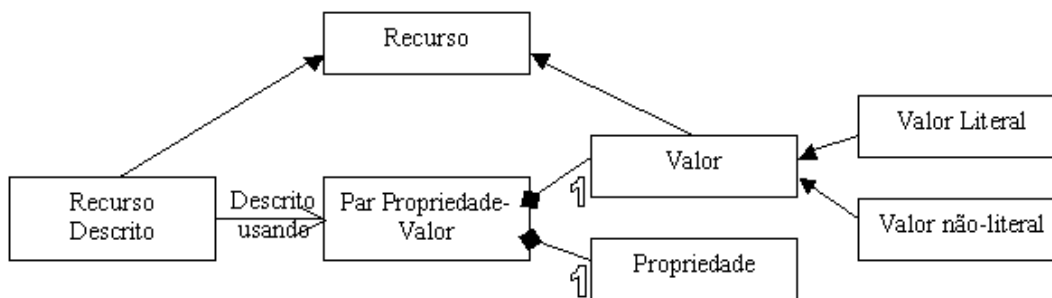
- Relação de generalização “é um (a)”.  
Exemplo: “um *valor* é um *recurso*”
- ◆ Relação de agregação “contém um(a)”.  
Exemplo: “uma *descrição* contém uma ou mais *declarações*”
- Representa a relação definida textualmente no gráfico.  
Exemplo: “um *recurso* pode ser instância de uma ou mais *classes*”.

- 1 Associação obrigatória de um único componente do tipo indicado.  
Exemplo: “uma *declaração* contém um URI *da propriedade*”
- 1.. n Associação obrigatória de pelo menos um componente do tipo indicado.  
Exemplo: “um *conjunto descritivo* contém uma ou mais *descrições*”
- 0.. n Associação opcional de um ou vários componentes do tipo indicado.  
Exemplo: “uma *propriedade* pode ser uma subpropriedade de uma ou mais *propriedades*”
- 0.. 1 Associação opcional de até um componente do tipo indicado.  
Exemplo: “um *valor string simples* pode ter até um *valor string de linguagem*”

As definições semânticas dos componentes do Modelo Abstrato estão nos anexos do trabalho.

### 2.6.1 Modelo de Recurso

O Modelo de Recurso DCMI configura-se conforme a figura a seguir.



**Figura 4** – Modelo de Recurso DCMI, delineado em Powell (2007a)

As relações entre os componentes do Modelo de Recurso presentes na figura 4 podem ser compreendidas da seguinte maneira:

- A descrição de um recurso é feita por um ou mais pares propriedade-valor.



- Cada par propriedade-valor é composta de uma propriedade e um valor.
- Um valor é a representação de uma entidade física, digital ou conceitual, ou ainda um literal, associados a uma propriedade na descrição de um recurso. Um valor também é considerado um recurso em si mesmo.
- Um literal (ou string) é uma entidade expressa por uma cadeia de caracteres Unicode para denotar um recurso. Opcionalmente, pode ser acompanhado de uma indicação do idioma em que o literal foi expresso.
- Uma propriedade representa aspecto pelo qual o recurso é descrito
- Um valor pode ser do tipo um valor literal ou um valor não-literal.
- Um valor literal é um valor que é um literal.
- Um valor não-literal é um valor que é uma entidade física, digital ou conceitual.

### **2.6.2 Modelo de Conjunto Descritivo**

O Modelo de Conjunto Descritivo DCMI configura-se conforme a figura a seguir.

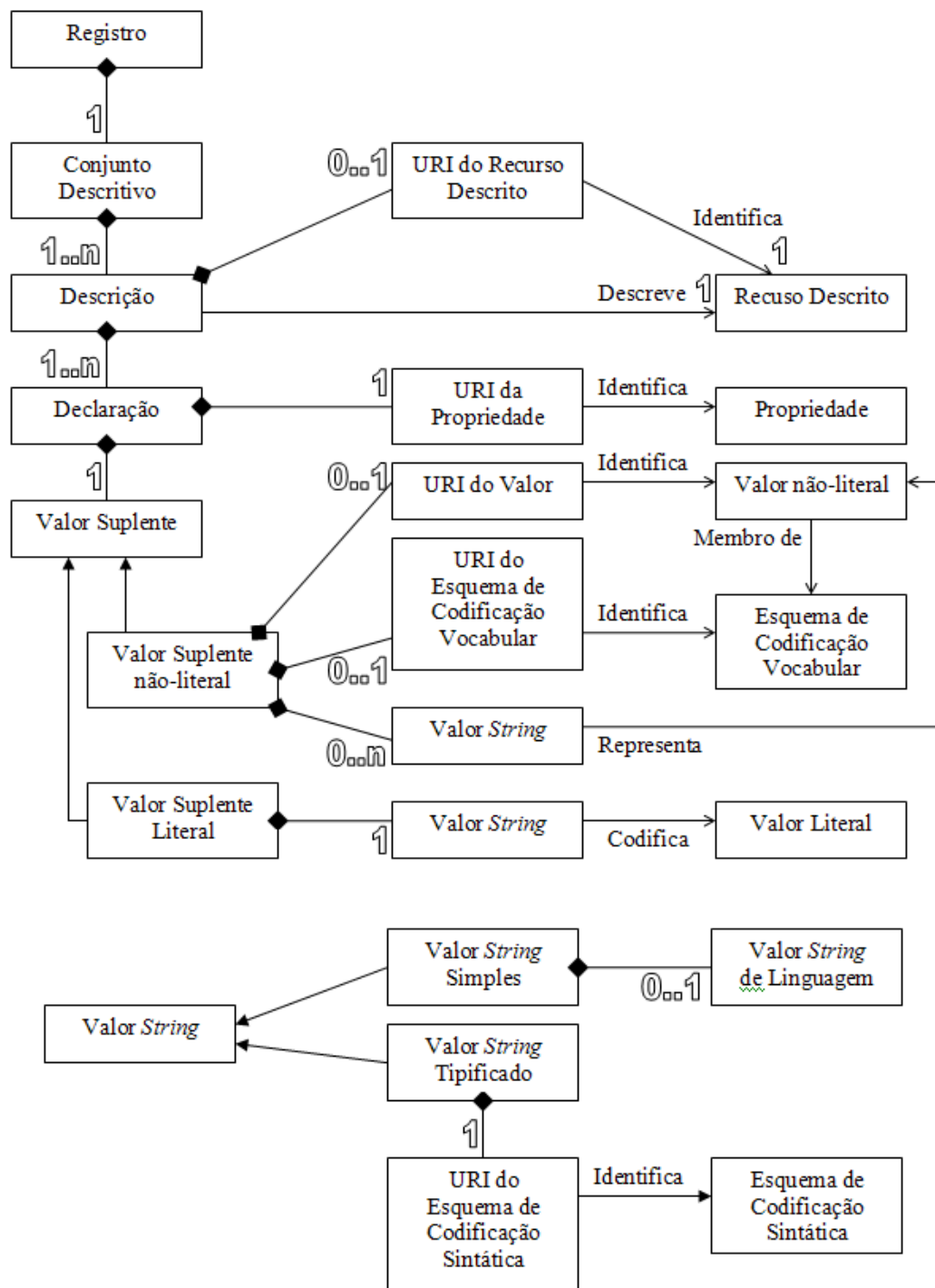


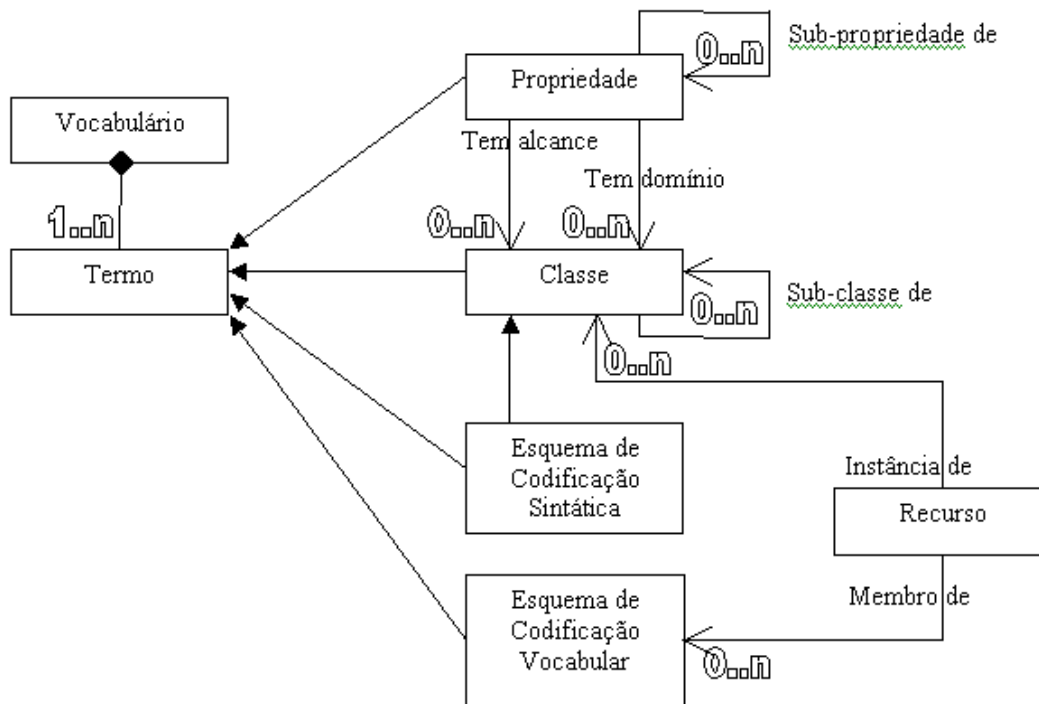
Figura 5 – Modelo de Conjunto Descritivo DCMII em Powell (2007a)

As relações entre os componentes do Modelo de Conjunto Descritivo presentes na figura 5 podem ser compreendidas da seguinte maneira:

- Um registro é composto por um conjunto descritivo.
- Um conjunto descritivo é o agrupamento de uma ou mais descrições.
- Cada descrição é feita de uma ou mais declarações que se referem a um único recurso descrito que, por sua vez, pode ser identificado por um URI do recurso descrito.
- Cada declaração é composta por um URI da propriedade que identifica uma propriedade e também por um valor suplente (*value surrogate*).
- Um valor suplente pode ser um valor suplente literal ou um valor suplente não-literal.
- Um valor suplente literal é composto por um valor string, ou seja, uma sequência de caracteres Unicode, que codifica um valor literal.
- Um valor suplente não-literal pode ser composto de uma única URI de valor que identifica valor não-literal. Esse valor não-literal é membro de um esquema de codificação vocabular, que também pode ser identificado por uma única URI do esquema de codificação vocabular (que também pode fazer parte do valor suplente não-literal). O valor suplente não-literal também pode ser composto por várias unidades de valor string, que neste caso representam o único valor não-literal permitido numa declaração.
- Um valor string pode ser um valor string simples ou um valor string tipificado:
- Um valor string simples é um valor destinado à leitura humana, como uma palavra, um ou numeral ou um código. Opcionalmente, pode ser acompanhado de uma (única) indicação do idioma (valor string de linguagem) em que esse valor foi expresso.
- Um valor string tipificado é qualificado por um esquema de codificação sintática via um URI do esquema de codificação sintática.

### **2.6.3 Modelo de Vocabulário**

O Modelo de Vocabulário DCMI configura-se conforme a figura a seguir.



**Figura 6** – Modelo de Vocabulário DCMI em Powell (2007a)

As relações entre os componentes do Modelo de Vocabulário presentes na figura 6 podem ser compreendidas da seguinte maneira:

- Um vocabulário é um conjunto de um ou mais termos. Cada termo é um membro de um ou mais vocabulários.
- Um termo é uma propriedade (elemento), classe, esquema de codificação vocabular ou esquema de codificação sintática.
- Cada propriedade pode ser relacionada a uma ou mais classes por uma relação tem domínio. Onde é declarado que uma propriedade tem tal relacionamento com uma classe e a propriedade é parte de um par propriedade/valor, segue-se que o recurso descrito é uma instância dessa classe.
- Cada propriedade pode ser relacionada a uma ou mais classes por uma relação tem alcance. Onde é declarado que uma propriedade tem tal relacionamento com uma classe e a propriedade é parte de um par propriedade/valor, segue-se que o valor é uma instância dessa classe.
- Cada recurso pode ser uma instância de uma ou mais classes.

- Cada recurso pode ser um membro de um ou mais esquemas de codificação vocabular.
- Cada classe pode ser relacionada a uma ou mais classes por uma relação do tipo subclasse de (onde as duas classes são definidas de tal forma que todos os recursos que são instâncias de uma subclasse são também instâncias da classe relacionada).
- Cada propriedade pode estar relacionada a uma ou mais propriedades por uma relação do tipo *subpropriedade de*. Onde é declarado que tal relacionamento existe, as duas propriedades são definidas de tal forma que seja onde for que uma subpropriedade seja parte de um par propriedade/valor descrevendo um recurso, segue-se que o recurso é também descrito usando um segundo par propriedade/valor feito da propriedade e do valor.
- Cada esquema de codificação sintática é uma classe (de literais).

## 2.7 O Perfil de Aplicação Dublin Core

Sistemas de Informação que apresentam necessidades muitas vezes singulares, mesmo em comparação com sistemas de objetivos e escopo similares. Nesses casos, a busca por um modelo único (universal) para representação descritiva de seus recursos tende a não satisfazer suas peculiaridades.

A utilização de um mesmo padrão de metadados e um mesmo modelo abstrato pode não ser medida suficiente para garantir a interoperabilidade entre sistemas. Para se alcançar uma interoperabilidade eficiente entre bases de dados, “acordos” estruturais mais concretos devem ser estabelecidos.

A definição de perfis de aplicações utilizando metadados provenientes de um ou mais padrões ou esquemas de metadados – aqueles elementos que melhor correspondam às necessidades a serem supridas – pode ser uma solução que combina o atendimento das demandas locais e a garantia de um maior ou menor grau de interoperabilidade com bases de dados afins, de acordo com características do perfil de aplicação. A adoção de padrões de metadados de ampla aceitação é um fator que aumenta as chances de sucesso de interoperabilidade entre sistemas.

Assim, sob a ótica do DCMI, um Perfil de Aplicação Dublin Core (DCAP) é “uma declaração de quais termos de metadados uma organização, recurso de informação, aplicação ou comunidade usuária usa em seus metadados”. (BAKER, 2003)

Visando estabelecer orientações que guiassem implementadores de metadados na construção de perfis de aplicação eficientes – ou seja, que favorecessem práticas comuns de metadados dentro de contextos comuns de uso –, o DCMI estabeleceu um conjunto de componentes para sua composição.

Segundo a Iniciativa de Metadados Dublin Core, um perfil deve apresentar:

*Requerimentos funcionais:* definição do escopo e dos objetivos do SI, e a descrição do que essa comunidade deseja realizar com sua aplicação;

*Modelo de domínio:* uma representação abstrata dos tipos de coisas a serem descritas no perfil e as relações entre elas;

*Perfil do conjunto de descritivo:* levantamento dos elementos de metadados a serem empregados na descrição dos recursos e sob quais condições; O registro de metadados resultante desse conjunto de propriedades representa uma instância válida de um perfil de aplicação;

*Diretrizes de uso:* as regras para uso do perfil de aplicação por aqueles que o implementarão;

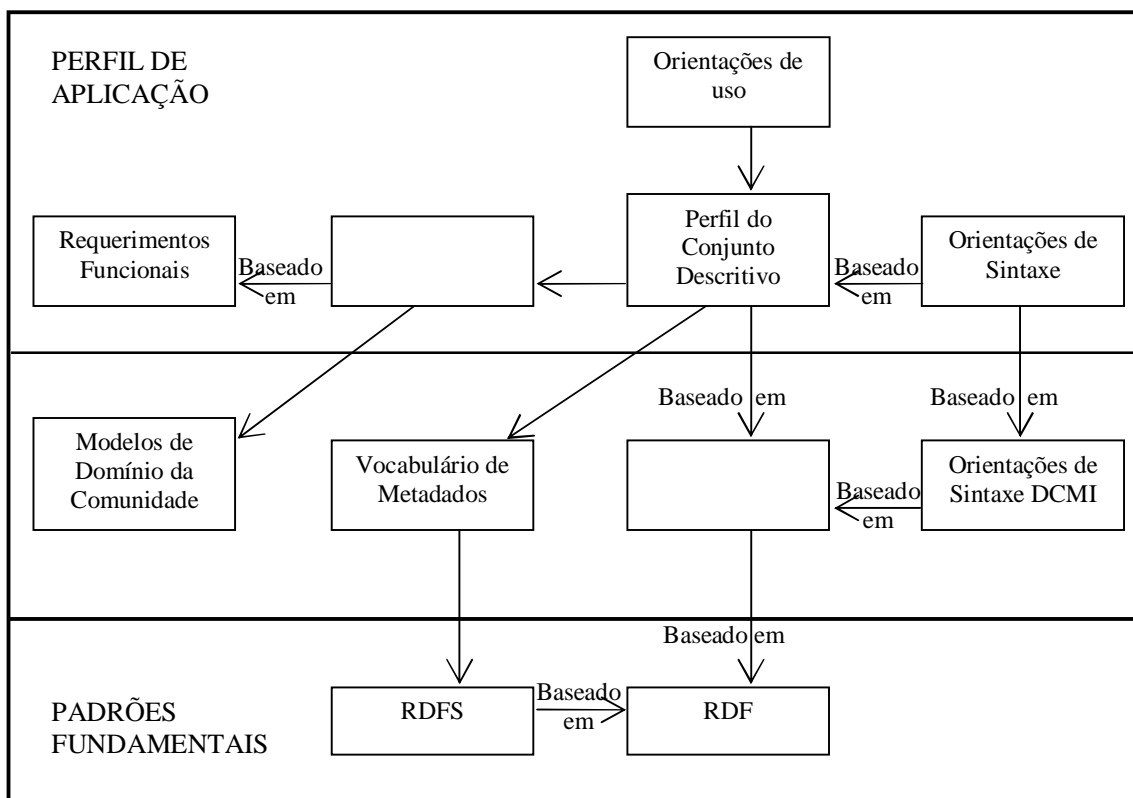
*Orientação para sintaxe e formatos de dados:* determinação da sintaxe para codificação do conteúdo dos metadados e de seus formatos. A sintaxe de codificação e o formato dos dados devem ser escolhidos visando possibilitar análise e manipulação automatizada dos dados.

### **2.7.1 Framework Singapura**

O modo como os componentes de um perfil de aplicação se interrelacionam é dado por uma estrutura para DCAPs, que foi apresentada em setembro de 2007, na Conferência

Internacional do Dublin Core e Aplicações de Metadados em Singapura, tendo assim ficado conhecida como Framework Singapura (figura 7).

Essa estrutura apresenta três níveis hierárquicos de especificações e as relações entre elas, de modo a evidenciar os componentes que participam na criação de um perfil de aplicação conformante com os princípios do DCMI, tais como boas práticas de documentação, a promoção do reuso de vocabulários de metadados existentes e o suporte para o ambiente de *Linked Data*.



**Figura 7** – Framework Singapura delineado por Nilsson, Baker e Johnston (2008)

Abaixo relacionamos cada uma das camadas e suas partes integrantes:

#### Nível do perfil de aplicação

No nível do perfil de aplicação estão os componentes que serão, de fato, desenvolvidos pelo implementador do esquema de metadados.

Como já pontuados na seção anterior, compõem esse nível:

- Orientações de uso do perfil de aplicação
- Requerimentos funcionais
- Modelo de domínio
- Conjunto Descritivo do Perfil
- Orientações de sintaxe e formato de data

#### Nível dos padrões de domínio

O nível dos domínios é composto pelos padrões e orientações que devem ser considerados durante a criação do perfil de aplicação. Compõem esse nível:

Modelos de Domínio da Comunidade: ao projetar o modelo de domínio do perfil de aplicação, o implementador deve se basear em modelos de domínio mais gerais prevalentes na comunidade de uso à qual seu cliente se insere. Essa é uma forma de garantir que o perfil de aplicação está de acordo com as práticas gerais dessa “comunidade estendida”, o que pode garantir a interoperabilidade entre diferentes aplicações que atuem nesse mesmo contexto.

Vocabulários de metadados: o Conjunto Descritivo do perfil faz uso de vocabulários formalmente definidos e mantidos por autoridades de agências competentes. Como exemplo desse tipo de vocabulário podemos citar o Conjunto de Elementos de Metadados Dublin Core (DCMES), *MARC Code List for Relators*, um vocabulário de funções de responsabilidade na descrição de recursos e *Library of Congress Subject Headings* (LCSH), uma lista de cabeçalhos de assunto, sendo os dois últimos mantidos pela Biblioteca do Congresso estadunidense.



Modelo Abstrato DCMI: as restrições estruturais do conjunto de descrições de um recurso, além de se pautar pelo modelo de domínio delineado para o perfil de aplicação, deve apresentar correspondência com as entidades e constructos definidos no Modelo Abstrato Dublin Core. A necessidade dessa correspondência é mais uma medida para promover a interoperabilidade entre perfis de aplicação ao se pautarem por um modelo de dados genérico comum.

Orientações de Sintaxe do DCMI: o Modelo Abstrato Dublin Core é independente de sintaxes particulares, embora a sintaxe usada para codificação de um perfil de aplicação deva ser compatível com esse modelo – incluindo-se o requerido uso de URIs para representação de suas propriedades. No entanto, o DCMI fornece orientações para codificação nas linguagens mais comumente usadas na web: XHTML, XML e RDF/XML.

#### Nível dos padrões fundamentais

Com o avanço de seus trabalhos, o DCMI tomou como princípio o desenvolvimento de ferramentas e promoção de práticas que possibilitem a troca de dados em ambiente de *Linked Data*. Essa ampliação do escopo da Iniciativa associa-se à união de esforços com a comunidade desenvolvedora do *Resource Description Framework* (RDF), que é a “fundação padrão para semântica processável por máquina de nosso tempo” (NILSSON; BAKER; JOHNSTON, 2008).

O Modelo Abstrato Dublin Core foi desenvolvido como uma “tradução” do modelo abstrato RDF, de forma a tornar mais simples sua compreensão pelos implementadores de metadados, uma vez que “a linguagem não familiar de modelação formal, a complexa documentação do RDF, e a dificuldade de sua representação em XML apresentava-se como um obstáculo para a adoção do RDF em larga escala.” (HARPER, 2010:23)

Dessa maneira, assim como um perfil de aplicação baseia-se no Modelo Abstrato Dublin Core, também se fundamenta nos padrões RDF e RDFS (*Resource Description Framework Schema*).

