

Universidade Federal de Ouro Preto - UFOP
Instituto de Ciências Exatas e Biológicas - ICEB
Departamento de Computação - DECOM

Proposta de uma Biblioteca Digital para Trabalhos
Monográficos do Departamento de Computação da
Universidade Federal de Ouro Preto

Aluno: Gustavo Henrique Braz da Silva
Matricula: 07.1.4160

Orientador: Guilherme Tavares de Assis

Ouro Preto
9 de novembro de 2012

Universidade Federal de Ouro Preto - UFOP
Instituto de Ciências Exatas e Biológicas - ICEB
Departamento de Computação - DECOM

Proposta de uma Biblioteca Digital para Trabalhos
Monográficos do Departamento de Computação da
Universidade Federal de Ouro Preto

Relatório de atividades desenvolvidas apresentado ao curso de Bacharelado em Ciência da Computação, Universidade Federal de Ouro Preto, como requisito parcial para a conclusão da disciplina Monografia I(BCC390).

Aluno: Gustavo Henrique Braz da Silva
Matricula: 07.1.4160

Orientador: Guilherme Tavares de Assis

Ouro Preto
9 de novembro de 2012

Resumo

Esta monografia apresenta uma proposta para a confecção de uma biblioteca digital voltada para a organização e o armazenamento, a princípio, de trabalhos monográficos do Departamento de Computação (DECOM) da Universidade Federal de Ouro Preto (UFOP), facilitando a preservação e a pesquisa dos mesmos pela comunidade acadêmica. Para a organização dos metadados, será utilizado o padrão *Dublin Core* [16] e para a interoperabilidade entre bibliotecas digitais e colheita de metadados será utilizado o protocolo OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*) [21]. Visando comprovar a eficácia e a eficiência da biblioteca digital proposta, a mesma será confeccionada e experimentos práticos serão realizados por usuários reais.

Palavras-chave: biblioteca digital, biblioteca virtual, *Dublin Core*, metadados.

Sumário

1	Introdução	1
1.1	Trabalhos Relacionados	3
2	Fundamentação Teórica	5
2.1	Visão geral sobre bibliotecas digitais	5
2.1.1	Metadados	6
2.1.2	Dublin Core	8
2.1.3	<i>Extensible Markup Language</i> XML	11
2.1.4	Protocolo OAI-PMH	12
2.2	BDBComp	14
2.3	BDCUNI	14
3	Desenvolvimento da Biblioteca Digital	15
3.1	Modelagem de Dados	15
4	Anexo - Script do Banco de Dados	18

Lista de Figuras

1	Exemplo de estrutura XML	12
2	Metadata Harvesting [25]	13
3	Modelagem de dados usando o modelo Entidade-Relacionamento	17
4	XML biblioteca digital DECOM/UFOP	17

Lista de Tabelas

1	Elementos Dublin Core	10
2	Elementos Qualified Dublin Core	11
3	Verbos do protocolo OAI-PMH	14
4	Mapeamento do <i>Dublin Core</i> para a biblioteca digital do DECOM . . .	16

1 Introdução

No sentido tradicional, uma biblioteca é um lugar no qual livros são conservados para as gerações futuras, de modo a oferecer continuidade à produção cultural e para formar a memória coletiva, encontrando-se organizados de forma a facilitar a localização dos mesmos pelas pessoas interessadas. De acordo com [15], a mais célebre e grandiosa biblioteca da Antiguidade, a de Alexandria, foi criada no século IV A.C. no Egito, tendo como ambição reunir em um só lugar todo o conhecimento da humanidade. Seu acervo era constituído de rolos de papiro manuscritos (aproximadamente 60 mil), contendo literatura grega, egípcia, assíria e babilônica.

Semelhante ao ideal da biblioteca de Alexandria, que aspirava reunir todo o conhecimento da humanidade em um só lugar, atualmente a rede mundial de computadores vem reunindo produções científicas como artigos, teses, monografias, e ideias e opiniões diversas. Com o desenvolvimento das redes computacionais, a leitura de documentos se faz cada vez mais por intermédio da tecnologia digital. A expansão da internet e o consequente desenvolvimento da tecnologia de redes eletrônicas vêm intensificando o fenômeno da explosão de documentos eletrônicos, ocasionando o aumento do volume de informações disponíveis. Assim, pode-se ter acesso a uma grande variedade de recursos informacionais como, por exemplo, *websites*, artigos eletrônicos, informações comerciais, culturais, artísticas, entre outros. Em vista dessa realidade, torna-se imprescindível o desenvolvimento de padrões que visem melhorar a organização dos recursos de informação.