### Resource Description Framework (RDF)

Preceituada como recomendação pelo W3C em 1999, RDF é uma linguagem criada para descrição de recursos da web, mais precisamente em ambiente de *Linked Data* – em que se visa favorecer inferências automáticas ou, em outras palavras, o "entendimento" do significado dessa descrição pela máquina.

Descrições RDF constituem-se de uma estrutura simples composta de sujeito, predicado e objeto, conhecida como "triplo" (*RDF triple*): o sujeito equivale ao recurso em si (sendo representado por um URI), o predicado equivale à propriedade desse recurso (especificando qual de seus aspectos está sendo declarado) e o objeto pode ser uma informação textual sobre o recurso, ou outros recursos (nesse caso, também representados por URIs).

### Resource Description Framework Schema (RDFS)

Declarações em RDF descrevem recursos por meio da atribuição de sujeito, predicado e objeto, porém não possui meios para determinar o significado desses componentes, nem as relações entre eles. Para isso, RDFS (também conhecida como "Vocabulário RDF") introduz os conceitos de classe, subclasse e subpropriedade (permitindo descrições hierárquicas), além dos conceitos de domínio e alcance, usados para determinar a que classes pertencem sujeito e objeto, respectivamente.

## **2.8 Níveis de interoperabilidade do Dublin Core**

Os níveis de interoperabilidade Dublin Core foram estabelecidos para classificar projetos de metadados conforme seus graus de compatibilidade com o padrão. "A intenção é prover uma 'escala de interoperabilidade', especificando as escolhas, custos e benefícios envolvidos em desenhar aplicações para aumento dos níveis de interoperabilidade". (NILSSON, 2009)

Os quatro níveis apresentados se sobrepõem, de maneira que um nível superior engloba as especificações dos níveis inferiores. De forma resumida, as exigências progridem da seguinte maneira:

- 1) Compartilhamento de vocabulários definidos em linguagem natural.
- 2) Compartilhamento de vocabulários formais.
- 3) Compartilhamento de vocabulários formais e uso de modelo de dados.
- 4) Compartilhamento de vocabulários formais, uso de modelo de dados e das regras referentes ao preenchimento dos valores.

Abaixo, a explanação de cada um desses níveis se dá de forma mais detalhada:

### **2.8.1 Nível 1: Compartilhamento de definição de termos**

O primeiro nível de interoperabilidade se refere às aplicações que compartilham as mesmas definições, em linguagem natural, para os elementos de metadados. O uso de URIs não é requerido, sendo assim considerada *informal*. “Atualmente, a maior parte das aplicações de metadados operam dessa maneira” (DCMI, c1995-2011).

Um esquema de metadados é considerado informalmente interoperável com o padrão Dublin Core se for possível efetuar seu mapeamento dentro do conjunto dos quinze elementos cerne. Neste nível a conformidade com o Modelo Abstrato DCM não é requerida.

### **2.8.2 Nível 2: Interoperabilidade semântica formal**

O segundo nível de interoperabilidade é marcado pelo uso de URIs na identificação de metadados. Apesar de a conformidade com o Modelo Abstrato Dublin Core ainda não ser requerida, é necessário que as relações entre os termos e suas regras de uso sejam declaradas formalmente, possibilitando inferências (conclusões lógicas automáticas).

Essa “semântica formalizada” deve ser condizente com a representada em modelos de grafos RDF – ou seja, independentemente do formato de codificação original, um completo

mapeamento dos dados para triplos RDF deve ser possibilitado. As relações de domínio, alcance, e subpropriedade, referentes aos termos em uso, também devem ser condizentes com aquelas estabelecidas em RDF.

### **2.8.3 Nível 3: Conjunto Descritivo de interoperabilidade sintática**

Neste nível de interoperabilidade é requerida a conformidade com o Modelo Abstrato DCMI.

Enquanto cada grafo RDF declara um único fato a respeito de um recurso informacional, o Modelo Abstrato DCMI acrescenta as noções de *descrição* e *conjunto descritivo*: uma descrição é composta de uma ou mais declarações, um conjunto descritivo por uma ou mais descrições. Dessa forma, registros são formados de maneira identificável, permitindo suas trocas entre sistemas. Um registro bibliográfico, por exemplo, pode ser formado por declarações que descrevam um livro, e declarações que descrevam seu autor.

Para manter a conformidade com o Modelo Abstrato Dublin Core, deve-se observar também a distinção entre “identificação” e “representação” de um valor, bem como possibilitar essa representação por meio de vários “valores *string*”. Por exemplo, enquanto a identificação de uma imagem pode ser realizada por meio de uma URL que aponte para ela em ambiente Web, sua representação pode ser realizada por meio de legendas em diferentes idiomas.

### **2.8.4 Nível 4: Interoperabilidade de Perfil de Conjunto Descritivo**

Este nível de interoperabilidade é caracterizado pela presença de todos os componentes especificados no Framework Singapura para Perfis de Aplicação Dublin Core.

O Perfil de Conjunto Descritivo expresso em linguagem XML é de especial importância, pois formaliza as restrições estruturais que definem o Conjunto Descritivo do Perfil de Aplicação, de modo a torná-lo analisável por máquina, promovendo a interoperabilidade automaticamente.

### **3. Proposta de um Perfil de Aplicação para descrição de publicações em quadrinhos**

Propõe-se aqui a construção de um modelo para um catálogo de publicações em quadrinhos, seguindo a metodologia de construção de um Perfil de Aplicação conforme definido pela Iniciativa de Metadados Dublin Core, e fazendo uso dos elementos de dados cerne do padrão do Dublin Core.

A restrição às propriedades desse padrão de metadados confere ao modelo um caráter intencionalmente generalista, possibilitando a composição de um registro descritivo compatível com o grau mínimo de interoperabilidade exigido para conformidade com a Iniciativa de Arquivos Abertos ou OAI (*Open Archives Initiative*).

Tal conformidade é perseguida a fim de possibilitar que tal catálogo de publicações em quadrinhos tire máximo proveito da estrutura de armazenamento e coleta de metadados proporcionado pelo Protocolo para Coleta de Metadados do OAI, ou OAI-PMH (Protocol for Metadata Harvesting), uma vez que todo e qualquer Provedor de Serviço que pretenda fazer uso do protocolo deve reconhecer os elementos cerne do Padrão Dublin Core, estando apto, assim, a dispor dos registros do catálogo de publicações em quadrinhos em sua integridade – ou seja, sem ser forçado a ignorar dados associados a metadados desconhecidos.

Essa implementação sugerida agruparia dados dispersos de registros bibliográficos de publicações em quadrinhos, presentes em bases de dados independentes, constituindo o que a literatura chama de Sistema de Federado de Bases de Dados ou Sistema de Informação Federado.

Buscou-se traçar, também, um esquema de metadados para servir de modelo de dados intermediário entre o perfil de aplicação baseado nos elementos cerne Dublin Core e os modelos de dados originais das bases de dados participantes da federação, com o objetivo de auxiliar na tradução de um esquema para o outro.

### 3.1 Sistemas de Informação Federados

Em artigo que define uma arquitetura de referência para sistemas de gerenciamento de bases de dados distribuídas, Sheth e Larson conceituam um Sistema Federado de Bases de Dados como “uma coleção de sistemas de bases de dados cooperantes, que são autônomas e possivelmente heterogêneas.” (SHETH; LARSON, 1990:228)

Nesse trabalho, obtemos as seguintes definições:

Um *Sistema de Base de Dados* (SBD) consiste de um *software*, chamado de *Sistema de Gerenciamento de Base de Dados* e uma ou mais bases de dados gerenciadas.

*Sistema de Base de Dados Federado* (SBDF) é um sistema criado para prover operações em bases de dados gerenciadas por Sistemas de Bases de Dados autônomas e possivelmente heterogêneas.

Assim, para esses autores, “sistemas consistindo de múltiplos SBDs, dos quais SBDFs são um tipo específico, podem ser caracterizados ao longo de três dimensões ortogonais: distribuição, heterogeneidade e autonomia.” (SHETH; LARSON, 1990:185)

*Distribuição* diz respeito à multiplicidade de bases de dados do sistema, que podem ser co-aloçadas ou descentralizadas, tanto geograficamente quanto em termos de *hardware*. No entanto, para fazerem parte de um mesmo SBDF devem estar interconectadas por um sistema de comunicação.

A *heterogeneidade* se dá em razão de diferenças tecnológicas, seja relacionada a *hardware* ou *software*. Ela também pode ocorrer devido a diferenças em modelos de dados ou semântica de dados.

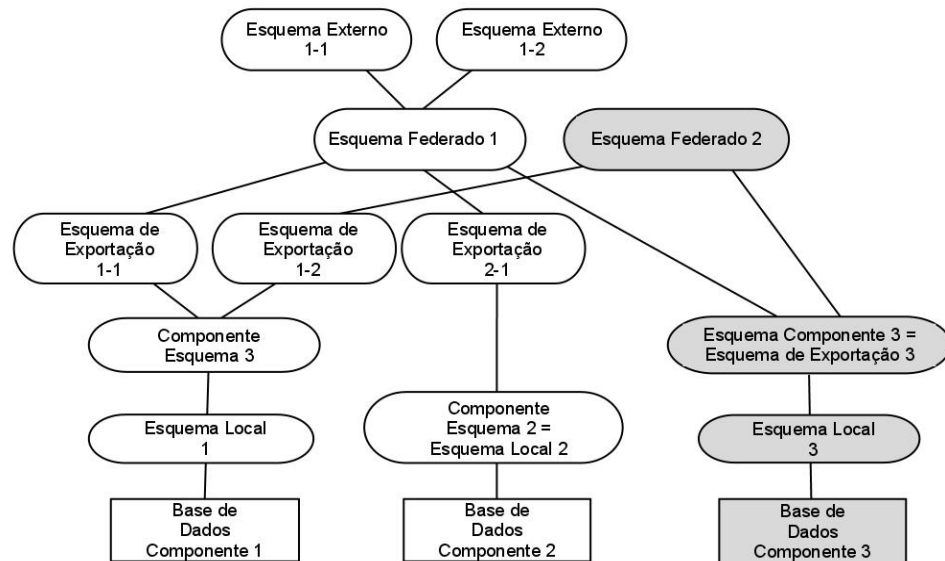
A *autonomia* está relacionada ao desenho das bases de dados, com a determinação dos dados que cada uma delas contém e sua modelação, de suas funcionalidades e estrutura física; assim como à capacidade de cada um dos Sistemas de Base de Dados componentes da federação em responder pela sua execução, mantendo operacionalidade de forma

independente da associação ao SBDF. Além disso, um SBD também é autônomo para optar por associar-se ou não a um Sistema de Base de Dados Federado.

A arquitetura de referência proposta por esses autores busca esclarecer questões relacionadas a sistemas de bases de dados, em que cada componente da arquitetura trata de um importante aspecto desse sistema, seja ele federado ou não. Assim, podemos nos concentrar em apenas alguns desses aspectos por vez, analisando seus componentes individualmente.

Uma arquitetura de referência provê uma estrutura pela qual entender, categorizar e comparar diferentes opções arquitetônicas para o desenvolvimento de sistemas de bases de dados federados. (SHETH; LARSON, 1990:192)

No que se refere às relações esquemas de metadados de um sistema federado de bases de dados, consideramos o modelo apresentado na figura 8, e em especial os componentes por nós destacados, como representativo da estrutura que compõe a proposta deste trabalho.



**Figura 8** – Relações entre esquemas de sistemas federados de bases de dados “com esquemas faltando em alguns níveis” (destaque nosso) propostos de Sheth e Larson (1990)

As relações de equivalência dos esquemas de metadados que compõem nossa proposta com aqueles apresentados por Sheth e Larson são as seguintes:

Esquema Local é o esquema original das bases de dados que se terão seus registros compartilhados. Em nossa proposta, elas correspondem àquelas participantes do sistema federado de bases de dados.

Esquema Componente ou Esquema de Exportação é algum esquema desenvolvido para possibilitar ou facilitar a conversão dos registros compartilhados para o esquema federado. Em nossa proposta, esse esquema corresponde ao esquema de metadados intermediário mencionado na seção anterior.

Esquema Federado é o esquema que todas as bases de dados participantes do sistema federado compartilham. Para nós, trata-se do Perfil de Aplicação proposto.

Um esquema externo se referiria a um modelo composto por um subgrupo de metadados, extraído de um esquema federado, criado para responder às especificidades de uma aplicação ou serviço para os quais o conjunto total de metadados não fosse necessário. No entanto, está fora do escopo deste trabalho a questão da utilização do modelo de catálogo para publicações em quadrinhos por serviços ou aplicações específicos.

Em uma abordagem mais recente sobre sistemas federados, Ferreira e Souto alertam que,

Como resultado do *Workshop Engineering Federated Database Systems*, em Barcelona na Espanha, a comunidade científica – observando a necessidade de integrar também diversificados sistemas digitais às bases de dados federadas – propõe uma nova terminologia: sistema de informação federado (SIF). (FERREIRA; SOUTO, 2006:26)

Um Sistema de Informação Federado, para Ferreira e Souto, comportaria dois tipos de sistemas de informação distribuídos: busca federada e sistema de coleta de metadados ou *harvesting*.

Nossa proposta baseia-se nesse segundo tipo de sistema de informação distribuído, e sobre ele nos ateremos mais na próxima seção.



## 3.2 A Iniciativa de Arquivos Abertos e o Protocolo OAI-PMH

A Iniciativa de Arquivos Abertos, ou OAI (*Open Archives Initiative*) surgiu como um empreendimento da comunidade de *e-prints* (artigos científicos disponibilizados digitalmente, antes ou depois de revisão por pares<sup>24</sup>), com a Convenção de Santa Fé, no Novo México, em outubro de 1999.

Foi proposta, nesse encontro, a transformação da comunicação acadêmica por meio da definição de aspectos técnicos e organizacionais de uma estrutura visando uma rápida disseminação e aumento da oferta de artigos na comunidade científica “na qual camadas livres ou comerciais pudessem ser estabelecidas”. (SOMPEL; LAGOZE, 2000)

Embora o apoio a esse trabalho continue sendo fundamental, a Iniciativa de Arquivos Abertos define-se correntemente como “uma organização e um esforço explicitamente em transição<sup>25</sup>”, com a estrutura tecnológica e padrões em desenvolvimento independentes do tipo de conteúdo, tendo por objetivo promover a interoperabilidade entre sistemas independentes, e sendo aberta a qualquer instituição interessada no compartilhamento de metadados.

Tal interoperabilidade é propiciada por meio do chamado Protocolo de Coleta de Metadados, ou OAI-PMH (*Protocol for Metadata Harvesting*).

Resumidamente, OAI-PMH é um mecanismo para coletar metadados formatados em XML de coleções distribuídas de metadados. Num sentido mais amplo, é uma estrutura para aumento da interoperabilidade que inclui um modelo arquitetural, orientações de implementação, um registro de implementadores, e um formato comum de metadados descritivos, além do próprio protocolo de coleta. (CAPLAN, 2004)

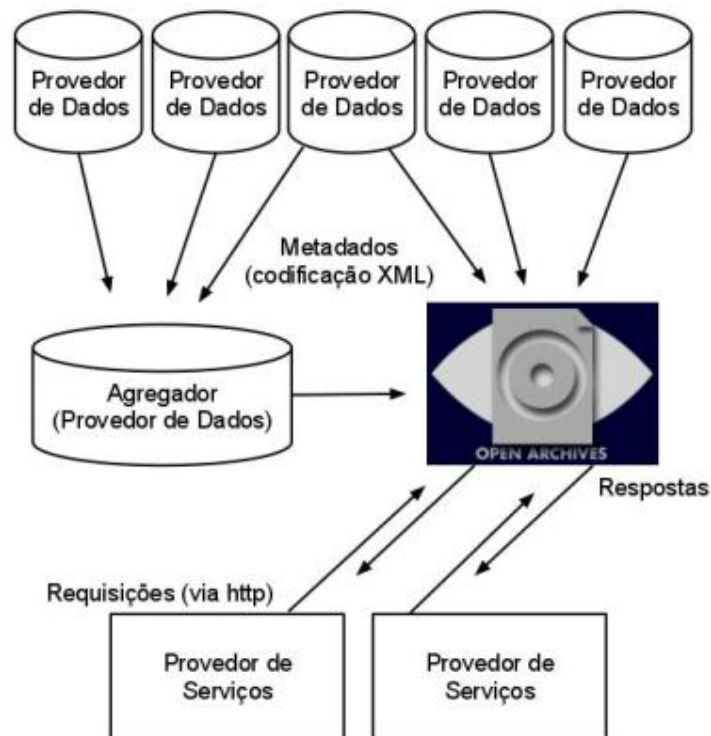
---

<sup>24</sup> *Eprints.org Network*, [s.d.]. Disponível em: <http://www.eprints.org/openaccess/self-faq/#What-is-Eprint>. Acesso em novembro de 2011.

<sup>25</sup> Open Archives Initiative. About OAI. Disponível em: <http://www.openarchives.org/OAI/OAI-organization.php>. Acesso em novembro de 2011.

Esse protocolo define como deve ser feita a transferência de metadados entre computadores de provedores de dados e de serviços, como ilustra a figura 9.

Os provedores de dados (*data providers*) são aplicações que buscam dados em bases de dados distribuídas e os disponibilizam aos provedores de serviços (*service providers*), que coletam esses dados para transmiti-los a usuários -- pessoas ou outras aplicações --, "usualmente na forma de serviços de pesquisa e recuperação" (CAPLAN, 2004). Essa comunicação ocorre por meio da Internet, sendo que as requisições desses dados, realizadas pelos provedores de serviço, baseadas no protocolo HTTP, e as respostas dos provedores de dados fornecidas em linguagem XML.



**Figura 9** – Plano de funcionamento do protocolo OAI-PMH

Seis tipos de requisições (ou “verbos”) podem ser estabelecidos na comunicação entre provedores. O provedor de serviços não é obrigado a usar a todos, mas um repositório precisa estar preparado para interpretar e responder a todos – ainda que com uma mensagem de erro. São eles:

*Identify* – Retorna informações sobre o provedor dos registros;

*ListSets* – Lista as categorias (opcionais) em que os documentos são classificados em um provedor de dados;

*ListMetadataFormats* – Lista os formatos de metadados dos documentos armazenados em um provedor de dados;

*ListRecords* – Lista os metadados dos registros armazenados no provedor de dados segundo um formato de metadados, especificando todos que pertencem a um set ou todos a partir de uma data

*ListIdentifiers* – Lista “identificadores únicos” dos registros armazenados em um provedor de dados, ou seja, códigos atribuídos aos registros que funcionam como URIs em relação aos repositórios;

*GetRecords* – Recupera um registro de metadados individual, mediante discriminação do identificador do registro e de seu formato.

Parâmetros ou argumentos de pesquisa – opcionais ou obrigatórios – podem ser existir para uso em conjunto com cada uma dessas requisições, como a limitação de um período de tempo referente à criação, modificação ou exclusão<sup>26</sup> de um registro (*datestamp*); ou a indicação do formato de metadados referente ao registro que se pretende recuperar (*metadataPrefix*).

O protocolo OAI-PMH demanda que todos os provedores de dados suportem o esquema de metadados dc\_oai – uma representação XML do DC Simple, cujos metadados compõem o *namespace* “oficial” do OAI. Garante-se, assim, uma interoperabilidade mínima baseada nos quinze elementos cerne do padrão Dublin Core. No entanto, outros esquemas de metadados podem ser combinados entre provedores de dados e serviços – desde que representados em XML.

O OAI PMH estabelece o *Dublin Core Metadata Element Set* como um conjunto mínimo de metadados a ser suportado pelos provedores de dados em resposta a uma solicitação de um provedor de serviços. No entanto, o provedor de serviços pode, a seu critério, oferecer outros formatos de metadados, mais amplos e complexos, como o MARC. (MARCONDES; SAYÃO, 2002)

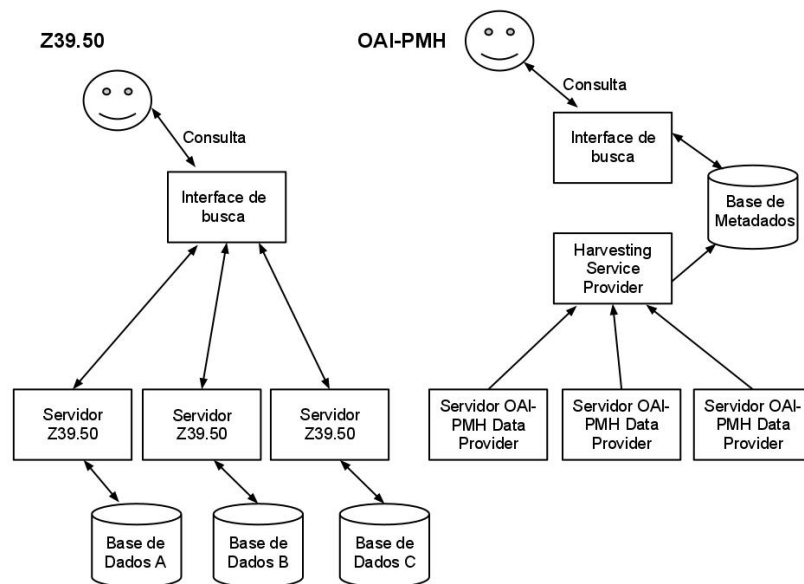
---

<sup>26</sup> Alguns repositórios podem manter o “cabeçalho” de um registro, mesmo depois de deletar o arquivo correspondente da base.

Essa característica do protocolo é apontada como uma desvantagem por Caplan, pois nessa operação “todos os metadados são convertidos para o formato de menor denominador comum, que pode não ser ótimo para certos propósitos” (CAPLAN, 2004), e chama a atenção para desafios como o controle de vocabulário na agregação de *sites* distintos.

Por outro lado, a autora afirma que o modelo de pesquisa operado pelo OAI se assemelha ao da Internet, em que “(meta)dados de sites distribuídos são agregados centralizadamente para pesquisa e recuperação” (CAPLAN, 2004), tornando o fator escalabilidade sua maior vantagem, quando comparado a modelos de pesquisa distribuídos, que é o caso do Z39.50, o qual “tende a quebrar com mais de uma dúzia de alvos [*targets*], enquanto centenas de sites podem ser agregados sob OAI.” (CAPLAN, 2004)

Marcondes e Sayão (2002) ilustram as diferenças nas estruturas dos protocolos OAI-PMH e Z39.50, como reproduzido na figura 10.



**Figura 10** – Quadro comparativo dos protocolos Z39.50 e OAI-PMH, adaptado de Marcondes e Sayão (2002)

O mecanismo pelo qual provedores de dados lidam com grandes volumes de resultados em uma pesquisa no protocolo OAI-PMH é conhecido como *resumption token*, e sua implementação por parte de repositórios é um requisito do sistema. Por meio desse mecanismo efetua-se o controle de fluxo dos dados, repartindo uma lista muito numerosa de

registros, marcando até que ponto da lista metadados já foram retornados e aguardando comandos para continuidade do processo.

### **3.3 Catalogação de publicações em quadrinhos**

As principais fontes de catalogação bibliográfica de publicações em quadrinhos disponíveis na Web são alimentadas por colecionadores voluntários, o que evidencia a ausência de iniciativas públicas no registro desse tipo de material – alvo de um histórico desprestígio como meio de comunicação.

O reconhecimento do valor das histórias em quadrinhos tem crescido nas últimas décadas, mas a lacuna em sua documentação em instituições como bibliotecas persiste. Segundo Scott Randall (1990:62), “bibliotecas nos Estados Unidos não catalogaram mais que um punhado de títulos antes de 1980”, e ainda:

A Biblioteca do Congresso Norte-americano nunca proveu catalogação para revistas em quadrinhos, como faz para quase todos as outras categorias de materiais publicados. Até o final da década de 1970, nenhum bibliotecário em qualquer parte do mundo poderia provar, usando qualquer catálogo padrão de biblioteca, que títulos como Mulher Maravilha, Super-homem ou O incrível homem-aranha existissem como entidades bibliográficas. (SCOTT, 1990:14-15)

O autor afirma existirem ainda outras razões pelas quais revistas em quadrinhos não são facilmente encontradas em bibliotecas, como a inconveniência do formato, que os torna extremamente frágeis, e o relativo baixo custo de aquisição, que possibilita que seus leitores potenciais possam comprá-los, não havendo necessidade de tomá-los emprestados em uma biblioteca pública.

Para Vergueiro,

Tendo sido objeto de perseguições e preconceitos em sua longa trajetória, hoje as histórias em quadrinhos já são encaradas de forma muito mais positiva por parcelas cada vez maiores da sociedade. Aos poucos, elas passaram a ser aceitas nos mais diversos ambientes educacionais, sendo utilizadas por professores de todas as áreas e

níveis de ensino e se tornando objeto de atenção de pesquisadores no mundo inteiro. Em conseqüência, passaram a fazer parte de coleções das bibliotecas, passando a representar um desafio para os profissionais da informação devido a pouca familiaridade que estes tinham (e têm) com esse produto editorial. (VERGUEIRO, 2005)

De uma forma geral, as bibliotecas que catalogam revistas em quadrinhos o fazem atribuindo um registro para toda a série de edições publicadas sob um mesmo título, de alguma maneira especificando quais de seus números ou volumes ela referente a esse título ela dispõe. Informações como dimensão e número de páginas são extraídos de uma edição exemplar.

As figuras 11 e 12 apresentam exemplos de registro de revistas em quadrinhos criados para bibliotecas. A primeira delas traz uma ficha catalográfica de revista em quadrinho pertencente ao acervo da Biblioteca da Universidade de Michigan<sup>27</sup>; a segunda, proveniente da Universidade Estadual de Bowling Green<sup>28</sup>, exibe o registro de um título em formato MARC, seguido da indicação das edições disponíveis do título em questão. Ambas as universidades são norte-americanas e disponibilizam seus registros na Web, através do *website* de suas bibliotecas.

```
Batman. -- New York : Detective Comics, 1940- . -- col. ill. ;  
26 cm. -- Began with no. 1 (Spring 1940), cf. Comic Book  
Price Guide. -- Later issues published by National  
Periodical Publications, then by DC Comics. -- LIBRARY HAS:  
no. 0, 3, 5-6, 17, 20, 22-23, 25, 27, 29, 33-35, 37-40,  
42-52, 54-55, 73, 75, 77-79, 99, 107-108, 113, 122, 126,  
128-135, 138-139, 141-379, 381-382, 384-514, 516-528,  
530-532, 535-552, 555-557, 559, 566-567, 586 (1940-2000).  
-- Call no.: PN6728.1.N3B3
```

**Figura 11** – Ficha catalográfica do acervo da Biblioteca da Universidade Estadual de Michigan

<sup>27</sup> Michigan State University – Special Comics Collection – website: <http://comics.lib.msu.edu/rri/index.htm>


<sup>28</sup> Bowling Green State University – Browne Popular Culture Library – website:  
<http://www.bgsu.edu/colleges/library/pcl/>

```

LEADER 00000nam 2200229Ia 4500
001 70688769
003 0CoLC
005 20060725091446.0
008 060725m2006999nyua 000 c eng d
040 BCU|cBCU
049 BCUp
090 PN6720.2.D33|bJ8386
245 10 Justice League unlimited
260 New York, NY :|bDC Comics,|cc2006-
300 v. :|bchiefly col. ill. ;|c26 cm
500 No. 1 (May 2006)-
520 The super heroes Superman, Wonder Woman, Batman, and
others join forces in the Justice League to combat
criminals, villains and would-be conquerors wherever they
threaten
550 0 Superheroes|xComic books, strips etc.
550 0 Comic books, strips, etc

```

View additional copies or search for a specific volume/copy

Location	Call Number	Status 
Pop Cult Bravard Comics	<a href="#">pop PN6720.2.D33 J8386</a> no.1 <a href="#">Map It</a>	Lib Use Only
Pop Cult Bravard Comics	<a href="#">pop PN6720.2.D33 J8386</a> no.2 <a href="#">Map It</a>	Lib Use Only
Pop Cult Bravard Comics	<a href="#">pop PN6720.2.D33 J8386</a> no.3 <a href="#">Map It</a>	Lib Use Only
Pop Cult Bravard Comics	<a href="#">pop PN6720.2.D33 J8386</a> no.4 <a href="#">Map It</a>	Lib Use Only
Pop Cult Bravard Comics	<a href="#">pop PN6720.2.D33 J8386</a> no.5 <a href="#">Map It</a>	Lib Use Only
Pop Cult Bravard Comics	<a href="#">pop PN6720.2.D33 J8386</a> no.6 <a href="#">Map It</a>	Lib Use Only
Pop Cult Bravard Comics	<a href="#">pop PN6720.2.D33 J8386</a> no.7 <a href="#">Map It</a>	Lib Use Only
Pop Cult Bravard Comics	<a href="#">pop PN6720.2.D33 J8386</a> no.8 <a href="#">Map It</a>	Lib Use Only
Pop Cult Bravard Comics	<a href="#">pop PN6720.2.D33 J8386</a> no.9 <a href="#">Map It</a>	Lib Use Only
Pop Cult Bravard Comics	<a href="#">pop PN6720.2.D33 J8386</a> no.10 <a href="#">Map It</a>	Lib Use Only

**Figura 12** – Registro MARC de revista em quadrinho da coleção da Universidade Estadual Bowling Green

Nos dois casos são claras as limitações que esse tipo de classificação impõe à recuperação, derivadas da generalização dos dados dos itens, ao contrário do registro de uma publicação monográfica, que é caracterizada, em todos os seus aspectos, singularmente.

Embora seja comum, quando se trata de uma série de revistas em quadrinhos, que todas as edições tenham as mesmas dimensões – e muitas vezes até o mesmo número de página –, isso nem sempre é verdadeiro, principalmente quando se trata de títulos de longa duração. Além disso, não é raro que outras características mudem no percurso de um periódico, como aquelas relativas à responsabilidade intelectual. No exemplo da ficha catalográfica, observa-se a indicação de alterações na editora do periódico: “*Later issues published by National Periodical Publications, then by DC comics*” (Edições mais antigas publicadas pela National Periodical Publications, em seguida por DC Comics); no entanto, a

um eventual pesquisador para o qual seja importante uma informação precisa sobre o publicador, essa base de dados não responderá a essa necessidade. No exemplo do registro MARC, a indicação de que o primeiro volume da série foi publicado em Maio de 2006 foi incluída como uma nota – campo em que usualmente não requer maiores cuidados com a padronização de seu conteúdo; apesar de haver *hyperlinks* acompanhando a indicação de cada um dos volumes da série sob a guarda da instituição, esses não levam a fontes de informação mais detalhadas sobre os itens.

Como se pode constatar, essas bibliotecas optaram por realizar a catalogação das revistas em quadrinhos no nível da série. Trata-se de uma escolha justificada provavelmente pelo que se considerou ser a melhor forma de tratar o material tendo em vista a relação custo-benefício. Esse tipo de julgamento que faz parte das funções de um bibliotecário, e uma decisão desse tipo será invariavelmente imbuído de “prós e contras”.

No entanto, o aumento de atenção que os quadrinhos tem conquistado na sociedade nas últimas décadas, especialmente no meio acadêmico, tornando-se fonte de informação e objeto de pesquisas em disciplinas variadas tende a culminar na necessidade de um maior detalhamento referente ao conteúdo das publicações, bem como aos aspectos que dizem respeito ao suporte, possibilitando o acompanhamento de transformações ocorridas nesse recinto.

Na medida em que as histórias em quadrinhos se tornaram um elemento de grande influência na cultura popular, também o interesse por elas aumentou em todas as áreas. Pesquisas acadêmicas sobre quadrinhos têm surgido nas áreas do conhecimento mais diversas, como história, sociologia, artes, literatura, antropologia, educação, etc., ampliando a disponibilidade de informações de qualidade sobre eles. (VERGUEIRO, 2005)

Assim, fontes de informação sobre publicações em quadrinhos que as examinam holisticamente auxiliam na resposta a essa demanda. Iniciativas importantes existem em linha, como mencionado anteriormente. Trata-se de bases de dados empreendidas voluntariamente, principalmente por colecionadores e fãs de histórias em quadrinhos – muitos deles consumidores ávidos dessa forma de expressão artística, podendo ser considerados “especialistas do ramo.”



De uma forma geral, esse tipo de base de dados de publicações em quadrinhos na Web divide as informações em seções, separando dados sobre origem e responsabilidade, conteúdo e suporte físico, de forma muito parecida com que registros bibliográficos são organizados em bibliotecas. Um esforço cooperativo é realizado para o preenchimento de formulários com as características de cada publicação. Tomando de empréstimo a terminologia adotada pelo modelo de dados bibliográfico FRBR, essas características são geralmente analisadas do ponto de vista da *manifestação*, abstraídas do *item* bibliográfico que os colaboradores da base de dados têm em mãos. Também é comum que essas bases de dados possibilitem a identificação dos exemplares por meio dos catalogadores que os possuem.

As figuras 13 e 14 exemplificam registros de publicações em quadrinhos existentes em *ComicBookDB.com*<sup>29</sup>, apresentando dados no nível da série e nível do item da série, respectivamente.

---

<sup>29</sup> ComicBookDB.com: the comic book database. Disponível em: <http://www.comicbookdb.com/index.php>

**Persépolis [BRA] (2004)**  
[Search for 'Persépolis \[BRA\]' on eBay](#)

**Publisher:** [Companhia das Letras](#)

**Publication Date:** October 2004 - March 2007

**Country:** Brazil

**Language:** Portuguese

**Notes:**  
None entered.

**Number of issues cataloged:** 5

[Cover Gallery](#)

Issue	Title	Story Arc	Cover Date
4	<a href="#">Persépolis 4</a>		March 2007
3	<a href="#">Persépolis 3</a>		March 2006
2	<a href="#">Persépolis 2</a>		April 2005
1	<a href="#">Persépolis 1</a>		October 2004
	<a href="#">TPB Persépolis Completo</a>		December 2007

[Edit This Title](#)

[Add an Issue to This Title](#)

[Add a batch of issues from this title to your collection](#)

[Add a batch of issues from this title to your wishlist](#)


[Add this title to your pull list](#)

[View the contribution history for this title](#)

[Report a problem regarding this title](#)

Figura 13 – Registro de publicação seriada em ComicBookDB

**Fruits Basket (2004) - vol. 05**  
Tokyopop Manga (Tokyopop)



[Change this cover or add a variant](#)

**Writer(s):**  
[Naka Hatake - 'Natsuki Takaya'](#)

**Penciller(s):**  
[Naka Hatake - 'Natsuki Takaya'](#)

**Inker(s):**  
[Naka Hatake - 'Natsuki Takaya'](#)

**Colorist(s):**  
(Story is monochromatic)

**Editor(s):**  
[Jake T. Forbes](#)

**Cover Artist(s):**  
[Naka Hatake - 'Natsuki Takaya'](#)  
[Aaron Suhr](#)

[Prev Issue](#) [Next Issue](#)

[Search for 'Fruits Basket' on eBay](#)

**Cover Date:** October 2004  
**Cover Price:** US \$9.99

**Issue Tagline:** None.

**Format:** Black & White; Manga; 208 pages

**Story Arc(s):** [Add/remove story arcs to this issue](#)

**Synopsis:**  
None entered.

**Reprinted/Collected in:**  
[Fruits Basket \(2004\) Ultimate Edition Vol. 03](#)

**Characters:** [Add/remove characters to this issue](#)  
[Saki Hanajima](#) [Kyo Sohma](#)  
[Ayame Sohma](#) [Shigure Sohma](#)  
[Hatori Sohma](#) [Yuki Sohma](#)

**Groups:** [Add/remove groups to this issue](#)

**Reviews:** There are no reviews for this issue. - [Add your review](#)

Figura 14 – Registro de item de série em ComicBookDC

Na Web, algumas bases de dados se destacam pelo volume de registros, como é o caso da *ComicBookDB.com*, apresentando 32.751 títulos e 231.270 edições; e *Grand Comics Database*<sup>30</sup>, com 61.987 títulos de séries e 710.872 edições (números aproximados relativos ao mês de novembro de 2011).

É importante observar que, como a maior parte dos recursos de informação disponíveis na Web, essas duas grandes bases de dados de publicações em quadrinhos são construídas em língua inglesa – idioma em que a maior parte das publicações registradas também foram escritas. Mas existem alternativas interessantes: em língua francesa, um banco

<sup>30</sup> *Grand Comics Database (GCD)*. Disponível em [www.comics.org](http://www.comics.org).

de dados comercial chamado *Bedetheque*<sup>31</sup> fornece um módulo para exportação para Web dos dados de registros de publicações em quadrinhos de seus clientes (indivíduos e vendedoras). Esses dados são incorporados ao *site* do *software*, conformando uma base unificada, que em novembro de 2011 contava com a marca de pelo menos 110.383 álbuns e 24.358 séries; em português, o site Guia dos Quadrinhos<sup>32</sup> oferece um banco de dados para registro de gibis publicados no Brasil, contando com 7.404 títulos e 80.378 edições no início de novembro de 2011.

Não é raro que os mesmos voluntários que alimentam essas bases de dados cooperativas registrem de alguma forma os itens de suas coleções particulares. Antes da popularização dos computadores pessoais, coleções podiam ser registradas, por exemplo, por meio de cartões adquiridos para essa finalidade, como ilustrado na figura 15<sup>33</sup>. Atualmente, diversos *softwares* são oferecidos na Web para que colecionadores mantenham organizadas suas aquisições em quadrinhos, possibilitando a exportação de registros em vários formatos.

TITLE: <i>Amazing Spider-Man</i>									
GRADING :	M/MINT	NM/NEAR MINT		FI/FINE		VG/VERY GOOD		G/GOOD	
	10	20	30	40	50	60	70	80	90
01	11	21	31	41	51	61	71	81	91
02	12	22	32	42	52	62	72	82	92
03	3	23	33	43	53	63	73	83	93
04	14	24	34	44	54	64	74	84	94
05	15	25	35	45	55	65	75	85	95
06	16	26	36	46	56	66	76	86	96
07	7	17	27	37	47	57	67	77	87
08	18	28	38	48	58	68	78	88	98
09	19	29	39	49	59	69	79	89	99

COLLECTION INVENTORY CARDS © ELDORADO COMICS 1978

Figura 15 – Cartão para inventario das edições de um título de revista em quadrinhos

Não é possível quantificar bases de dados pessoais *offline*, bem como obter números seguros das bases de dados em linha desse tipo de registro, considerando a variedade de escopo, formatos e linguagens em que podem ser encontradas. Entretanto, assumindo a

<sup>31</sup> *Bedetheque*. Disponível em <http://www.bedetheque.com/>.

<sup>32</sup> *Guia dos Quadrinhos*. Disponível em: <http://www.guiadosquadrinhos.com/>.

<sup>33</sup> Imagem obtida em Westfield Comis Blog! *KC column: closed for inventory*. [s.d.] Disponível em: <http://westfieldcomics.com/blog/interviews-and-columns/kc-column-closed-for-inventory/>

potencial existência de um sem número de bases de dados caseiras, de pouca ou nenhuma visibilidade na Web, vislumbramos neste trabalho a possibilidade de integrá-las em uma base de dados comum, assim como aquelas criadas cooperativamente por voluntários e as pertencentes a bibliotecas e outras organizações – enfim, a integração de qualquer base de dados de publicações em quadrinhos em meio digital, que sigam ou possam vir a seguir algumas convenções conceituais, técnicas e estruturais definidas adiante.

As vantagens mais claras e imediatas desse tipo de integração estão na visibilidade que tal base de dados proporcionaria, pela potencial diversidade de registros de publicações em quadrinhos, em termos de origem geográfica, temáticos e de linguagem, indicando uma amplitude de expressividades culturais inerentes a esse meio de comunicação, que não é comumente encontrada em coleções de escopo mais delimitado; na promessa de economia de tempo de usuários com a intenção de poupá-lo ao fazer uso de uma única interface de busca em suas pesquisas; e no reaproveitamento do trabalho já realizado do levantamento de dados sobre tais publicações.

### **3.3.1 Publicações em quadrinhos: características**

As primeiras revistas em quadrinhos (ou gibis, como são comumente conhecidas no Brasil) surgiram como reimpressões de tiras em quadrinhos publicadas em jornais que, nos Estados Unidos, tornaram-se comuns a partir do início do século 20. Seu formato característico estabeleceu-se na década de 1930.

A revista em quadrinhos [*comic book*] como um meio distinto data de meados dos anos 1930. Antes disso, existiram reimpressões de coleções de tiras de jornais, que começaram a aparecer na virada do século. (...) O formato físico da revista em quadrinhos como a conhecemos hoje também começaram como reimpressões, sendo o primeiro título regularmente publicado foi *Famous Funnies*, com início em 1934. (SCOTT, 1990:106)

Esse novo meio de comunicação obteve sucesso imediato, principalmente com a população mais jovem e aberta para experimentar novidades, configurando-se como público majoritário desse tipo de literatura (SCOTT, 1990). Características próprias desse produto, aliadas a demandas crescentes de produção e seu consumo acelerado favoreceram o

surgimento de uma forma “industrializada” de criação, dividindo seu processo em etapas, realizadas por diferentes artistas, de acordo com suas habilidades.

Diversas funções estão, assim, comumente envolvidas na criação de uma revista em quadrinhos, principalmente daquelas que historicamente obtiveram maior sucesso de vendas, fidelizando uma parcela significativa do mercado de consumo. Além dos previsíveis roteirista e ilustrador, na divisão do trabalho encontram-se balonistas, coloristas, arte-finalistas, tradutores, editores, etc. Uma revista em quadrinhos típica pode conter uma história, a parte de uma história ou várias histórias parciais ou integrais – a cada uma dessas histórias correspondendo um diferente time de criadores.

No entanto, a identificação desses agentes nem sempre é possível. De uma forma geral, editoras de revistas em quadrinhos que detêm os direitos autorais sobre suas histórias e personagens não têm interesse em dar crédito aos membros de suas equipes criadoras. Segundo Vergueiro, “a prática de elucidar, nas próprias publicações, a verdadeira autoria das histórias, não foi correntemente adotada durante bem mais da metade dos mais de cem anos de existência dos quadrinhos como meio de comunicação de massa.” (VERGUEIRO, 2003). Ainda segundo esse autor, esse quadro começou a mudar sob influência global do mercado norte-americano a partir dos anos sessenta, “quando a editora *Marvel Comics*, pelas mãos de *Stan Lee*, passou a investir fortemente na divulgação dos nomes de desenhistas e roteiristas como meio de criar um ambiente familiar no relacionamento com os leitores. (VERGUEIRO, 2003)

Outra peculiaridade das revistas em quadrinhos, como produto editorial, reside em certo grau de informalidade na determinação e exposição de informações sobre a própria edição publicada. Editoras de revistas em quadrinhos não apresentaram preocupação significativa em estabelecer padrões, mesmo para edições pertencentes a uma mesma série, como por exemplo, a uniformidade de títulos ou métodos que assegurassem uma indicação lógica e inequívoca do seqüenciamento dessas edições ao longo do tempo.

Sob muitos aspectos, os *gibis* representam um mercado totalmente caótico, sem qualquer tipo de padronização em relação à numeração, uniformidade dos títulos ou continuidade, com almanaques e números especiais intercalando títulos regulares, podendo tanto receber uma numeração própria como seguir a mesma seqüência

numérica do título principal, numa balbúrdia difícil de compreender por aqueles que não estão familiarizados com o meio. (VERGUEIRO, 2005)

Embora esse tipo de postura ainda seja usual, seus efeitos são comumente amenizados pela presença da seção de créditos (ou *indicia*), onde se podem encontrar informações como designação jurídica do publicador, endereços corporativos, número e data de publicação da revista, periodicidade e natureza dos direitos autorais relativos à série, além de um título genérico, diferente daquele que consta na capa da revista.

A *indicia* quase sempre começa com o título do livro em quadrinhos impresso em letras maiúsculas, e este título é a base para a consistência na bibliografia de quadrinhos. O título de capa de uma revista em quadrinhos pode parecer variar muito ao iniciante, mas o título da *indicia* é conservador em relação a mudanças. (SCOTT, 1990:115-6)

O surgimento do ISSN<sup>34</sup> como o meio padrão de identificação de publicações seriadas, e sua adoção por editoras de revistas em quadrinhos, pode ter contribuído para a fixação de um título genérico para séries de revistas em quadrinhos, uma vez que o código numérico fornecido como valor do ISSN deve ligar-se a uma forma padronizada do título da série em questão – o que é conhecido como "título-chave". (ISSN, c2008) Assim, o número de ISSN tem a função de qualificar o título de uma publicação, diferenciando-a de publicações homônimas.

As publicações em quadrinhos consideradas no escopo deste trabalho pertencem à tipologia definida por Vergueiro (2005), com exceção de *quadrinhos em jornais* e *fanzines*; o tipo definido como *publicações variadas* fica restrito às publicações que permitam identificação de título, publicador e data de publicação.

Portanto, as publicações em quadrinhos consideradas aqui são dos seguintes tipos:

**Gibis:** nome popular atribuído no Brasil para revistas em quadrinhos, trata-se de publicações periódicas (frequentemente mensais), voltadas para o público infantil e juvenil. São tradicionalmente encontradas em bancas de jornal, em formato 13,5 x 19 cm, impressas

---

<sup>34</sup> ISSN (*International Standard Serial Number*) é um código numérico de oito dígitos usado para identificação de publicações seriadas, definido como padrão internacional pela norma ISO 3297, publicada em 1975, contendo sua definição e regras de aplicação.

em papel frágil, de pouca durabilidade, proporcionando um baixo custo. No entanto, eles também podem ser encontrados em formatos maiores e com qualidade material superior, seguindo padrões editoriais norte-americanos dos chamados *comic books*. Gibis podem dar origem a suplementos, edições especiais e almanaques que não sigam necessariamente o seqüenciamento numérico ou dêem continuidade a narrativa do periódico.

**Álbuns e edições encadernadas:** Suas histórias geralmente começam e terminam na mesma edição, com eventuais continuidades decorrendo do sucesso comercial de uma publicação pioneira. Normalmente seu custo é mais alto, devido à melhor qualidade do papel, impressão e encadernação. Ao contrário dos gibis, os álbuns são geralmente encontrados para vendas em livrarias.

**Graphic novels, maxi e minisséries:** fisicamente, de uma forma geral, se aproximam do conceito de álbum, porém caracterizam-se pela busca de uma maior qualidade gráfica e uma produção mais elaborada em termos de roteiro e arte, que surgiram como uma forma de revitalizar o gênero e atrair novos leitores. Podem constituir-se de um único volume ou uma série limitada – minisséries compõem-se geralmente de três a seis números; maxisséries, em torno de 16 números.

**Publicações variadas:** histórias em quadrinhos elaboradas para uso em publicidade, propaganda política ou como meio de transmissão de mensagens para públicos específicos, patrocinadas por instituições governamentais ou outros tipos de organizações.

### **3.4 PAPuQ: Perfil de Aplicação para Publicações em Quadrinhos**

Perfis de Aplicação (PA) são vistos primariamente como uma forma de documentação (HEERY; PATEL, 2000), que auxilia comunidades no processo de planejamento e implementação de práticas de metadados, e fornece subsídios para manutenção e avaliação dessas práticas.

Os perfis de aplicação provêm meios para expressar desde os princípios mais elementares de organização de um serviço de informação – como a própria determinação de sua razão de ser –, até a potencial automatização desse serviço em meio digital, com a definição de linguagens para interpretação por máquina dos dados comportados.

Nesta seção propomos, em linhas gerais, o desenvolvimento de um Perfil de Aplicação para Publicações em Quadrinhos – PAPuQ –, visando atender a demanda da comunidade de fãs, colecionadores e pesquisadores por algumas informações pontuais sobre esse tipo de publicação. Considerando a amplitude de acesso que a Web disponibiliza, outros tipos de público, imprevisíveis, podem vir a se beneficiar desse serviço.

Naturalmente, como qualquer serviço de informação, sua natureza é dinâmica e melhoramentos, incrementos, alterações de escopo podem ocorrer para atendimento satisfatório de seu público – original ou transformado. Por isso mesmo, a documentação do Perfil de Aplicação pode ser de significativa importância.

A construção do PAPuQ baseia-se nas especificações da Iniciativa de Metadados Dublin Core para Perfis de Aplicação, compondo-se assim de cinco partes:

1. Requerimentos Funcionais
2. Modelo de Domínio
3. Perfil do Conjunto Descritivo
4. Diretrizes de Uso
5. Orientação para Sintaxe e Formatos de Dados

Discorre-se sobre cada uma dessas partes nas próximas seções.

### **3.4.1 Requerimentos Funcionais**

Os objetivos de um Sistema de Informação são definidos em um perfil de aplicação por meio dos Requerimentos Funcionais. A definição do próprio SI proposto pode ser a melhor maneira de introduzir suas funções:



*PAPuQ é um modelo de dados para catálogo bibliográfico voltado a publicações em quadrinhos seriadas ou monográficas, que visa uma descrição generalista dessas publicações, a fim de favorecer a coincidência de sua estrutura conceitual com o modelo de dados de outros catálogos de escopo similar, e que faz uso de propriedades declaradas em namespace mantido pela Iniciativa de Metadados Dublin Core, buscando favorecer a interoperabilidade entre sistemas por meio do uso de padrões bem estabelecidos.*

Assim, podemos considerar como objetivo primário do PAPuQ a realização das funções de um catálogo bibliográfico, como definidas pela IFLA (*International Federation of Library Associations and Institutions*) na recomendação de um nível básico de funcionalidade para os registros criados por agências bibliográficas nacionais.

Segundo *Requerimentos Funcionais para Registros Bibliográficos*<sup>35</sup>, ou FRBR (*Funcional Requirements for Bibliographic Records*), os quatro usos elementares dos dados bibliográficos são:

- Uso dos dados para ACHAR materiais que correspondam aos critérios de pesquisa estipulados pelo usuário;
- Uso dos dados obtidos para IDENTIFICAR uma entidade;
- Uso dos dados para SELECIONAR uma entidade que é apropriada às necessidades do usuário;
- Uso dos dados para OBTER acesso à entidade descrita.

Essas são funções gerais definidas para quaisquer catálogos bibliográficos. O nível de detalhamento em que as buscas podem ser realizadas e os meios de identificação, seleção e obtenção dos recursos descritos variam conforme peculiaridades dos catálogos.

No caso da implementação do PAPuQ aqui proposta, como um Sistema de Informação Federado, pretende-se possibilitar o reaproveitamento de registros existentes evitando a necessidade por reformulações muito sofisticadas na estrutura dos dados das bases

---

<sup>35</sup> IFLA Study Group on the *Funcional Requirements for Bibliographic Records*, 1998

componentes. Para tanto, é preciso que o registro do PAPuQ seja apto a representar desde aqueles advindos de bases de dados componentes mais simplesmente estruturadas: aquelas que servem a coleções pessoais, nas quais, por não ser comum haver mais de um exemplar da mesma publicação, tem no registro que descreve a publicação também a representação de um seu exemplar (aquele que pertence ao colecionador).

Nas bases de dados construídas de forma cooperativa encontradas na Web, as funcionalidades de inclusão de dados são usualmente similares às encontrados nos *softwares* de biblioteca atuais: o catalogador confere a existência do registro de uma determinada publicação; se essa publicação já foi adicionada à base, ele pode editar o registro acrescentando ou corrigindo dados; se o que existe na base é um registro parecido com aquele que o catalogador pretende inserir, ele pode duplicá-lo e então promover as alterações necessárias para caracterizar a publicação que tem em mãos, com reduzido esforço; se esse registro não existe, o catalogador procede a sua completa descrição – algumas bases de dados fornecem listagens de valores previamente incluídos, que tendem a ser recorrentes, para campos como “Editora”, “Criador” ou “Personagem”, o que gera uma espécie de vocabulário controlado, facilitando a padronização dos dados, embora esse controle normalmente não seja rigoroso ou autoritário.

Essa estrutura descritiva, embora mais racional, por evitar duplicação desnecessária de dados, diminuindo a ocorrência de inconsistências – problemas “naturais” nesse tipo de trabalho, principalmente quando realizado por diferentes pessoas, ainda que essas sigam as mesmas orientações –, postula o exame da situação de cada publicação na base de dados, individualmente, e inviabiliza o reaproveitamento do trabalho dos voluntários, que tenham sido realizados previamente, em suas bases pessoais.

Na proposta de implementação sugerida para o PAPuQ, os registros de publicações em quadrinhos de cada componente do Sistema de Informação Federado entra como um item em sua integridade, independentemente da prévia inclusão de dados de uma mesma publicação.

A desvantagem óbvia dessa abordagem está na duplicação de dados, e a propendente geração de inconsistências que podem refletir na eficácia dos resultados de busca: dados provenientes de diferentes bases, que versem sobre a mesma publicação, mas

que por alguma razão diverjam morfológicamente (como em caso de erros de digitação), deixam de ser recuperados em conjunto quando uma das formas variantes é usada como termo de busca. Outro problema que pode surgir dessa questão morfológica está relacionado com a ordenação alfabética de um conjunto de resultados de busca.

Por outro lado, a vantagem consiste na simplificação do processo de agrupar resultados de, virtualmente, qualquer base de dados de publicações em quadrinhos. Os problemas inerentes a essa estrutura descritiva podem ser neutralizados por uma definição de escopo que tire o máximo proveito de suas qualidades e evite investir em questões a que essa estrutura não pode responder adequadamente. A limitação do modelo do registro para campos descritivos básicos (no caso do Protocolo OAI-PMH elementos cerne do padrão Dublin Core) parece ser uma boa solução.

Pode-se considerar, então, finalidades imediatas da coleta dos registros no Sistema Federado de Bases de Dados através do PAPuQ:

- Identificar exemplares de publicações em quadrinhos, por meio de pontos de acesso;
- Identificar proprietário desses exemplares;
- Identificar localização desses exemplares por meio do proprietário.

Pontos de acesso previstos incluem título, editora, ISSN ou ISBN, país de origem, idioma, data de publicação, criadores e contribuidores (responsáveis e co-responsáveis pela produção da obra) e personagens.

Prevê-se, assim, que os registros dessa base de dados são úteis idealmente para localização de exemplares de obras já conhecidos pelos usuários, sobre as quais possuam informações que podem ser usadas na construção de termos de busca, retornando resultados objetivos, embora a pesquisa de campos como “personagem”, “artista”, “país de origem” ou “idioma”, possibilite a descoberta de obras até então desconhecidas.

É importante ressaltar que tal Sistema de Informação Federado não pretende substituir outros tipos de bases de dados, como as bases cooperativas existentes na Web e os catálogos de publicações em quadrinhos de outras organizações. Ao invés disso, de uma

relação de complementaridade pode ser estabelecida: a identificação do proprietário de uma publicação pode indicar ao usuário onde obter informações mais detalhadas sobre ela – como proprietários de exemplares localizados ou a base de dados na Web que é fonte do registro, e possivelmente o apresenta com mais riqueza de detalhes.

Assim, pelo seu caráter referencial, entende-se que a implementação do PAPuQ, aqui proposta, pode ser útil principalmente para descrição de publicações oriundas da indústria editorial tradicional (físicas, não digitais), uma vez que os dados necessários para localização de uma instância do PAPuQ são os mesmos que poderiam ser usados para obter de forma direta o recurso existente em meio multimidiático, como os chamados *webcomics*, por meio de outros serviços de busca.

### **3.4.2 Modelo de Domínio**

Como já visto, modelos são representações limitadas de alguma coisa, criadas com um propósito específico, que determinam a ênfase em determinados aspectos da coisa representada.

Um modelo de dados define a estrutura e o significado de dados, e um modelo de domínio refere-se a essa estrutura e significados delimitados dentro de um domínio de interesse constituído.

West (2003:10) identifica uma hierarquia de modelos de dados, estabelecendo uma distinção entre modelo externo (ou visão) e modelo conceitual: para ele, um modelo externo olha para o mundo de uma perspectiva particular, existindo muitas visões externas possíveis no mundo, que podem se sobrepor e não têm que ser compatíveis; um modelo conceitual é subjacente ou neutro, capaz de suportar qualquer visão externa que caiba em seu escopo. Segundo o autor, os modelos de dados da maior parte das aplicações, hoje, tem uma visão de mundo de uma perspectiva daquela aplicação, e são, assim, modelos externos.

Analogamente, no Framework Singapura, encontramos o Modelo de Domínio e o Modelo de Domínio da Comunidade: o Modelo de Domínio da Comunidade situa-se na camada dos Padrões de Domínio, e serve de orientação para criar a visão a ser usada na

aplicação. Ao se pautar pelo Modelo de Domínio da Comunidade, essa visão manterá a coerência com representações dos dados mais abstratas; o Modelo de Domínio situa-se na camada do perfil de aplicação, representando mais concretamente as coisas que os dados descrevem e suas relações.

Assim, na criação de um Modelo de Domínio, entende-se que um implementador usará o Modelo de Domínio da Comunidade para, a partir dele, desenvolver um modelo de dados mais específico de idêntica esfera de interesse.

O modelo de dados do PAPuQ descreve registros bibliográficos. Sua comunidade está relacionada a um tipo especial de bibliografia: as publicações em quadrinhos. A visão de sua aplicação busca realizar a descrição dessas publicações de forma simples, permitindo sua expressão através dos elementos cerne do padrão Dublin Core, e baseando-se também nos requerimentos funcionais definidos para o Perfil de Aplicação, conforme preconizado pelo *Framework Singapura*.

Assim, o Modelo de Domínio da PAPuQ descreve duas coisas: exemplar e colecionador, por meio da estrutura de relacionamentos e léxico mostrado na figura 16. A semântica do léxico do Modelo de Domínio é apresentada em seguida, no quadro 6.

Este léxico apresenta da maneira menos ambígua possível as definições dos elementos que compõem o modelo conceitual, desempenhando um papel fundamental quanto à economia de tempo, minimização dos problemas de comunicação e, conseqüentemente, promoção de um diálogo mais eficiente e consistente acerca do domínio modelado. (GUIZZARDI, 2005:36)



Figura 16 – Modelo de Domínio PAPuQ

<b>Entidade do Modelo de Domínio</b>	<b>Descrição</b>
<i>Código local</i>	Referência não-ambígua do registro, dentro de um contexto (neste caso, o sistema federado de bases de dados).
<i>Exemplar</i>	Título pelo qual o recurso é conhecido.
<i>Artista</i>	Agente responsável por criar ou contribuir na criação do o recurso.
<i>Editora</i>	Agente responsável pela publicação do recurso.
<i>Local de origem</i>	Indicação geográfica de origem (comumente associado à Editora).
<i>Dados de conteúdo</i>	Dados que caracterizam o conteúdo do recurso.
<i>Data de publicação</i>	Data associada ao evento de publicação do recurso.
<i>Descrição física</i>	Discriminação das características materiais do recurso.
<i>Direitos autorais</i>	Declaração de direitos autorais do recurso.
<i>Código internacional</i>	ISBN ou ISSN do recurso.
<i>Tipo</i>	Categoria ou gênero do recurso.
<i>Colecionador</i>	Fonte da qual o recurso é procedente.
<i>Forma de contato</i>	Informações para contatar o colecionador.

**Quadro 6** – Semântica das entidades do Modelo de Domínio PAPuQ

### 3.4.3 Perfil do Conjunto Descritivo

O aspecto mais concreto de um Perfil de Aplicação, tratado nesta etapa de sua construção, refere-se a sua “descrição física”: após a definição dos Requerimentos Funcionais do sistema e da elaboração de seu Modelo de Domínio, procede-se para o levantamento das propriedades a serem usadas para representar o modelo.

Essa representação física é realizada com o uso de metadados formalmente declarados em *namespaces*. Como visto, a adoção de padrões de metadados favorece a interoperabilidade entre sistemas distintos disponíveis em rede – como no caso da implementação sugerida para o PAPuQ, a compor-se registros bibliográficos provenientes de bases de dados distribuídas –, permitindo sua incorporação a esse ambiente digital, de forma a beneficiar-se de suas potencialidades.

Essa parte do processo de construção de um Perfil de Aplicação também é a mais controversa em comparação a catalogação realizada tradicionalmente por bibliotecários, que agora, com a presença imperativa da tecnologia, “passam a integrar novos conhecimentos e práticas, no esforço multidisciplinar de se prover o acesso à informação da forma mais ágil e eficaz possível”. (BAPTISTA, 2007)

Os usos da descrição bibliográfica continuam os mesmos, desde a concepção da biblioteca como serviço de informação e fonte de materiais bibliográficos para o público – utilizando-se dos catálogos para descrição dos itens de acervo, a fim de suprir as demandas dos usuários. A grande diferença da atualidade está nos meios e no volume da produção da informação e na natureza eletrônica das ferramentas utilizadas para assegurar sua utilidade.

Interessa-nos, assim, transferir os conhecimentos empregados nas atividades que envolvem a catalogação para essas novas formas de suporte e novos processos de trabalho, a fim de prosseguir contribuindo para a organização da informação.

Bibliotecários em todos os setores têm um importante papel a desempenhar na configuração da Internet através da descoberta de recursos, e ao familiarizar-se com padrões de catalogação da Internet, assegurar que registros consistentes são construídos. (SHARP, 2002:104)

Entre os padrões de catalogação da Internet encontram-se os padrões de metadados e, portanto, para fins de construção de Perfis de Aplicação, é importante conhecer alternativas existentes e suas características, a fim de optar por aquelas que melhor respondem às necessidades de um sistema. Em outras palavras,

(...) para o uso adequado dos padrões de metadados é necessário conhecer as particularidades e diferenças entre suas estruturas e níveis de especificidade, e principalmente conhecer o princípio que norteou sua criação, ou seja, o escopo pelo qual e para qual finalidade foi desenvolvido. Essas características determinam, não só a aplicação dos padrões de metadados como também direciona o seu uso, de tal forma que a recuperação baseada em metadados possa ser realizada com maior qualidade. (ALVES; SANTOS, 2009)

Os elementos cerne do Padrão Dublin Core já foram definidos para representação das propriedades do PAPuQ, mas outros poderiam ser usados (como pode ser visto no modelo estendido do perfil de aplicação, tratado adiante).

Um modelo físico representa um modo pelo qual os dados são fisicamente armazenados. Podem existir muitos modelos físicos de um modelo conceitual. O requerimento é que um modelo físico seja hábil para apoiar o modelo conceitual. (WEST, 2003)

### **Escolha dos elementos para compor o Perfil de Aplicação**

A definição dos elementos de metadados cerne do Padrão Dublin Core apóia-se em amplo consenso internacional no uso dessas propriedades para descrições bibliográficas, sendo o próprio padrão “inspirado” padrão MARC. Uma simplificação, como visto, favoreceria seu uso para descrição de recursos da Web por parte de seus próprios criadores.

Em nível nacional, podemos apontar os casos da Embrapa, que “vem sentindo a crescente necessidade de estabelecimento de padrões de descrição de conteúdo de recursos de informação e a geração de metadados.” (SOUZA; MOURA; SANTOS, 2005) e da Biblioteca Digital de Teses e Dissertações da USP<sup>36</sup>, cuja base de dados foi estruturada utilizando-se o conjunto de elementos desse modelo de metadados. (ROSETTO; NOGUEIRA, 2002)

O uso do conjunto dos quinze elementos cerne do padrão Dublin Core também é requerido pela Iniciativa de Arquivos Abertos, para fins de coleta por meio do protocolo OAI-PMH, o que torna sua adoção um fator *sine qua non* no contexto da implementação para o PAPuQ aqui sugerida.

Assim, a análise da semântica dos elementos de dado desse padrão, apresentada no quadro 7, nos permite identificar de que forma esses metadados podem ser usados para atingirmos os objetivos do nosso perfil de aplicação, apontados por seus Requerimentos Funcionais, enquanto mantemos conformidade com o Modelo de Dados igualmente definido.

---

<sup>36</sup> Biblioteca de Teses da USP. Disponível em: <http://www.teses.usp.br>



Como notado anteriormente, o significado dos elementos cerne Dublin Core é bastante genérico, apto a englobar o universo semântico de propriedades que descrevem recursos dos mais variados tipo e domínios do conhecimento. Portanto, qualquer alteração nessa semântica para responder a particularidades descritivas locais, que por ventura sejam necessárias, será de caráter restritivo, nunca extensivo.

<b>Dublin Core</b>	<b>PAPuQ</b>	<b>Descrição</b>
<i>Title</i>	Exemplar	Nome pelo qual um recurso pode ser identificado
<i>Creator</i>	Criador	Agente primariamente responsável pela criação do recurso
<i>Contributor</i>	Contribuidor	Agente que contribui para criação do recurso
<i>Subject</i>	Personagem	Agente que figura na narrativa contida no recurso
<i>Date</i>	Data de Publicação	Data associada à publicação do recurso
<i>Publisher</i>	Editora	Entidade responsável pela produção material do recurso
<i>Coverage</i>	País de origem	País em que o recurso teve origem
<i>Language</i>	Idioma	Idioma do conteúdo intelectual do recurso
<i>Format</i>	Formato	Configuração física do recurso
<i>Rights</i>	Direitos	Informação sobre os direitos relacionados ao recurso e a seu uso
<i>Relation</i>	Código da publicação	Código ISBN ou ISSN relacionado ao recurso
<i>Description</i>	Notas	Dados adicionais sobre o recurso
<i>Source</i>	Colecionador	Indivíduo ou entidade que dispõe do recurso
<i>Type</i>	Tipo	Tipo do recurso (reservado para termos Imagem e Texto do vocabulário de termos DCMIType)
<i>Identifier</i>	Identificador	Identificação não ambígua do recurso no contexto de um Sistema de Informação

**Quadro 7** – Semântica das propriedades PAPuQ e equivalência com elementos Dublin Core

Abaixo precisamos as relações entre os elementos Dublin Core e seus usos no âmbito do Perfil de Aplicação para Publicações em Quadrinhos – PAPuQ.

### **Exemplar**

Segundo o AACR2, “Título é a palavra, frase, caractere, ou grupo de caracteres, que normalmente aparecem num item, dando nome a este ou à obra nele contida.” (RIBEIRO, 2006)

Entre os tipos produtos editoriais em quadrinhos existentes, encontram-se publicações seriadas e monográficas, as quais apresentam características físicas que são peculiares a esses tipos, que também podem influenciar na determinação de seu título.

Em ambos os casos, no entanto, o Título pode ser constituído de vários dados, sendo os mais comuns um título principal, subtítulo e indicações de parte. Especialmente para revistas em quadrinhos, que são comumente publicações seriadas, a indicação de parte é determinante na identificação de um exemplar, por este fazer parte de um conjunto de publicações em que sua ordem seqüencial é relevante.

Assim, o elemento Título do padrão Dublin Core corresponde à propriedade Exemplar do PAPuQ, o qual pode ser constituído de Título, Subtítulo, e indicações de parte – freqüentemente número e volume para revistas em quadrinhos.

### **Criador e Contribuidor**

Entre os tipos de publicações em quadrinhos, os gibis – e em especial os gibis de heróis – são, tradicionalmente, produzidos por grupos de agentes, relativamente numerosos: roteiristas, ilustradores, balonistas, coloristas, arte-finalistas, tradutores, editores, etc.

Como visto anteriormente, durante décadas essas funções não foram creditadas. Mas mesmo com a tendência iniciada em contrário, muitas vezes, tais créditos não são suficientes para vincular, de forma efetiva, a obra aos artistas por ela responsáveis. Isso acontece, pois esse tipo de revista em quadrinhos consiste, tipicamente, na apresentação de um personagem, ou um grupo de personagens, que não estão sob a “tutela” de seus inventores – ainda que esses últimos continuem sendo reconhecidos como seus criadores. Times de criação se revezam para dar vida a personagens existentes de longa data, sob uma linha editorial definida pela empresa que detém os direitos de exploração dos personagens.

Assim, a autoridade, nesses casos, não se mostra a forma mais apropriada de identificar esse tipo de publicação seriada. Situação análoga ocorre em relação à publicação de tiras em jornais, como observa Randall Scott:

Se existe uma maneira “correta” de listar quadrinhos, tem de ser por título. Cartunistas mudam, escritores vão e vem, mas o título de uma tira identifica a continuidade do começo ao fim. (SCOTT, 1990:70)

Esse quadro muda significativamente nos outros tipos de publicações em quadrinhos: álbuns, minisséries, maxisséries, *graphic novels*, identificadas como produções de um único artista ou, quando trabalho coletivo, usualmente obra de duplas formadas por roteiristas e ilustradores.

Pode-se caracterizar, assim, uma hierarquia dessas funções, em que roteirista e ilustrador comporiam os principais responsáveis intelectuais por uma obra em quadrinhos.

No Dublin Core, *Criador* e *Contribuidor* são as propriedades de metadados disponibilizadas para inclusão desses agentes. Ao condicionar a descrição de um registro aos elementos cerne desse padrão, abdica-se de discriminar cada uma das funções envolvidas na geração de uma obra em quadrinhos. A perda da especificidade, no entanto, pode ser atenuada pela convenção na distribuição dessas funções entre essas duas propriedades.

Propõe-se, assim, que agentes identificados como roteiristas ou ilustradores sejam equivalentes ao elemento Criador, e os demais papéis cumpridos pelos demais agentes, sejam equivalentes ao elemento Contribuidor.

Espera-se, ao fazer uso dessa oportunidade de refinamento dentro do conjunto de elementos Dublin Core Simple, prover maior eficácia em resultados de busca.

### **Personagem**

Novamente, este Perfil de Aplicação procura descrever publicações em quadrinhos para identificação de exemplares, da forma mais genérica possível, a fim de habilitar-se a

receber registros provenientes de repositórios que apresentam diversificadas modelagens de dados.

Esse fato, aliado ao número reduzido de propriedades de metadados do padrão Dublin Core, demanda que se opte pela captação dos valores que, além de significativos para a identificação de uma publicação, também contem com uma maior possibilidade de coincidências nas diferentes bases, tornando a alimentação dos campos dos registros mais consistentes.

Inúmeras possibilidades de conceitos, representações e opiniões são inerentes à atribuição de temas a qualquer tipo de recurso, não sendo diferente com publicações em quadrinhos. Diferentes pontos de vista sobre um mesmo aspecto podem ser – e geralmente são – obtidos de diferentes catalogadores. A busca por uma convenção mínima referente à descrição temática de publicações em quadrinhos seria um grande desafio, considerando a multiplicidade de fontes que se pretende agregar.

Assim, no PAPuQ, o elemento *Assunto* foi destinado a receber os valores referentes a personagens. Trata-se de um dado relevante do ponto de vista da identificação de uma publicação, e que também pode ser útil para fins de descoberta de recursos. E a designação de um personagem, ao contrário da atribuição de um tema, não é subjetiva. Pode-se variar, entre bases de dados, o grau de pormenorização em que os personagens são considerados para inclusão de registros – entre zero e todos, destacando-se a possibilidade de inclusão apenas dos personagens principais no registro. Mas excetuando-se as variações morfológicas que possa ocorrer na nomeação de um mesmo personagem, como nos casos de nomeação do personagem em diferentes idiomas, o uso de *personagem* como valor da propriedade Assunto conta com a vantagem da maior objetividade.

### **Data de Publicação**

Restringe-se, no PAPuQ, o elemento Data do padrão Dublin Core para receber, primariamente, informação que se refira à Data de Publicação. Como especificado mais adiante neste trabalho, no entanto, trata-se de um campo de valor obrigatório, devido a sua relevância para fins de identificação de um recurso bibliográfico. Por isso, entendemos que

certa flexibilização para inclusão desse valor é importante, uma vez que, nem sempre, essa informação consta de uma publicação – mormente as mais antigas.

Dessa maneira, a indicação de uma data de publicação presumida parece-nos aceitável. Espera-se que aplicações sejam desenvolvidas com mecanismos que possibilitem diferenciar esses dois tipos de dados, por meio de especificação presente no documento XML Schema do esquema de metadados intermediário sugerido.

### **Editora**

Assim como determina a semântica do elemento Publicador do padrão Dublin Core, a propriedade Editora, do PAPuQ, denota a organização responsável pela produção de um item.

### **País de origem**

Embora a semântica do elemento Cobertura do padrão Dublin Core faça menção ao *conteúdo* do recurso na atribuição de uma sua extensão temporal ou espacial, não interpretamos como uma definição de temática do recurso, que se pautar por essas extensões, mas à indicação dos fatores temporais ou espaciais que exercem influência sobre a formulação desse conteúdo, quando de sua origem.

No PAPuQ, esse elemento Dublin Core restringe-se à indicação espacial e equivale ao local de origem do recurso como suporte físico, uma vez que Perfil de Aplicação descreve basicamente, um exemplar, e não a obra que lhe deu origem.

Embora, tradicionalmente, a indicação geográfica na catalogação de um recurso bibliográfico relacione-se primariamente ao endereço da sede de uma Editora – comumente designando a cidade –, propõe-se generalizar o valor do campo para país, entendendo que esse tipo de descrição é mais interessante para fins de descoberta de recursos em um sistema de informação global, além de possibilitar termos de busca mais intuitivos.

### **Idioma**

Assim como determina a semântica do elemento Linguagem do padrão Dublin Core, o valor da propriedade Idioma, no PAPuQ, corresponde ao idioma em que o conteúdo do recurso é expresso.

Não se deve confundir o idioma em que o conteúdo de um recurso é expresso, com o idioma usado para catalogação do recurso (que pode ser indicado, por exemplo, pelo idioma dos metadados).

É interessante lembrar para as propriedades contendo *literais*, conforme a terminologia do Modelo Abstrato Dublin Core, é possível indicar o idioma referente a seu valor – o que na linguagem XML é realizado por meio do atributo XML:lang, conforme observado na figura 2.

### **Formato**

Assim como determina a semântica do elemento Formato do padrão Dublin Core, a propriedade Formato, do PAPuQ, refere-se à inserção de dados que caracterizem o recurso fisicamente.

Orienta-se o uso desse campo para um conjunto de três espécies de dados, retiradas do próprio item, de forma explícita ou implicitamente indicadas: número de páginas, ou extensão; tamanho, ou dimensão; e tipo de coloração.

### **Direitos autorais**

Analogamente a semântica do elemento Direito do padrão Dublin Core, a propriedade Direitos Autorais é usado para uma indicação referente aos direitos autorais de uma publicação em quadrinhos.

### **Código da publicação**

No padrão Dublin Core, o elemento Relação é indicado para estabelecer uma ligação entre dois recursos.

Esse elemento poderia ser explorado de inúmeras maneiras para descrição de um recurso bibliográfico, mas procurando focar na captação de dados que tenham um maior potencial na identificação de uma publicação em quadrinhos, aliado a um maior potencial de consistência de valores, optamos por utilizar esse metadado Dublin Core na indicação de ISSN e ISBN.

ISBN (*International Standard Book Number*) é um código numérico de dez ou (a partir de 2007) treze dígitos usados para identificar livros e outros produtos relacionados (como audiolivros). ISSN (*International Standard Serial Number*) é um código numérico de oito dígitos usado para identificação de publicações seriadas. Ambos os códigos foram definidos por norma ISO nos anos 1970 e são gerados por uma agências nacionais.

O ISBN é um ponto de acesso útil para recuperação de publicações específicas, não sendo adequado para descoberta de recurso. A um valor ISSN, no entanto, corresponde um conjunto continuado de publicações, possibilitando que, por meio dele, levantem-se todos os exemplares existentes de uma mesma série – incluindo-se aí aqueles até então desconhecidos pelo usuário.

### **Notas**

Analogamente à semântica do elemento Descrição do padrão Dublin Core, a propriedade Notas é usada para inserção de dados conter dados, em linguagem natural, considerados relevantes para identificação do item descrito, que não couberam nos demais campos.

Não há uma estrutura definida para essa propriedade, e portanto seu potencial informativo num contexto global é, de uma forma geral, reduzido. No entanto, entende-se que se trata de um expediente importante em um sistema de informação de caráter generalista, que pode favorecer, ainda que modestamente, um conhecimento maior sobre o item por parte do usuário do sistema.

### **Colecionador e informações de contato**

O elemento Fonte do padrão Dublin Core foi definido para identificação do proprietário do item descrito numa instância de registro PAPuQ, assim como indicações de meios para contatá-lo. Ao contrário do código de identificação do item, os valores dessa propriedade de identificação devem ser fornecidos pelo catalogador.

A identificação do catalogador pode ser expressa por um valor literal, em linguagem natural, constando, por exemplo, do nome do indivíduo ou organização que possui o item descrito. Porém, a fim de permitir o processamento automatizado das informações de contato, por parte de aplicações interessadas nesses dados, recomenda-se que entradas para e-mail, número de telefone e *website* sejam estruturadas de acordo com padrões internacionais.

### **Tipo**

Como forma de promover, em mais uma frente, o alinhamento às boas práticas para metadados, definidas pela Iniciativa de Metadados Dublin Core, convencionou-se aplicar os valores Texto e Imagem ao elemento Tipo.

Avanços no desenvolvimento de implementações do PAPuQ podem ser realizadas com a escolha ou criação de um vocabulário de gêneros para publicações em quadrinhos, que sirvam como conteúdo para o elemento Tipo.

### **Código do item**

No PAPuQ, o elemento Identificador é reservado para conter uma identificação única do item dentro do contexto de uma sua implementação.

A forma idealizada para gerar esse código prevista nas implementações PAPuQ consiste na combinação de um identificador único do colecionador, obtido no momento de sua inscrição em um Sistema de Informação, e um número seqüencial de entrada do registro nesse sistema.

### **Restrições estruturais das propriedades**



O Perfil do Conjunto Descritivo compõe-se não apenas das propriedades escolhidas para descrição dos recursos do sistema, mas também da definição de suas condições de uso, em termos de cardinalidade e restrição de valores.

Cardinalidade diz respeito ao número de vezes que uma propriedade pode aparecer. O uso de uma propriedade na descrição de um recurso pode ser *opcional*, *mandatório* ou *condicional*. Além disso, seu uso pode ou não ser *repetível*.

Restrição de valores se refere à limitação dos valores atribuídos às propriedades, a um grupo específico. Esse limite pode ser manifesto pela determinação de uso dos termos de um vocabulário controlado.

O formato dos registros que as bases de dados componentes de um Sistema de Informação Federado pretendem compartilhar deve conformar-se às condições definidas pelo Perfil de Aplicação utilizado para agrupar esses registros.

O quadro 8 apresenta dos elementos Dublin Core e as indicações de obrigatoriedade, repetição e restrições de valor de cada elemento, no contexto do PApUQ: uma propriedade classificada como mandatória, é aquela considerada imprescindível para identificação de um exemplar. Embora, de um modo geral, todas as propriedades sejam, para isso, relevantes, nem sempre é possível obter todos os dados a elas correspondentes – as propriedades classificadas como opcionais antecipam-se a esse fato.

<b>Elemento</b>	<b>Obrigatoriedade</b>	<b>Repetição</b>	<b>Restrição de Valor</b>
<b>Dublin Core</b>			
Identificador	Mandatório	Não repetível	Código gerado para cada registro do Sistema Federado
Título	Mandatório	Repetível	Texto (alfanumérico)
Assunto	Opcional	Repetível	Texto (alfanumérico)
Criador	Opcional	Repetível	Texto (alfanumérico) com formatação convencional sugerida
Contribuidor	Opcional	Repetível	Texto (alfanumérico) com formatação convencional sugerida
Publicador	Mandatório	Repetível	Texto (alfanumérico)
Cobertura	Opcional	Repetível	Termos extraídos de vocabulário controlado
Linguagem	Mandatório	Repetível	Termos extraídos de vocabulário controlado
Data	Mandatório	Não repetível	Formato estruturado de data
Formato	Opcional	Repetível	Texto (alfanumérico) com formatação convencional sugerida
Direitos	Opcional	Repetível	Texto (alfanumérico)
Relação	Mandatório se aplicável	Repetível	Identificador oriundo registro internacional
Tipo	Mandatório	Repetível	Termos extraídos de vocabulário controlado
Fonte	Mandatório	Repetível	Texto (alfanumérico)
Descrição	Opcional	Repetível	Texto (alfanumérico)

**Quadro 8** – Equivalência entre propriedades PAPuQ e elementos Dublin Core com suas restrições estruturais

### **Moldes de descrição e de declaração**

As informações sumarizadas no quadro TT conformam uma instância válida do Perfil de Aplicação, ou seja, em nosso caso, elas definem o formato de um registro bibliográfico PAPuQ. No entanto, elas devem ser especificadas mais detalhadamente para possibilitar que essas instruções sejam inequivocamente transpostas para uma linguagem processável por máquina – no caso da implementação aqui sugerida para o PAPuQ, a

linguagem de marcação XML. A inserção do catálogo bibliográfico em meio digital demanda, assim, um trabalho interdisciplinar.

À medida que as tecnologias da informação foram sendo criadas, disponibilizadas e aperfeiçoadas, os sistemas de representação e recuperação de informações documentais assistiram a uma extrapolação dos limites dos tradicionais catálogos referenciais em fichas, alcançando as bases de dados online. Essa nova situação resultou em uma grande mudança no ambiente dos sistemas de representação, assistindo-se a um gradual envolvimento de bibliotecários com profissionais oriundos de outros campos de conhecimento, destacando-se dentre estes o pessoal da ciência da computação e da lingüística e também os produtores comerciais de bases de dados, nos mais diversos campos de conhecimento (ALVARENGA, 2001)

O Perfil do Conjunto Descritivo descreve a estrutura de um conjunto descritivo, que por sua vez compõem um registro, em conformidade com as relações representadas Modelo Abstrato Dublin Core.

Na abordagem DCMI, um registro de metadados é baseado no Modelo do Conjunto Descritivo (ele mesmo parte do Modelo Abstrato DCMI), e o desenho do registro é detalhado no Perfil do Conjunto Descritivo [ou DSP – *Description Set Profile*] usando uma linguagem de restrição DSP. Para cada Descrição e Declaração de um registro, o DSP define um molde [*template*], e cada molde contém restrições relevantes especificando detalhes técnicos como repetição dos elementos ou restrições em valores permitidos. (COYLE, 2009)

Existem dois níveis de moldes em um Perfil do Conjunto Descritivo:

*Molde de Descrição*: contém todos os Moldes de Declaração de uma descrição, bem como as restrições sobre o recurso descrito

*Molde de Declaração*: contém as restrições referentes a uma propriedade e ao valor a ela atribuído, que constituem uma declaração.

A especificação conferida pelos Moldes de Descrição e Declaração do PAPuQ estão anexadas neste trabalho.

## ***Crosswalk* e o esquema de metadados intermediário**

Um sistema de coleta de metadados (*harvesting*) apresenta um terminal que concentra os metadados das bases de dados independentes e frequentemente heterogêneas, que participam de uma federação.

Para que os metadados sejam recebidos por esse agregador, de modo a abastecer “um repositório global com eles e oferecendo uma interface única de busca” (FERREIRA; SOUTO, 2006:27) é preciso que haja um trabalho prévio de homogeneização dos esquemas de metadados de repositórios distintos.

Reconciliar metadados criados em diferentes ambientes é um desafio maior entre alguns esforços que tem sido realizados para mapear elementos de metadados equivalentes entre diferentes esquemas de metadados. Esses mapeamentos podem ser exibidos como tabelas e são conhecidos como *crosswalks*. (CHUDAMANI:174)

*Crosswalk* é "um mapeamento de elementos, semântica e sintaxe de um esquema de metadados para outro" (NISO, c2004:11). Trata-se do procedimento de “tradução” que suporta a habilidade de motores de busca pesquisar o conteúdo de bases de dados distribuídas.

O Perfil de Aplicação de Publicações em Quadrinhos é, no contexto de sua implementação, o esquema de metadados “destino” em relação aos esquemas de metadados “origem” das bases de dados componentes do Sistema Federado sugerido neste trabalho; ou seja, os modelos de dados das bases participantes do sistema devem ser traduzidos para o PAPuQ.

Em um estudo sobre as dificuldades no emprego dos quinze elementos cerne Dublin Core na descrição de recursos, Park e Childress (2009:736) apontam que uma das mais salientes características desse esquema de metadados concerne a sua simplicidade e abrangência, apropriadas para facilitar descrição e descoberta de recursos de qualquer domínio. Porém, concluem os autores, se por um lado tal característica propicia sua adoção em larga escala, tornando-o um dos padrões de metadados mais usados; por outro lado, comparado com outros padrões de metadados, falta riqueza de semântica ao Dublin Core.

Essas dificuldades de uso encontradas, que são devidas a ambigüidades conceituais e sobreposições semânticas apresentadas pelos metadados, guardam analogia com as dificuldades para tradução dos esquemas locais, com campos de dados mais específicos, para um esquema reduzido aos quinze elementos cerne do padrão Dublin Core. Essa busca pela equivalência entre esquemas de metadados que apresentam níveis de detalhamento distintos em seus modelos de dados ainda é, e talvez sempre seja, uma atividade humana essencial na promoção da interoperabilidade entre sistemas de informação.

Atualmente, o mapeamento de esquemas de metadados visando interoperabilidade ainda requer um enorme esforço intelectual, mesmo com toda a assistência que a tecnologia computacional pode prover. Ainda assim, é claro que a tecnologia tem provido muitas possibilidades para facilitar e acentuar esforços de interoperabilidade semântica de metadados. O que ainda é preciso são as preparações intelectuais para apoiar o objetivo da interoperabilidade e ferramentas práticas para levar a cabo tais atividades. Se a comunidade da informação busca prover um acesso otimizado para toda informação disponível através de bibliotecas digitais e repositórios, profissionais da informação devem dar alta prioridade para a tarefa de criar e manter o mais alto nível de interoperabilidade entre serviços de informações novos e correntes. (ZENG; CHAN, 2006)

Segundo Chan e Zeng (2006), duas abordagens têm sido usadas: o *crosswalking absoluto*, em que se requer o mapeamento exato entre os elementos envolvidos (se a equivalência não for total, não há *crosswalk* possível para o elemento de origem); e o *crosswalking relativo*, em que todos os elementos de um esquema de origem são mapeados para o esquema de destino, ainda que a equivalência semântica não seja idêntica. "A abordagem de *crosswalking* relativo parece trabalhar melhor quando se mapeia de um esquema complexo para um esquema simples (ex. do MARC para DC, mas não vice versa)". (CHAN; ZENG, 2006)

A vantagem dessa segunda abordagem sobre a primeira reside na efetivação da transferência de todos os dados, que se dependessem de uma equivalência semântica completa para compor o novo esquema, seriam deixados para trás. Na tradução de um esquema de metadados com uma semântica mais rica, para um esquema de metadados de semântica mais genérica, no entanto, pode ocorrer interpretação variada sobre o significado das propriedades, causando inconsistências nos resultados unificados de busca.

Como uma forma de atenuar a possibilidade desse tipo de inconsistência na tradução dos esquemas das bases de dados componentes do Sistema de Informação Federado, foi elaborado um esquema de metadados intermediário, nomeado Pré-PAPuQ.

Trata-se de um esquema de metadados para configuração de um registro mais detalhado, oferecendo opções de equivalência semântica potencialmente mais próximas àquelas dos campos encontrados nas bases independentes. O quadro abaixo exemplifica o processo de mapeamento para os esquemas de metadados Pré-PAPuQ e PAPuQ de campos do registro apresentado na figura 17, a partir de sua codificação em formato MARC (conforme figura 18).

MARC	Pré-PAPuQ	PAPuQ	Dado
20\$a	ISBN	Código da Publicação	8533610610
100\$a	Nome	Criador	Quino
245\$a	Título	Exemplar	Mafalda vai embora /
260\$c	Ano	Data de Publicação	1999
260\$b	Editora	Editora	Martins Fontes,
300\$c	Tamanho	Formato	29cm.
300\$a	Páginas	Formato	48p.
300\$b	Coloração	Formato	il. col.

Quadro 9 – Mapeamento de registro em formato MARC para PAPuQ

<b>Autor:</b>	Quino, 1932-
<b>Título / Barra de autoria:</b>	Mafalda vai embora / Quino ; [coordenação da tradução e texto final Monica Stahel]. -
<b>Imprenta:</b>	São Paulo : Martins Fontes, 1999.
<b>Descrição física:</b>	48p. : il. col. ; 29cm. -
<b>Série:</b>	(Álbuns da Mafalda ; 11)
<b>Notas:</b>	Texto em quadrinhos.
<b>ISBN:</b>	8533610610 (enc.)
<b>Assuntos:</b>	Histórias em quadrinhos.
<b>Classificação Dewey:</b>	741.5
<b>Edição:</b>	19
<b>Indicação do Catálogo:</b>	IV-2,4,39/INF F
<b>Registro Patrimonial:</b>	985.748 DL 13/01/2000
<b>Sigla do Acervo:</b>	DRG
<input type="button" value="Nova Busca"/> <input type="button" value=" &lt;&lt;"/> <input type="button" value="&lt;&lt;"/> <input type="button" value="&gt;&gt;"/> <input type="button" value="&gt;&gt; "/> <input type="button" value="&gt;"/> <input type="button" value="Ficha"/> <input type="button" value="Marc"/> <input type="button" value="Lista"/> <input type="button" value="Índices"/> <input type="button" value="Avançada"/> <input type="button" value="Análise"/>	

Figura 17 – Interface de usuário de registro pertencente ao catálogo online da Biblioteca Nacional

```

LDR 00694cam0022002177 4500
001 100011117005389063
003 Br
005 20040914121555.8
008 100011s1999 spba 000 0 por d
020 __ |a 8533610610 (enc.)
082 04 |2 19 |a 741.5
092 __ |a IV-2,4,39/INF |d F
100 0_ |a Quino, |d 1932-
245 10 |a Mafalda vai embora / |c Quino ; [coordenação da tradução e texto final Monica Stahel]. -
260 __ |a São Paulo : |b Martins Fontes, |c 1999.
300 __ |a 48p. : |b il. col. ; |c 29cm. -
440 _1 |a (Álbuns da Mafalda ; |v 11)
500 __ |a Texto em quadrinhos.
650 04 |a Histórias em quadrinhos.
852 __ |a DRG
949 __ |a 985.748 DL 13/01/2000

```

**Figura 18** – Registro pertencente ao catálogo *online* da Biblioteca Nacional em formato MARC

A intermediação desse esquema de metadados estendido também objetiva servir de meio para favorecer a inclusão de regras de transformação de registros que garantam a uniformidade na distribuição dos dados coletados, em sua transferência para as propriedades do PAPuQ.

Como exemplo, podemos citar o caso de publicações seriadas em quadrinhos, como o os gibis, cuja correta identificação exigirá que também seja inserido, juntamente com o título da revista, sua indicação de seqüência (número, volume, ou ambos). Isso ocorre porque, entre os elementos Dublin Core, não existe um que seja apropriado a representar essa informação, em nossa interpretação.

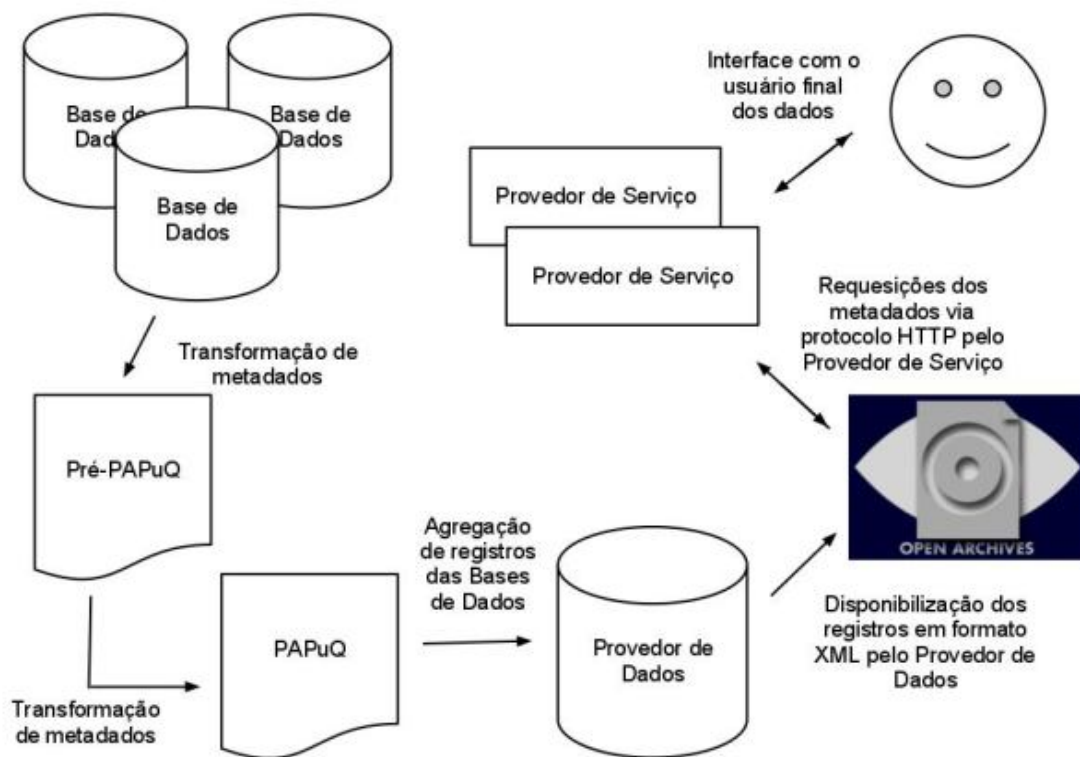
Assim, o fornecimento de campos para receber as informações de número e volume, no esquema de metadados Pré-PAPuQ, possibilita a criação de regras de redistribuição dos dados quando de sua transferência para os elementos cerne Dublin Core – essas regras seriam executadas por um aplicativo.

Como ilustração, suponhamos a existência de três repositórios de publicações em quadrinhos, *A*, *B* e *C*, que insiram os dados referentes ao seqüenciamento de publicações seriadas de formas distintas em suas bases de dados, como mostrado no quadro 10.

	ESQUEMA ORIGINAL		ESQUEMA Pré-PAPuQ		ESQUEMA PAPuQ	
	Propriedade	Valor	Propriedade	Valor	Propriedade	Valor
<b>A</b>	<i>Título</i>	Batman, 12, v.3	<i>Título</i>	Batman	<i>Exemplar</i>	Batman (3) #12
			<i>Parte</i>	3		
			<i>Parte da parte</i>	12		
<b>B</b>	<i>Nome</i>	Batman	<i>Título</i>	Batman	<i>Exemplar</i>	Batman (3) #12
	<i>Edição</i>	3/12	<i>Parte</i>	3		
			<i>Parte da parte</i>	12		
<b>C</b>	<i>Revista</i>	Batman	<i>Título</i>	Batman	<i>Exemplar</i>	Batman (3) #12
	<i>Número</i>	12	<i>Parte</i>	3		
	<i>Volume</i>	3	<i>Parte da parte</i>	12		

**Quadro 10** – Sucessivos mapeamentos de metadados para padronização dos valores de registros

O esquema de metadados Pré-PAPuQ é visto com mais detalhes na próxima seção. No âmbito da implementação do Perfil de Aplicação proposto neste trabalho, as transformações dos modelos de dados locais para o esquema de metadados Pré-PAPuQ, e deste último para o modelo de dados do PAPuQ é representado na pela figura abaixo.



**Figura 19** – Plano de transformações de esquemas de metadados da implementação de base de dados proposta



Resguardadas oportunidades para seu aprimoramento e com o mapeamento de suas propriedades para metadados devidamente declarados em *namespaces*, o Pré-PAPuQ pode servir, no âmbito da Iniciativa de Arquivos Abertos, como um esquema de metadados mais rico para uso de Provedores de Serviços que reconheçam esses *namespaces*, oferecendo interfaces mais especializadas que a possibilitada pelo uso do padrão Dublin Core Simple.

### **3.4.4 Diretrizes de uso**

Diretrizes de uso oferecem instruções para criação de registros de metadados ou, no caso da implementação sugerida para o PAPuQ, para mapeamento dos registros existentes em repositórios para o Sistema de Informação Federado.

Um Perfil do Conjunto Descritivo define o “o que” de um Perfil de Aplicação; as Diretrizes de Uso provêm o “como” e o “porque”. Idealmente, eles explicam cada propriedade e antecipam as decisões que devem ser feitas no curso da criação de um registro de metadados. (COYLE, 2009)

Essa documentação para criadores de metadados apresenta algumas das mesmas informações encontradas no Perfil do Conjunto Descritivo, mas em uma forma mais amigável para o leitor humano.

As regras numeradas na próxima seção orientam para o mapeamento dos registros das bases de dados componentes para as propriedades do esquema de metadados mediador Pré-PAPuQ, que tem a função de realizar uma filtragem e redistribuição dos dados desses registros, para então, de forma mais consistente, serem recebidas no PAPuQ.

Algumas dessas regras podem ser tidas como prescrição de melhores práticas para representação descritiva do PAPuQ, não interferindo efetivamente no processo de conversão de formato do registro; outras representam requerimentos para validação dessa transformação, via, por exemplo, documento XML Schema a ser definido para o esquema de metadados mediador. Com finalidade meramente ilustrativa, por tratar-se da simulação de uma interface humana da inclusão de dados no esquema Pré-PAPuQ, um formulário que define os campos

do modelo do esquema de metadados mediador é dada na figura 20, seguida da indicação de alguns processos a serem executados por aplicações, para promover a filtragem e redistribuição mencionadas.

Trata-se de orientações gerais para padronização de registros bibliográficos, embasadas no conhecimento empírico de um conjunto de características mais frequentemente encontrado em publicações em quadrinhos.

Necessidades de detalhamento na padronização da descrição contida nos registros, podem ser discutidas em fóruns abertos, pelos interessados, como geralmente acontece em iniciativas de catalogação cooperativa, encontradas na Web.

A noção mais importante que o implementador de um Perfil de Aplicação deve ter se refere à semântica da terminologia adotada no esquema de metadados mediador do PAPuQ, como apresentada no quadro 11.

<b>Campos Pré-PAPuQ</b>	<b>Descrição</b>
Título	Transcrição do título do exemplar.
Subtítulo	Expressão ou palavra que se segue ao Título, complementando-o.
Parte	Primeiro ou único grau de divisão do recurso no contexto de uma série de publicações a qual pertença.
Parte da Parte	Segundo grau de divisão do recurso no contexto de uma série de publicações a qual pertença.
Personagem	Agente que figura a narrativa apresentada no recurso.
Nome	Nome completo do responsável intelectual pelo recurso.
Primeiro Nome	Primeiro nome do responsável intelectual pelo recurso.
Sobrenome	Sobrenome do responsável intelectual pelo recurso.
Roteirista	Agente responsável por escrever a história.
Ilustrador	Agente primariamente responsável pelas ilustrações.
Colorista	Agente responsável por prover cores às ilustrações.
Balonista	Agente responsável por inserir as palavras nos balões e nas legendas de uma história em quadrinhos.
Arte-finalista	Agente responsável pelo aprimoramento de ilustrações esboçadas pelo ilustrador.
Tradutor	Agente responsável por traduzir uma história.
Editor	Agente responsável por controlar o conteúdo e direção de uma história em quadrinhos.
Compilador	Agente responsável pela escolha e organização de histórias de variadas fontes para serem publicadas em uma obra.
Colaborador	Agente que de alguma (outra) maneira contribua para produção de uma publicação.
Editora	Instituição responsável pela publicação do recurso.
Código do País	Código do país de origem do recurso.
Código do Idioma	Código do idioma em que o recurso é expresso.
Data de Publicação	Data associada ao evento de publicação do recurso.
Número de Páginas	Indicação do número de páginas de um recurso.
Tamanho	Indicação da altura de um recurso.
Coloração	Indicação do tipo de coloração de um recurso.
Direitos Autorais	Declaração de direitos autorais do recurso.
Código da Publicação	Código normalizado correspondente a ISBN ou ISSN.
ISSN	Código internacional normalizado para publicações seriadas.
ISBN	Código internacional normalizado para publicações monográficas.
Tipo	Categoria em que o recurso pode ser alocado.
Colecionador	Indivíduo ou instituição que dispõe do recurso.
Endereço	Endereço do indivíduo ou instituição que dispõe do recurso.
Telefone	Telefone do indivíduo ou instituição que dispõe do recurso.
E-mail	<i>E-mail</i> do indivíduo ou instituição que dispõe do recurso.
Website	<i>Website</i> do indivíduo ou instituição que dispõe do recurso.
Notas	Observações acerca do item, que auxiliem em sua identificação.

**Quadro 11** – Semântica das propriedades do esquema de metadados Pré-PAPuQ

## Instruções de mapeamento

As instruções de mapeamento a seguir são dadas para cada campo do esquema de metadados Pré-PAPuQ.

### **Agentes de criação de conteúdo do recurso**

Existem duas formas de mapeamento de dados referentes aos responsáveis intelectuais pelo conteúdo de um recurso: pelo uso do campo Nome, ou pelo uso combinado dos campos Primeiro Nome e Sobrenome.

Prefira realizar a inserção dos dados pelo uso combinado dos campos Primeiro Nome e Sobrenome, pois, dessa maneira, uma aplicação pode se encarregar da estruturação do valor na transferência desses dados para as propriedades Criador ou Contribuidor do PAPuQ.

De outra maneira, melhor prática recomendada para uso do campo Nome é formatar seu valor com a fórmula *SOBRENOME, PRIMEIRO NOME*, garantindo a uniformidade entre registros, e facilitando a ordenação alfabética dos nomes dos agentes.

O Pré-PAPuQ também fornece meios de indicar a exata função de um Criador ou Contribuidor. Esses dados não podem ser transferidos para o PAPuQ, pois devido a sua conformidade com o padrão Dublin Core Simple, o refinamento de uma propriedade não é possível. No entanto, esse mapeamento é um meio de manter essa especificidade, tendo em mente a possibilidade de usos alternativos futuros para os dados contidos no Pré-PAPuQ.

A classificação dessas funções é dada por um subconjunto de siglas extraídas do vocabulário controlado de funções de responsabilidade *MARC Code List for Relators*, desenvolvido e mantido pela Biblioteca do Congresso nos Estados Unidos.

<b>Função no Pré-PAPuQ</b>	<b>Função no MARC Code List for Relators</b>	<b>Código MARC Code List for Relators</b>
Roteirista	<i>Author</i>	aut
Ilustrador	<i>Illustrator</i>	ill
Colorista	<i>Colorist</i>	clr
Balonista	<i>Calligrapher</i>	cil
Arte-finalista	<i>Draftsman</i>	drm
Tradutor	<i>Translator</i>	trl
Editor	<i>Editor</i>	edt
Compilador	<i>Compiler</i>	com
Colaborador	<i>Contributor</i>	ctb

**Quadro 12** – Equivalência entre designações de responsabilidade do Pré-PAPuQ e do *MARC Code List for Relators*

### **Título e Subtítulo**

Os valores das propriedades Título e Subtítulo são as transcrições do título e do subtítulo de exemplar, como se encontra na publicação.

Como melhor prática, recomenda-se que, no caso de publicações monográficas, essa transcrição corresponda ao título presente na página de rosto do item; no caso de revistas em quadrinhos, que se transcreva esses dados como apresentados na seção de créditos (ou *indícia*).

### **Parte e Parte da Parte**

Como visto, no PAPuQ, um item representa o exemplar de uma publicação em quadrinhos, independentemente da série à qual porventura pertença. No entanto, muitas vezes, a discriminação da posição que um exemplar ocupa dentro de uma série é necessária sua correta identificação.

O PAPuQ prevê até dois níveis de graduação que demarcam a posição de um item dentro de uma série que o compreenda: Parte e Parte da parte.

Valores que correspondam a Parte e Parte da parte podem ser expressos por variada nomenclatura. No caso de gibis, esses valores equivalem, usualmente, a Volume e Número, em publicações monográficas podem ser representados ainda por termos como Tomo ou Parte.

De qualquer maneira, para fins de mapeamento, o mais importante é observar a hierarquia na designação desses valores, garantindo a uniformidade entre registros.

Buscando favorecer a internacionalidade em implementações do Perfil de Aplicação de Publicações em Quadrinhos, uma boa prática recomendada é a expressão desses níveis de graduação preferencialmente por números arábicos.

### **Personagem**

Valores para personagens não são estruturados de qualquer maneira.

Melhor prática recomendada é repetir o campo para cada uma das entradas de personagem, a fim de facilitar buscas e ordenação de registros que sejam realizadas por meio dele.

### **Editora**

Valores de Editora não são estruturados de qualquer maneira.

Melhor prática recomendada é repetir o campo para cada uma das entradas de editora, e que tal valor seja, preferencialmente, extraído da seção de crédito do item bibliográfico – a *indicia* –, no caso de revistas em quadrinhos; e da página de rosto no caso de uma publicação monográfica.

### **Código do País**

O valor referente ao local de publicação deve ser estruturado conforme norma ISO 3166 – Alpha-3, que estabelece um código de três caracteres alfabéticos para cada país listado.

Repositórios que tenham essa informação registrada de outra maneira devem realizar a adaptação dos valores presentes em suas bases para os respectivos códigos de país, antes do mapeamento para o Pré-PAPuQ.

### **Data de Publicação**

O valor referente à data de publicação deve ser estruturado conforme formato ISO 8601 ou padrão W3CDTF, criado como uma simplificação dessa norma ISO para o ambiente Web. Uma aplicação se encarregaria da estruturação do valor na transferência dos dados para o elemento Data de Publicação do PAPuQ.

Três níveis de detalhamento podem ser aplicados à data de publicação: *ano, mês e ano*, ou *dia, mês e ano*.

Uma especificação de data deve ser decomposta em seus elementos dia, mês e ano a fim de ser corretamente mapeada para os respectivos campos do esquema de metadados Pré-PAPuQ.

### **Número de páginas, Tamanho e Coloração**

De maneira a prover valores padronizados, ao menos no contexto de um idioma de catalogação específico, às indicações de Número de Páginas, Tamanho e Coloração, optou-se por adotar a estruturação de descrição física de itens bibliográficos postulada pela norma AACR2. Dessa maneira:

O número de páginas de uma publicação é indicada em números arábicos, seguida, em descrições realizadas em língua portuguesa, pela abreviação “p.”. Exemplo: *60 p.*

A dimensão de uma publicação é indicada pela medida da altura do item, em centímetros, seguida, em descrições realizadas em língua portuguesa, pela abreviação “cm”. Exemplo: *19 cm.*

O tipo de coloração é limitado a dois tipos: *colorido* e *preto e branco*. Em descrições realizadas em língua portuguesa, esses valores devem ser indicados pelas formas *color.* e *p&b*, respectivamente.

### **Código do Idioma**

O valor referente ao campo Idioma deve ser estruturado conforme norma ISO 639 – Alpha-2, que consiste de uma listagem<sup>37</sup> de códigos de 3 caracteres alfabéticos para identificação das línguas.

Os códigos de linguagem foram delineados para uso por bibliotecas, serviços de informação e publicadores para indicar linguagem na troca de informação, especialmente em sistemas computadorizados. Esses códigos têm sido largamente usados na comunidade bibliotecária e podem ser adotados por quaisquer aplicações que requeiram a expressão codificada de linguagem por terminologias e lexicografias. O conjunto de códigos *alpha-2* foi delineado para uso prático para a maior parte das linguagens do mundo que são freqüentemente representadas no corpo total da literatura do mundo. (LIBRARY OF CONGRESS, 2011)

Repositórios que tenham essa informação registrada de outra maneira devem realizar a adaptação dos valores presentes em suas bases para os respectivos códigos de linguagem, antes do mapeamento para o Pré-PAPuQ.

O campo pode ser repetido para múltiplos valores, como nos casos em que o conteúdo de uma publicação é bilíngüe.

### **Direitos**

Melhor prática definida para este campo consiste na transcrição da declaração *copyright* impressa no exemplar, usualmente consistindo dos dados ano/detentor do direito, acompanhados do símbolo de *copyright* ©.

---

<sup>37</sup> Codes for the representation of names of languages (Library of Congress). Disponível em: [http://www.loc.gov/standards/iso639-2/php/English\\_list.php](http://www.loc.gov/standards/iso639-2/php/English_list.php). Acesso em novembro de 2011.



A propriedade pode ser repetida para o caso de múltiplas ocorrências, como quando há uma declaração de *copyright* para o texto e outra para as ilustrações, ou para casos em que estão presentes tanto a declaração de *copyright* da publicação original, como a declaração de *copyright* da versão traduzida da obra.

Classificações gerais do tipo de licença referentes a uma obra, como *Creative Commons*, *Licença Pública Geral*, *Copyright* ou *Copyleft* são preteridas em favor da reprodução literal da informação contida no item, como forma de prover o preenchimento do campo maneira mais objetiva, evitando a necessidade da determinação de um vocabulário específico para garantia da consistência dos valores dessa propriedade.

### **Código da Publicação**

ISBN e ISSN são códigos atribuídos a categorias de publicação auto-excludentes, não sendo permitido, então, que uma mesma publicação receba ambos os tipos de código. Porém, algumas vezes, publicações realizadas por parcerias de Editoras recebem um código referente a cada uma delas. Nesses casos, a propriedade pode ser repetida.

Valores referentes a ISBN e ISSN devem ser inseridos sem pontuações.

### **Notas**

Valores de Notas não são estruturados de qualquer maneira.

Informações adicionais sobre o exemplar podem se referir dados como preço, no caso das revistas em quadrinhos, ou número de edição, no caso de monografias.

### **Colecionador e Endereço**

Valores de Colecionador e Endereço não são estruturados de qualquer maneira.

Pelo menos um valor que identifique a entidade que dispõe do exemplar descrito é obrigatória. Assim, melhor prática recomendada para o campo Colecionador é identificá-lo de

forma clara e sucinta. Mas essa identificação também pode ser fornecida pelos campos Endereço, Telefone, E-mail ou Website.

### Telefone, E-mail e Website

Tecnicamente, valores dos campos Telefone, E-mail e Website são *valores string*, conforme terminologia do Modelo Abstrato Dublin Core (cf. anexo).

Porém, a fim de permitir o processamento automatizado das informações de contato, por parte de aplicações interessadas nesses dados, recomenda-se que entradas para e-mail, número de telefone e *website* sejam estruturadas por meio de URIs, seguindo sintaxe definida por RFCs, que especificam um formato para cada um desses tipos de dados, como pode ser observado no quadro 13.

Forma de Sintaxe contato	Exemplo	Norma
Telefone	“tel”:+código global(código local) número de telefone	tel:+55(11)9469-6664 RFC 3966 <sup>38</sup>
E-mail	“mailto”:endereço-de-email	mailto:salenka@gmail.com RFC 2368 <sup>39</sup>
Website	“URL”:protocolo:endereço-do-site	URL:http://www.bn.br RFC 1738 <sup>40</sup>

Quadro 13 – Sintaxe para formas de contato Pré-PAPuQ

### Tipo

A princípio, o uso do campo Tipo é reservado para os termos Texto e Imagem do vocabulário de tipos Dublin Core (DCMIType). A repetição do campo é requerida para comportar cada um desses termos.

<sup>38</sup> Request for Comments 3966 – Esquema Tel. Disponível em: <http://tools.ietf.org/html/rfc3966>

<sup>39</sup> Request for Comments 2368 – Esquema Mailto. Disponível em: <http://tools.ietf.org/html/rfc2368>

<sup>40</sup> Request for Comments 1738 – Esquema URL. Disponível em: <http://www.ietf.org/rfc/rfc1738.txt>

### Pré-PAPuQ

Tipo   Repetir campo

Título   Repetir campo

Subtítulo

Parte  Parte da Parte

Editora   Repetir campo

Data de Publicação     Data presumida  
Dia Mês Ano

Código do Idioma   Repetir campo

Código do país

Código da Publicação    Repetir campo

Número de páginas  Altura

Coloração

Personagem   Repetir campo

Direitos Autorais   Repetir campo

Nome   Repetir campo

Primeiro Nome  Sobrenome

Função criativa

Notas   Repetir campo

---

Colecionador

Endereço

Telefone   Repetir campo

E-mail   Repetir campo

Website   Repetir campo

**Figura 20** – Campos do esquema de metadados Pré-PAPuQ

Instruções de processamento desejáveis para serem executadas por uma aplicação:

1) Inserção ordenada dos dados de Título, SubTítulo, Parte e Parte da Parte, respectivamente, na composição da propriedade Exemplar do PAPuQ;

2) Inserção ordenada dos dados de Dia, Mês e Ano na composição da propriedade Data de Publicação do PAPuQ, conforme formato ISO 8601 e, por exemplo, pode executar inserção automática no campo Notas de informação sobre presunção de data, se for o caso;

3) Aplicação define mecanismo para que as duas formas de inserção de nome do responsável intelectual sejam auto-excludentes.

4) Redução do valor dos campos Código do Idioma, Código do País e Função Criativa aos códigos correspondentes a idioma, país e função indicados entre parênteses.

### 3.4.5 Orientação para Sintaxe e Formatos de Dados

Um Perfil de Aplicação bem elaborado é sintaticamente neutro. Isso quer dizer que, desde que as relações e os valores definidos em um Perfil sejam completamente expressos, qualquer codificação de sintaxe pode ser empregada para o processamento por máquina dos dados contidos em cada instância.

A Iniciativa de Metadados Dublin Core desenvolveu, até o momento, codificações (*bindings*) para expressar um Perfil de Aplicação pelo uso de XML<sup>41</sup>, RDF<sup>42</sup> e HTML/XHTML<sup>43</sup>. Essas codificações são uma tradução para linguagem de máquina, das características determinadas para formatação de um registro de acordo com a definição das funcionalidades da base de dados que o contém, de seu escopo, de restrições estruturais e vocabulares. No entanto, a tradução dessas representações codificadas demanda um conhecimento especializado, assim como é especializado o conhecimento que provê um levantamento adequado e satisfação das necessidades dos usuários dos dados.

Tal representação extrapolou a alçada exclusiva de profissionais da informação, especialmente os bibliotecários, passando a exigir uma integração de diferentes tipos

---

<sup>41</sup> Guidelines for implementing Dublin Core in XML. *Dublin Core Metadata Initiative*. Disponível em: <http://dublincore.org/documents/dc-xml-guidelines/>

<sup>42</sup> Expressing Dublin Core using the Resource Description Framework (RDF). *Dublin Core Metadata Initiative*. Disponível em: <http://dublincore.org/documents/dc-rdf/>.

<sup>43</sup> Expressing Dublin Core metadata using HTML/XHTML meta and link elements. *Dublin Core Metadata Initiative*. Disponível em: <http://dublincore.org/documents/dc-html/>

de profissionais, entendidos como atores envolvidos no processo mais amplo da organização da informação. Trata-se, efetivamente, de um conjunto de atividades interdisciplinares, colaborativas e descentralizadas, as quais demandam integração de esforços, conhecimentos, e utilização intensiva de tecnologia. (BAPTISTA, 2007)

Na implementação do Perfil de Aplicação para Publicações em Quadrinhos sugerida neste trabalho, instâncias do PAPuQ devem ser geradas em formato XML, como requerido pelo protocolo OAI-PMH. Um documento XML Schema é apresentado em anexo, contendo os elementos e algumas restrições estruturais para validação dos registros, conforme definições do Perfil de Aplicação.

## Conclusão

Um amplo conjunto de transformações, causadas pelo desenvolvimento e uso de tecnologias eletrônicas pela sociedade, tem modificado as noções de tempo e espaço da informação. Essas transformações afetam profundamente a vida de toda sociedade, alterando modos de produção, consumo, trabalho, e mesmo de relacionamentos entre as pessoas.

Diretamente envolvido com essa questão, tal situação tem causado a “mudança no perfil profissional do bibliotecário e determinado a necessidade de integração entre diferentes atores envolvidos no tratamento e organização da informação”. (BAPTISTA, 2007:177) Paradoxalmente, a interdisciplinaridade na organização de produtos do conhecimento torna-se condição para se alcançar esse objetivo, enquanto cada vez mais o sentido desses produtos precisa ser atribuído por meio de abordagens especializadas.

Entre os diferentes atores envolvidos no tratamento e organização da informação, encontra-se o indivíduo comum, o tradicional consumidor da informação, que passa também à condição de produtor, em uma intensidade sem paralelo histórico. Os quinze elementos cerne padrão Dublin Core surgiram do entendimento de que a contribuição desses produtores nessa atividade é imprescindível. E foi com o uso desse conjunto de elementos de metadados que propomos um Perfil de Aplicação, que estrutura a descrição de publicações em quadrinhos.

O crescente reconhecimento social do valor artístico, educacional e antropológico dessa mídia é evidenciado pela instituição de gibitecas e inclusão progressiva dessas publicações em bibliotecas públicas em âmbito nacional (VERGUEIRO, 2004), além da constatação de comunidades de fãs, colecionadores e pesquisadores atuando veementemente em rede, operando na construção de bases de dados especializadas na Web, conotando que se trata de um meio de comunicação com características bastante próprias.

Assim, buscou-se aliar a estruturação provida pelo uso de metadados descritivos a registros disponíveis em rede, de forma dispersa, na sugestão de um repositório referencial de publicações em quadrinhos, para auxiliar a localização desses recursos, por meio do reuso de conteúdos existentes, fazendo uso do protocolo OAI-PMH. Porém, podemos concluir que

mesmo a utilização da tecnologia mais avançada não é garantia de obtenção dos resultados desejados.

Foi antevisto que o OAI-PMH possibilitaria uma larga automatização, uma abordagem de baixo esforço humano para o processo de agregar metadados. No entanto, facilitar o transporte de metadados não é o mesmo que facilitar a geração e uso de metadados de uma ponta a outra no contexto de uma biblioteca digital. A criação e uso de metadados continuam em larga escala um processo mediado pelo humano, e fatores humanos afetam o sucesso do processo. (COLE; FOULONNEAU, 2007:185-6)

A qualidade na representação de recursos, a fim de favorecer necessidades informacionais dos usuários, é objetivo de qualquer Sistema de Informação, seja em ambiente convencional ou digital, buscando prover “agilização e autonomia operacional no acesso à informação”. (BAPTISTA, 2007:187) Em rede, principalmente, onde novos conteúdos são gerados diariamente, em grandes quantidades, é preciso considerar meios e técnicas para otimizar esse processo. A experiência de longa data adquirida por profissionais à frente de bibliotecas e centros de documentação não pode ser desprezada.

Enquanto a internet se apresenta como novas possibilidade de disponibilização e processamento de informações eletrônicas, bibliotecários e especialistas da informação se esforçam para desenvolver métodos para descrição, organização e recuperação de objetos digitais acessados remotamente. (ALVARENGA, 2001)

Para Pombo, a mutação resultante da introdução dessas novas tecnologias sublinham ainda mais o caráter pragmático da classificação documental e biblioteconômica, de “intuitos funcionais imediatos” (POMBO, 1997:10). A autora afirma que “na biblioteconomia como na documentação como em todos os outros domínios, as melhores soluções pragmáticas são ainda, e necessariamente, aquelas que se propõem resolver os problemas de eficiência prática em termos teóricos e conceptuais”. (POMBO, 1997:13)

Tal vocação alude diretamente ao ambiente dinâmico que caracteriza a Internet e a Web, onde essas competências podem e devem contribuir no provimento de sentido a sua incomensurável rede documental.

## Referências

- ALVARENGA, L. A Teoria do Conceito Revisitada em conexão com ontologias e metadados no contexto das bibliotecas tradicionais e digitais. *DataGramaZero* [Periódico online] v. 2, n. 6, dez, 2001. Disponível em: [http://www.dgz.org.br/dez01/Art\\_05.htm](http://www.dgz.org.br/dez01/Art_05.htm). Acesso em: novembro de 2011.
- ALVES, M. D. R.; SOUZA, M. I. F. Estudo de correspondência de elementos metadados: Dublin Core e Marc 21. *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, v. 4, n. 2, jan./jun., 2007, p. 20-38. Disponível em: [http://www.sbu.unicamp.br/seer/ojs/index.php/sbu\\_rci/article/viewFile/358/237](http://www.sbu.unicamp.br/seer/ojs/index.php/sbu_rci/article/viewFile/358/237). Acesso em: novembro de 2011.
- ALVES, R. C. V. *Web Semântica: uma análise focada no uso de metadados*. Local, 2005. Dissertação (Mestrado) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista. Disponível em: [http://www.athena.biblioteca.unesp.br/exlibris/bd/bma/33004110043P4/2005/alves\\_rcv\\_me\\_mar.pdf](http://www.athena.biblioteca.unesp.br/exlibris/bd/bma/33004110043P4/2005/alves_rcv_me_mar.pdf). Acesso em: novembro de 2011.
- ALVES, R. C. V.; SANTOS, P. Metadados em ciência da informação: considerações preliminares sobre padrões para a construção normalizada de representações. In: *Congreso de la CiberSociedad, IV*, 2009, (Congresso online). Crisis analógica, futuro digital. Grupo de trabajo B-10: Visualização de redes e de outros dados. Disponível em: <http://www.cibersociedad.net/congres2009/>. Acesso em outubro de 2011.
- ARAÚJO, V. M. R. H. Sistemas de Informação: nova abordagem teórico-conceitual. *Ciência da Informação*, v. 24, n. 1, 1995. Disponível em: <http://revista.ibict.br/index.php/ciinf/article/view/532/484>. Acesso em novembro de 2011.
- BAKER, T. DCMI Usage Board Review of Application Profiles. *Dublin Core Metadata Initiative*, 2003. Disponível em: <http://dublincore.org/usage/documents/2003/02/11/profiles/>. Acesso em: setembro de 2011.
- BAKER, T. Revisions to DCMI metadata terms. *Dublin Core Metadata Initiative*, 2008. Disponível em: <http://dublincore.org/usage/decisions/2008/dcterms-changes>. Acesso em setembro de 2011.
- BAPTISTA, D. M. O impacto dos metadados na representação descritiva. *Revista ACB: Biblioteconomia em Santa Catarina*, Florianópolis, v. 12, n. 2, p. 177-190, jul./dez., 2007. Disponível em: <http://dialnet.unirioja.es/servlet/oaiart?codigo=2684571>. Acesso em: novembro de 2011.
- BARRETO, A. A. Os Destinos da Ciência da Informação: entre o cristal e a chama. *DataGramaZero* [Periódico online], n. 0 1999. Disponível em: [http://www.dgz.org.br/dez99/Art\\_03.htm](http://www.dgz.org.br/dez99/Art_03.htm). Acesso em: novembro de 2011.
- BARRETO, C. M. *Modelo de metadados para a descrição de documentos eletrônicos na web*. Rio de Janeiro, 1999. Dissertação (Mestrado) Instituto Militar de Engenharia. Disponível em: <http://www.de9.ime.eb.br/dissertacoes/1999-Cassia.pdf>. Acesso em agosto de 2001.
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Semantic Web. *Scientific American*, may, 2001. Disponível em: <http://www.jeckle.de/files/tblSW.pdf>. Acesso em: outubro de 2011.
- CAMPELLO, B. S.; MAGALHÃES, M. H. A. *Introdução ao controle bibliográfico*. Brasília, DC: Brinquet de Lemos, 1997.
- CAPLAN, P. OAI-PMH. *Computers in libraries*; Feb 2004; 24, 2; ABI/INFORM Trade & Industry, p. 24. Disponível em: [http://books.google.com/books?id=azlVAAAAYAAJ&q=oai+pmh+caplan&dq=oai+pmh+caplan&hl=en&ei=qd nMT0W\\_KomLgweGs-zkDQ&sa=X&oi=book\\_result&ct=result&resnum=4&ved=0CEAQ6AEwAw](http://books.google.com/books?id=azlVAAAAYAAJ&q=oai+pmh+caplan&dq=oai+pmh+caplan&hl=en&ei=qd nMT0W_KomLgweGs-zkDQ&sa=X&oi=book_result&ct=result&resnum=4&ved=0CEAQ6AEwAw). Acesso em: julho de 2011.



- CHAN, L. M.; ZENG, M. L. Metadata interoperability and standardization – a study of methodology, part I: achieving interoperability at the schema level. *D-Lib Magazine*, v. 12, n. 6, jun., 2006. Disponível em: <http://www.dlib.org/dlib/june06/chan/06chan.html>. Acesso em: setembro de 2011.
- CHUDAMANI, K. S.; NAGARATHNA, H. C. *Metadata interoperability for cataloguing of e-journals: a comparative study of Libsys, MARC21, and Conser*. [s.d.] Disponível em: <http://shodhganga.inflibnet.ac.in/dxml/bitstream/handle/1944/1389/25.pdf?sequence=1>. Acesso em: agosto de 2011.
- COLE, T. W.; FOULONNEAU, M. *Using the Open Archives Initiative Protocol for Metadata Harvesting*. Westport, CT: Third Millennium Cataloging, 2007.
- COSTA, S. M. S. Mudanças no processo de comunicação científica: o impacto do uso de novas tecnologias. In: MUELLER, Suzana Pinheiro Machado; PASSOS, Edilenice. (Orgs). *Comunicação científica*. Brasília: Departamento de Ciência da Informação da Universidade de Brasília, 2000. p. 95-105. Disponível em: [http://repositorio.bce.unb.br/bitstream/10482/1443/1/CAPITULO\\_MudancaProcessoComunicacao.pdf](http://repositorio.bce.unb.br/bitstream/10482/1443/1/CAPITULO_MudancaProcessoComunicacao.pdf). Acesso em: novembro de 2011.
- COYLE, K. Guidelines for Dublin Core Application Profiles. *Dublin Core Metadata Initiative*, 2009. Disponível em: <http://dublincore.org/documents/2009/05/18/profile-guidelines>. Acesso em: julho de 2011.
- CROCKER, D. Making standards the IETF way. *StandardView*, v. 1, n. 1, 1993. Disponível em: <http://www.imc.org/making-standards>. Acesso em: outubro de 2011.
- DCMI. Dublin Core Metadata Element Set, Version 1.1. *Dublin Core Metadata Initiative*, 2010. Disponível em: <http://dublincore.org/documents/2010/10/11/dces/>. Acesso em: outubro de 2011.
- DCMI. Metadata basics. *Dublin Core Metadata Initiative*, c1995-2011. Disponível em: <http://dublincore.org/metadata-basics/>. Acesso em: outubro de 2011.
- DCMI Media Wiki. Use guide. *Dublin Core Metadata Initiative*, 2011. [This page was last modified on 6 September 2011, at 11:35] Disponível em: [http://wiki.dublincore.org/index.php/User\\_Guide](http://wiki.dublincore.org/index.php/User_Guide). Acesso em: 26 outubro de 2011
- DCMI USAGE BOARD. DCMI Metadata Terms. *Dublin Core Metadata Initiative*, 2010a. Disponível em: <http://dublincore.org/documents/2008/01/14/dcmi-terms/>. Acesso em: março de 2011.
- DCMI USAGE BOARD. DCMI Type Vocabulary. *Dublin Core Metadata Initiative*, 2010b. Disponível em: <http://dublincore.org/documents/dcmi-type-vocabulary/>. Acesso em: setembro de 2011.
- DUVAL, E. et al. Metadata principles and practicabilities. *D-Lib Magazine*, v. 8, n. 4, apr., 2002. Disponível em: <http://www.dlib.org/dlib/april02/weibel/04weibel.html>. Acesso em: novembro de 2011.
- DUVAL, E. Metadata standards: what, who & why. *Journal of Universal Computer Science*, v. 7, n. 7, jul., 2001, p. 591-601. Disponível em: [http://www.jucs.org/jucs\\_7\\_7/metadata\\_standards\\_what\\_who/Duval\\_E.pdf](http://www.jucs.org/jucs_7_7/metadata_standards_what_who/Duval_E.pdf). Acesso em: novembro de 2011.
- DYKES, L. *XML for dummies*, 4th ed. Hoboken, NJ : Wiley Publishing, Inc., c2005.
- FERREIRA, A. B. H. *Novo dicionário aurélio: século XXI*. [Dicionário Aurélio Eletrônico] MGB Informática Ltda, 1999.
- FERREIRA, S. M. S. P.; SOUTO, L. F. Dos sistemas de informação federados à federação de bibliotecas digitais. *Revista Brasileira de Biblioteconomia e Documentação*, Nova Série, São Paulo, v. 2, n. 1, jan./jun., 2006, p.23-40. Disponível em: <http://revista.ibict.br/pbcib/index.php/pbcib/article/view/223>. Acesso em: novembro de 2011.

GILLILAND, A. J. Setting the stage . In: *Introduction to Metadata – online edition, version 3.0*. J. Paul Getty Trust, c2008. Disponível em: [http://www.getty.edu/research/publications/electronic\\_publications/intrometadata/setting.html](http://www.getty.edu/research/publications/electronic_publications/intrometadata/setting.html). Acesso em: outubro de 2011.

GUIZZARDI, G. *Ontological foundations for structural conceptual models*. Twente, Netherlands, 2005. Thesis (doctorate) – Twente University of Technology. Disponível em [www.loa-cnr.it/Guizzardi/cap3.pdf](http://www.loa-cnr.it/Guizzardi/cap3.pdf). Acesso em agosto de 2011.

HARPER, C. A. Dublin Core Metadata Initiative: beyond the element set. *Information Standards Quarterly*, v. 22, n. 1, winter, 2010. Disponível em: [http://www.niso.org/publications/isq/free/FE\\_DCMI\\_Harper\\_isqv22no1.pdf](http://www.niso.org/publications/isq/free/FE_DCMI_Harper_isqv22no1.pdf). Acesso em: novembro de 2011.

HEERY, R.; PATEL, M. Application profiles: mixing and matching metadata schemas. *Ariadne*, n. 25, 24-Sep-2000 [última modificação: Monday, 04-Oct-2004 11:48:26 UTC]. Disponível em: <http://www.ariadne.ac.uk/issue25/app-profiles/>. Acesso em: novembro de 2011.

HILLMAN, D. Using Dublin Core. *Dublin Core Metadata Initiative*, 2005. Disponível em: <http://dublincore.org/documents/2005/11/07/usageguide/>. Acesso em: outubro de 2011

HJELMSLEV, L. (1961) *Prolegômenos a uma teoria da linguagem*. São Paulo: Perspectiva, 1975.

HOBERMAN, S. *Data Modeling Made Simple: A Practical Guide for Business and IT Professionals*, 2nd edition. Bradley Beach, NJ : Technics Publications, 2009.

HODGE, G. *Metadata made simpler: a guide for libraries*. Bethesda, MD: NISO Press, 2001. Disponível em: [http://www.ncsi.iisc.ernet.in/raja/is214/214-2001-2002/Metadata\\_Bklt.pdf](http://www.ncsi.iisc.ernet.in/raja/is214/214-2001-2002/Metadata_Bklt.pdf). Acesso em: novembro de 2011.

ISO. *Information technology – metadata registries (MDR) – part 3: registry metamodel and basic attributes*, 2nd edition. Geneva, Switzerland, 2003. Disponível em: [http://standards.iso.org/ittf/PubliclyAvailableStandards/c031367\\_ISO\\_IEC\\_11179-3\\_2003\\_\(E\).zip](http://standards.iso.org/ittf/PubliclyAvailableStandards/c031367_ISO_IEC_11179-3_2003_(E).zip). Acesso em: outubro de 2011.

MARCONDES, C. H.; SAYÃO, L. F. Documentos digitais e novas formas de cooperação entre sistemas de informação em c&t. *Ciência da Informação*, Brasília, v. 31, n. 3, set./dez., 2002. Disponível em: <http://www.scielo.br/pdf/ci/v31n3/a05v31n3.pdf>. Acesso em: novembro de 2011.

*Metadata Watch Report* [Periódico online] v. 5, aug., 2001. Disponível em: <http://www.schemas-forum.org/metadata-watch/fifth/section3.html>. Acesso em novembro de 2011.

MOODY, D.; WALSH, P. Measuring the value of information: an asset valuation approach. In: *European Conference on Information Systems (ECIS)*, 7th, 1999. Copenhagen Business School, Frederiksberg, Denmark, 23-25 June, 1999. Disponível em: <http://www.info.deis.unical.it/~zumpano/2004-2005/PSI/lezione2/ValueOfInformation.pdf>. Acesso em: novembro de 2011.

NAGAMORI, M.; SUGIMOTO, S. Using metadata schema registry as a core function to enhance usability and reusability of metadata schemas. In: *International Conference on Dublin Core and Metadata Applications*, 2007 (Proceedings). Singapore. Disponível em: <http://dcpapers.dublincore.org/ojs/pubs/article/view/869/865>. Acesso em: novembro de 2011.

NILSSON, M. Interoperability Levels for Dublin Core Metadata. *Dublin Core Metadata Initiative*, 2009. Disponível em: <http://dublincore.org/documents/2009/05/01/interoperability-levels>. Acesso em outubro de 2011.

NILSSON, M. *From interoperability to harmonization in metadata standardization: designing an evolvable framework for metadata harmonization*. Stockholm, Sweden, 2010. Thesis (Doctoral) – KTH Royal Institute of Technology. Disponível em: <http://www.dissertations.se/dissertation/fd9948f471/>. Acesso em: novembro de 2011.

NILSSON, M.; BAKER, T.; JOHNSTON, P. The Singapura Framework for Dublin Core Application Profiles. *Dublin Core Metadata Initiative*, 2008. Disponível em <http://dublincore.org/documents/2008/01/14/singapore-framework/>. Acesso em: setembro de 2011.

NISO. *A framework of guidance for building good digital collections*, 3rd ed. Baltimore, MD : National Information Standards Organization, 2007. Disponível em: <http://framework.niso.org/node/24>. Acesso em novembro de 2011.

NISO. *Understanding Metadata*. Bethesda, MD: NISO Press, c2004. Disponível em: <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>. Acesso em: novembro de 2011.

OLIVEIRA, V. S. *Buscando interoperabilidade entre diferentes bases de dados: o caso da Biblioteca do Instituto Fernandes Figueira*. Dissertação (Mestrado) – Escola Nacional de Saúde Pública Sérgio Arouca - FIOCRUZ. Disponível em: <http://www.arca.fiocruz.br/bitstream/icict/1318/1/oliveiravsm.pdf>. Acesso em: novembro de 2011.

OUKSEL, A.M.; SHETH, A. *Semantic interoperability in global Information Systems : a brief introduction to the research area and the special section*. ACM SIGMOD Record, v. 28, n. 1, mar., 1999, p. 5-12. Disponível em: <http://knoesis.wright.edu/library/download/OS99-SemanticInteropIS.pdf>. Acesso em: novembro de 2011.

PAEPCKE, A. et al. Interoperability for digital libraries worldwide: problems and directions. *Stanford Digital Library Technologies Project*, 1998. Disponível em: <http://ilpubs.stanford.edu:8090/317/1/1998-24.pdf>. Acesso em: novembro de 2011.

PARK, J.; CHILDRESS, E. Dublin Core metadata semantics: an analysis of the perspectives of information professionals. [Published online before print in 29 May 2009]. *Journal of Information Science*, v. 35, n. 6, dec., 2009, p. 727-739. Disponível em: <http://jis.sagepub.com/content/35/6/727>. Acesso em: agosto de 2011

PATEL, M. et al. Semantic interoperability in digital library systems. *DELOS Network of Excellence on Digital Libraries Project*, 2005. [file last modified: Tuesday, 20-sep-2005 15:18:32 UTC] Disponível em: <http://delos-wp5.ukoln.ac.uk/project-outcomes/SI-in-DLs/>. Acessado em: maio 2010.

POMBO, O. *Da classificação dos seres à classificação dos saberes*. [Comunicação apresentada no Colóquio sobre Classificação, organizado pela revista Leituras, da Biblioteca Nacional de Lisboa, em 21 de outubro de 1997]. Disponível em <http://www.educ.fc.ul.pt/hyper/resources/opombo-classificacao.pdf>. Acessado em: março de 2011.

POWELL, A. et al. DCMI Abstract Model. *Dublin Core Metadata Initiative*, 2007a. Disponível em: <http://dublincore.org/documents/2007/06/04/abstract-model/>. Acesso em: outubro de 2011.

POWELL, A. et al. Namespace policy for the Dublin Core Metadata Initiative (DCMI). *Dublin Core Metadata Initiative*, 2007b. Disponível em: <http://dublincore.org/documents/2007/07/02/dcmi-namespace>. Acesso em: outubro de 2011.

RIBEIRO, A. M. C. M. *Catálogo de recursos bibliográficos: AACR2R em MARC21*. – 3. ed. Brasília : Ed. do Autor, 2006.

ROCHA, R. P. Metadados, Web Semântica e Categorização Automática. *Em Questão*. Porto Alegre, v. 10, n. 1, jan./jun., 2004, p. 109-121. Disponível em: [http://www6.ufrgs.br/emquestao/pdf\\_2004\\_v10\\_n1/EmQuestaoV10\\_N1\\_2004\\_art07.pdf](http://www6.ufrgs.br/emquestao/pdf_2004_v10_n1/EmQuestaoV10_N1_2004_art07.pdf). Acesso em: novembro de 2011.

ROSETTO, M.; NOGUEIRA, A. H. Aplicação de elementos metadados Dublin Core para descrição de dados bibliográficos on-line da Biblioteca Digital de Teses da USP. In: *Seminário Nacional de Bibliotecas Universitárias*, XII, 2002. Universidade Federal de Pernambuco. Disponível em: <http://www.sibi.ufrj.br/snbu2002/oralpdf/82.a.pdf>. Acesso em: outubro de 2011.

ROWLEY, J. *A biblioteca eletrônica* - 2 ed. de Informática para bibliotecas. Tradução de Antonio Agenor Brinquet de Lemos. Brasília: Brinquet de Lemos/Livros, 2002.

SCOTT, R W. *Comics librarianship: a handbook*. Jefferson, North Carolina : McFarland, 1990.

SENSO, J. A.; ROSA PIÑERO, A. El concepto de metadato: algo más que descripción de recursos electrónicos. *Ciência da Informação*, Brasília, v. 32, n. 2, maio/ago., 2003, p. 95-106. Disponível em: <http://revista.ibict.br/index.php/ciinf/article/view/126/107>. Acesso em: novembro de 2011.

SHARP, K. Internet librarianship: traditional roles in a new environment. In: *Libraries in the Information Society / International Federation of Library Associations and institutions*. Edited by Tatiana V. Ershova and Yuri E. Hohlov. München : Saur, 2002.

SHETH, A. P.; LARSON, J. A. Federated Database Systems for Managing Distributed, heterogeneous, and Autonomous Databases. *ACM Computing Surveys*, v. 22, n. 3, sep., 1990. Disponível em: <http://knoesis.org/library/download/SL90.pdf>. Acesso em: novembro de 2011.

SOMPPEL, H.; LAGOZE, C. The Santa Fe Convention of the Open Archives Initiative. *D-Lib Magazine*, v. 6, n. 2, feb., 2000. Disponível em: <http://www.dlib.org/dlib/february00/vandesompel-oai/02vandesompel-oai.html>. Acesso em novembro de 2011.

SOUZA, M. I. F.; MOURA, M. F.; SANTOS, A. D. (2005) apud ALVES, M. D. R.; SOUZA, M. I. F. Estudo de correspondência de elementos metadados: Dublin Core e Marc 21. *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, v. 4, n. 2, p. 20-38, jan./jun., 2007. Disponível em: [http://www.sbu.unicamp.br/seer/ojs/index.php/sbu\\_rci/article/viewFile/358/237](http://www.sbu.unicamp.br/seer/ojs/index.php/sbu_rci/article/viewFile/358/237). Acesso em: novembro de 2011.

VERGUEIRO, W. Autorias e responsabilidades nos quadrinhos: uma discussão provocativa. *Omelete* [website], 28 de maio de 2003. Disponível em: <http://omelete.uol.com.br/quadrinhos/autorias-e-responsabilidades-nos-quadrinhos-uma-discussao-provocativa/>. Acesso em outubro de 2011.

VERGUEIRO, W. Gibitecas brasileiras: um espaço para sonhos. *Omelete* [website], 5 de janeiro de 2004. Disponível em: <http://omelete.uol.com.br/quadrinhos/gibitecas-brasileiras-um-espaco-para-sonhos/>. Acesso em: novembro de 2011.

VERGUEIRO, W. Histórias em quadrinhos e serviços de informação: um relacionamento em fase de definição. *DataGramaZero* [Periódico online] v. 6, n. 2, abr., 2005. Disponível em: [http://www.dgz.org.br/abr05/Art\\_04.htm](http://www.dgz.org.br/abr05/Art_04.htm). Acesso em: novembro de 2011.

W3C. (1999) *Resource Description Framework (RDF) Model and Syntax Specification*. W3C Recommendation 22 February 1999 [Last updated: Date: 1999/02/24 14:45:07] Disponível em: <http://w3.org/TR/1999/REC-rdf-syntax-19990222>. Acesso em: outubro de 2011.

W3C. (2006) *Namespaces in XML 1.1 (Second Edition)*. W3C Recommendation 16 August 2006. Disponível em: <http://www.w3.org/TR/2006/REC-xml-names11-20060816>. Acesso em: outubro de 2011.

WEST, M. *Developing high quality data models* – version 2.0 Issue 2.1. European Process Industries STEP Technical Liaison Executive / Julian Fowler (editor), 2003. Disponível em: <http://www.matthew-west.org.uk/documents/princ03.pdf>. Acesso em: novembro de 2011.

ZENG, M. L.; CHAN, L. M. Metadata interoperability and standardization – a study of methodology, part II: achieving interoperability at the record and repository levels. *D-Lib Magazine*, v. 12, n. 6, Jun., 2006. Disponível em: <http://www.dlib.org/dlib/june06/chan/06chan.html>. Acesso em setembro de 2011.

## **Anexos**

## Semântica do Modelo Abstrato DCMI\*

O Modelo Abstrato DCMI faz uso dos seguintes termos:

**Classe:** Um grupo contendo membros que possuem atributos, comportamentos, relações ou semântica em comum; um tipo de categoria.

**Conjunto descritivo:** Um conjunto de uma ou mais *descrições*.

**Declaração:** Uma instanciação de um par *propriedade-valor* feito de um *URI de propriedade* (um *URI* que identifica uma *propriedade* e um *valor suplente*).

**Descrição:** Uma ou mais declarações sobre em, e somente um, *recurso*.

**Elemento:** Sinônimo de *propriedade*. É importante notar que a palavra “elemento” também é comumente usada para se referir ao componente de uma marcação estrutural dentro de um documento XML.

**Esquema de codificação sintática:** Um conjunto de “strings” e um conjunto de regras associado que descreve um mapeamento entre esse conjunto de “strings” e um conjunto de *recursos*. As regras de mapeamento podem definir como a “string” é estruturada (por exemplo, DCMI Box) ou eles podem simplesmente enumerar todas as “strings” e os *recursos* correspondentes (por exemplo, ISO 3166).

**Instância de:** Uma relação entre um *recurso* e uma *classe* indicando que o *recurso* é uma instância.

**Linguagem valor “string”:** Uma tag ISO de linguagem que indica a linguagem do *valor* “string”.

**Literal:** Uma entidade que usa uma cadeia de caracteres Unicode como uma forma lexical, juntamente com uma linguagem-tag opcional ou “datatype”, para denotar um *recurso* (i.e. “literal” como definido pelo RDF).

**Membro de:** Um relacionamento entre um *recurso* e um *esquema de codificação vocabular* que indica que o *recurso* é um membro de um conjunto.

**Par propriedade-valor:** A combinação de uma *propriedade* e um *valor*, usado para descrever uma característica de um *recurso*.

**Propriedade:** Um aspecto específico, característica, atributo ou relação usada para descrever *recursos*.

**Recurso descrito:** Um recurso descrito por uma *descrição*.

**Recurso:** Tudo o que pode ser identificado. Exemplos familiares incluem um documento eletrônico, uma imagem, um serviço (por exemplo, “previsão do tempo”), e uma coleção de outros *recursos*. Nem todos os *recursos* são recuperáveis em rede; por exemplo, seres humanos, corporações, conceitos e livros físicos de uma biblioteca podem ser considerados *recursos*.

**Registro:** O instanciamento de um *conjunto descritivo*, criado em concordância com uma das diretrizes de codificação DCMI (por exemplo, XHTML meta tags, XML e RDF/XML).

**Sub-classe de:** Um relacionamento entre duas *classes* que indica que as duas *classes* são definidas de tal forma que todos os *recursos* que são instâncias da sub-*classe* são também instâncias da *classe* relacionada.

**Sub-propriedade de:** Um relacionamento entre duas *propriedades* que indica que as duas *propriedades* são definidas de tal forma que, quando a sub-*propriedade* é parte de um par

---

\* POWELL, A. et al. DCMI abstract model. *Dublin Core Metadata Initiative*, 2007-06-04. Disponível em: <http://dublincore.org/documents/2007/06/04/abstract-model/>. Acesso em novembro de 2011.

*propriedade-valor* descrevendo um *recurso*, então o *recurso* é também descrito usando um segundo par *propriedade-valor* feito da *propriedade* e do *valor*.

**Tem alcance:** Uma relação entre uma *propriedade* e uma *classe* indicando que se a *propriedade* é parte de um par *propriedade-valor*, então o *valor* é uma instância dessa *classe*.

**Tem domínio:** Uma relação entre uma *propriedade* e uma *classe* indicando que se a *propriedade* é parte de um par *propriedade-valor*, então o *recurso descrito* é uma instância dessa *classe*.

**Term:** Uma *propriedade (elemento)*, *classe*, *esquema de codificação vocabular* ou *esquema de codificação sintática*.

**URI de esquema de codificação sintática:** Um *URI* que identifica um *esquema de codificação sintática*.

**URI de propriedade:** Um *URI* que identifica uma única *propriedade*.

**URI de recurso descrito:** Um *URI* que identifica um *recurso descrito*.

**URI de valor:** Um *URI* que identifica um *valor*.

**URI:** Um “*Uniform Resource Identifier*” ou um IRI “*Internationalized Resource Identifier*”. Pela perspectiva do Modelo Abstrato DCMI, a equivalência entre URIs é definida como no RDF.

**Valor “string” simples:** um *valor “string”* sem um *URI de esquema de codificação sintática* relacionado.

**Valor “string” tipificado:** Um *valor “string”* com um *URI de esquema de codificação sintática* associado.

**Valor “string”:** Um *literal*, opcionalmente associado com um *URI de esquema de codificação sintática* ou uma *linguagem de valor “string”*. Em um *valor literal suplente* um *valor “string”* codifica o *valor*; em um *valor suplente não literal* um *valor “string”* representa o *valor*.

**Valor literal suplente:** Um *valor suplente* para um *valor literal*, feito de exatamente um *valor “string”* (um *literal* que codifica um *valor*).

**Valor literal:** Um *valor* que é um *literal*.

**Valor não literal:** Um *valor* que é uma entidade física, digital ou conceitual.

**Valor suplente não literal:** Um *valor suplente* para um *valor não literal*, feito de uma *URI de propriedade* (uma *URI* que identifica uma *propriedade*, zero ou um *URI de valor* (um *URI* que identifica o *valor não literal* associado com a *propriedade*), zero ou um *URI de esquema de codificação vocabular* (uma *URI* que identifica um *esquema de codificação vocabular* do qual o *valor* é um membro), e zero ou mais *valores “string”* (*literais* que representam um *valor*).

**Valor:** A entidade física, entidade conceitual ou *literal* (um *recurso*) que é associada a uma *propriedade* quando um par *propriedade-valor* é usado para descrever um *recurso*.

**Valor suplente:** Um *valor literal suplente* ou um *valor suplente não literal*.

**Vocabulário:** Um conjunto de um ou mais *termos*.

**Esquema de codificação vocabular:** Um conjunto enumerado de *recursos*.

**URI de esquema de codificação vocabular:** Um *URI* que identifica um *esquema de codificação vocabular*.

## Moldes de Descrição e Declaração do PAPuQ

```
DescriptionSet: PAPuQ
  Description template: exemplar
  minimum = 1; maximum = 1
  Statement template: titulo
  minimum = 1; maximum = unlimited
  Property: http://purl.org/dc/elements/1.1/title
  Type of Value = "literal"
  Statement template: data_de_publicacao
  minimum = 1; maximum = 1
  Property: http://purl.org/dc/elements/1.1/date
  Type of Value = "literal"
  Syntax Encoding Scheme URI =
http://purl.org/dc/terms/W3CDTF
  Statement template: idioma
  minimum = 1; maximum = unlimited
  Property: http://purl.org/dc/elements/1.1/language
  Type of Value = "non-literal"
  takes list = yes
  Syntax Encoding Scheme URI =
http://purl.org/dc/terms/ISO639-2
  Statement template: personagem
  minimum = 0; maximum = unlimited
  Property: http://purl.org/dc/elements/1.1/subject
  Type of Value = "literal"
  Statement template: criador
  minimum = 0; maximum = unlimited
  Property: http://purl.org/dc/elements/1.1/creator
  Type of Value = "literal"
  Statement template: contribuidor
  minimum = 0; maximum = unlimited
  Property:
http://purl.org/dc/elements/1.1/contributor
  Type of Value = "literal"
  Statement template: editora
  minimum = 1; maximum = unlimited
  Property:
http://purl.org/dc/elements/1.1/publisher
  Type of Value = "literal"
  Statement template: direitos_autorais
  minimum = 0; maximum = unlimited
  Property: http://purl.org/dc/elements/1.1/rights
  Type of Value = "literal"
  Statement template: formato
  minimum = 1; maximum = unlimited
```



Property: <http://purl.org/dc/elements/1.1/format>  
Type of Value = "literal"  
Statement template: notas  
minimum = 0; maximum = unlimited  
Property:  
<http://purl.org/dc/elements/1.1/description>  
Type of Value = "literal"  
Statement template: codigo\_internacional  
minimum = 0; maximum = unlimited  
Property: <http://purl.org/dc/elements/1.1/relation>  
Type of Value = "literal"  
Statement template: local\_de\_origem  
minimum = 0; maximum = 1  
Property: <http://purl.org/dc/elements/1.1/coverage>  
Type of Value = "non-literal"  
takes list = yes  
Syntax Encoding Scheme URI =  
<http://purl.org/dc/terms/ISO3166-3>  
Statement template: tipo  
minimum = 2; maximum = 2  
Property: <http://purl.org/dc/elements/1.1/type>  
Type of Value = "non-literal"  
takes list = yes  
Syntax Encoding Scheme URI =  
<http://purl.org/dc/dcmitype/>  
Statement template: coleccionador  
minimum = 1; maximum = unlimited  
Property: <http://purl.org/dc/elements/1.1/source>  
Type of Value = "literal"

## XML Schema PAPuQ

### Proposta de XML Schema para publicações de quadrinhos

```
<?xml version="1.0" encoding="utf-8" ?>
<!--Created with Liquid XML Studio Developer Edition (Trial)
9.0.11.3078 (http://www.liquid-technologies.com)-->
<schema xmlns="http://www.w3.org/2001/XMLSchema"
xmlns:dc="http://purl.org/dc/elements/1.1/">
  <group name="exemplar">
    <sequence>
      <element name="title" minOccurs="1" maxOccurs="unbounded">
        <complexType>
          <simpleContent>
            <extension base="string">
              <attribute name="lingua" type="language"
/>
            </extension>
          </simpleContent>
        </complexType>
      </element>
    </sequence>
  </group>
  <element name="date" type="string" />
  <group name="idioma">
    <sequence>
      <element name="language" maxOccurs="unbounded">
        <simpleType>
          <restriction base="string">
            <enumeration value="ara" />
            <enumeration value="art" />
            <enumeration value="chi" />
            <enumeration value="eng" />
            <enumeration value="epo" />
            <enumeration value="fre" />
            <enumeration value="ger" />
            <enumeration value="grn" />
            <enumeration value="heb" />
            <enumeration value="ile" />
            <enumeration value="ina" />
            <enumeration value="ita" />
            <enumeration value="jpn" />
            <enumeration value="kor" />
            <enumeration value="lat" />
            <enumeration value="pol" />
            <enumeration value="por" />
            <enumeration value="san" />
            <enumeration value="spa" />
            <enumeration value="tlh" />
            <enumeration value="tup" />
            <enumeration value="tur" />
            <enumeration value="vol" />
            <enumeration value="yor" />
            <enumeration value="zul" />
            <enumeration value="zxx" />
          </restriction>
        </simpleType>
      </element>
    </sequence>
  </group>
</schema>
```

```

</group>
<element name="coverage">
  <simpleType>
    <restriction base="string">
      <enumeration value="ago" />
      <enumeration value="arg" />
      <enumeration value="aus" />
      <enumeration value="aut" />
      <enumeration value="bel" />
      <enumeration value="bol" />
      <enumeration value="bra" />
      <enumeration value="can" />
      <enumeration value="col" />
      <enumeration value="crc" />
      <enumeration value="hrv" />
      <enumeration value="cub" />
      <enumeration value="dnk" />
      <enumeration value="egy" />
      <enumeration value="fra" />
      <enumeration value="deu" />
      <enumeration value="gtm" />
      <enumeration value="hti" />
      <enumeration value="hkg" />
      <enumeration value="idn" />
      <enumeration value="isr" />
      <enumeration value="ita" />
      <enumeration value="jam" />
      <enumeration value="jpn" />
      <enumeration value="kwt" />
      <enumeration value="lbn" />
      <enumeration value="mex" />
      <enumeration value="moz" />
      <enumeration value="nld" />
      <enumeration value="pak" />
      <enumeration value="pry" />
      <enumeration value="per" />
      <enumeration value="pol" />
      <enumeration value="prt" />
      <enumeration value="rus" />
      <enumeration value="sen" />
      <enumeration value="sgp" />
      <enumeration value="esp" />
      <enumeration value="swe" />
      <enumeration value="tls" />
      <enumeration value="tur" />
      <enumeration value="gbr" />
      <enumeration value="usa" />
      <enumeration value="ury" />
      <enumeration value="ven" />
      <enumeration value="esh" />
      <enumeration value="yem" />
      <enumeration value="zmb" />
    </restriction>
  </simpleType>
</element>
<group name="editora">
  <sequence>
    <element name="publisher" type="string" minOccurs="0"
maxOccurs="unbounded" />
  </sequence>
</group>

```

```

    <group name="formato">
      <sequence>
        <element name="format" minOccurs="0"
maxOccurs="unbounded">
          <complexType>
            <simpleContent>
              <extension base="string">
                <attribute name="lingua" type="language"
/>
              </extension>
            </simpleContent>
          </complexType>
        </element>
      </sequence>
    </group>
    <group name="criador">
      <sequence>
        <element name="creator" minOccurs="0"
maxOccurs="unbounded">
          <simpleType>
            <restriction base="string">
              <enumeration value="aut" />
              <enumeration value="ilu" />
            </restriction>
          </simpleType>
        </element>
      </sequence>
    </group>
    <group name="contribuidor">
      <sequence>
        <element name="contributor" minOccurs="0"
maxOccurs="unbounded">
          <simpleType>
            <restriction base="string">
              <enumeration value="ill" />
              <enumeration value="edt" />
              <enumeration value="com" />
              <enumeration value="trl" />
              <enumeration value="clr" />
              <enumeration value="c11" />
              <enumeration value="drm" />
              <enumeration value="ctb" />
            </restriction>
          </simpleType>
        </element>
      </sequence>
    </group>
    <group name="personagem">
      <sequence>
        <element name="subject" minOccurs="0"
maxOccurs="unbounded">
          <complexType>
            <simpleContent>
              <extension base="string">
                <attribute name="lingua" type="language"
/>
              </extension>
            </simpleContent>
          </complexType>
        </element>
      </sequence>

```

```

    </group>
    <group name="direitos_autorais">
      <sequence>
        <element name="rights" minOccurs="0"
maxOccurs="unbounded">
          <complexType>
            <simpleContent>
              <extension base="string">
                <attribute name="lingua" type="language"
/>
              </extension>
            </simpleContent>
          </complexType>
        </element>
      </sequence>
    </group>
    <group name="codigo_publicacao">
      <sequence>
        <element name="relation" type="string" minOccurs="0"
maxOccurs="unbounded" />
      </sequence>
    </group>
    <group name="notas">
      <sequence>
        <element name="description" minOccurs="0"
maxOccurs="unbounded">
          <complexType>
            <simpleContent>
              <extension base="string">
                <attribute name="lingua" type="language"
/>
              </extension>
            </simpleContent>
          </complexType>
        </element>
      </sequence>
    </group>
    <group name="tipo">
      <sequence>
        <element name="type" type="string" minOccurs="2"
maxOccurs="2" />
      </sequence>
    </group>
    <group name="colecacao">
      <sequence>
        <element name="source" type="string" minOccurs="1"
maxOccurs="unbounded" />
      </sequence>
    </group>
  </schema>

```