Diante desse novo cenário, uma biblioteca digital apresenta-se como alternativa para a divulgação de produções científicas de forma consistente e organizada. Além de implicar um novo conceito para o armazenamento eletrônico de dados, ela melhora a forma de disseminação dos mesmos, que independe de localização física ou de horário. Assim, nesse contexto conceitual, estão embutidos a criação, a aquisição, a distribuição e o armazenamento de documentos sob a forma digital.

O aspecto fundamental para as bibliotecas digitais é a interoperabilidade: potencialidade de redes tecnológicas em trocar e compartilhar documentos, pesquisas e serviços. Para que essa interoperabilidade aconteça, é necessário reunir informações em padrões nacionais e internacionais, para que auxiliem no relacionamento entre dados. Tais informações aparecem na forma de metadados: características adicionais sobre os dados tornando-os mais significativos; tais características sobre os dados não estão normalmente incluídas nos dados propriamente ditos. Os metadados são utilizados para facilitar o entendimento, o uso e o gerenciamento de dados. Os metadados necessários para este fim variam conforme o tipo de dados e o contexto de uso. Assim, no contexto de uma biblioteca digital, onde os dados são o conteúdo das publicações, os metadados a respeito de uma publicação normalmente incluem uma descrição do conteúdo, o autor, a data de publicação e o tipo.

Bibliotecas digitais, segundo a *Digital Library Federation*¹, são sistemas eletrônicos de armazenamento de informações com capacidade para selecionar, estruturar, oferecer acesso intelectual, interpretar, distribuir, preservar a integridade e garantir a persistência, ao longo do tempo, das coleções de trabalhos digitais, para que estejam legíveis e disponíveis para uso por uma comunidade ou um conjunto de comunidades defini-

¹Consórcio de bibliotecas e agências relacionadas que usam tecnologias de informação eletrônica para estender suas coleções e serviços.

das [10]. Uma biblioteca digital estende e amplia os recursos e serviços disponíveis nas bibliotecas convencionais, possibilitando o intercâmbio de informações a nível mundial.

As bibliotecas digitais não se limitam, apenas, em armazenar arquivos texto. Segundo [13], “bibliotecas digitais envolvem a integração de sistemas complexos, incluindo coleção de documentos com estruturas, mídias e conteúdos variados, além de uma mistura de componentes de *hardware* e *software* interoperando, ao longo de diferentes estruturas de dados, algoritmos de processamento e múltiplas pessoas, comunidades e instituições com diferentes objetivos, políticas e culturas”.

O curso de Ciência da Computação da Universidade Federal de Ouro Preto (UFOP), criado no ano de 1992, possui em sua matriz curricular a matéria Monografia, que é obrigatória para a conclusão do curso. A partir dessa matéria, é produzida uma certa quantidade de trabalhos monográficos, que geram um grande volume de publicações oriundos do Departamento de Ciência da Computação (DECOM). Essas produções ficam armazenadas apenas de maneira tradicional, ou seja, na forma de papel impresso. Essa forma de armazenamento gera um volume de material de difícil acesso e má preservação. Com a utilização de uma biblioteca digital para o armazenamento de tais documentos técnico-científicos do departamento, é possível proporcionar o acesso remoto e ininterrupto aos mesmos, trazendo agilidade na obtenção das informações desejadas e auxiliando na preservação dos documentos originais. Com isso, a visibilidade desses documentos pode ser proporcionada, já que comunidades científicas nacionais e internacionais poderão acessá-los a qualquer momento de forma simultânea. Essa necessidade de armazenar, organizar e divulgar tais documentos técnico-científicos motivou a proposta de criação de uma biblioteca digital para o DECOM.

Portanto, o objetivo principal desta monografia consiste em se confeccionar uma biblioteca digital para o DECOM/UFOP com a finalidade de armazenar e disponibilizar, em princípio, trabalhos monográficos e relatórios técnicos para a comunidade acadêmica. Tal biblioteca digital seguiu o *Dublin Core* [16], como padrão de metadados, e o protocolo OAI [21] para a interoperabilidade desses dados, apresentados na seção 2.1. A biblioteca digital foi construída de tal forma que, futuramente, possa ser estendida, possibilitando o armazenamento de outros tipos de documentos.

Para comprovar a eficácia e eficiência da biblioteca digital proposta, a metodologia consistiu na realização de experimentos práticos envolvendo diferentes usuários: professores e alunos do DECOM. Foram feitos testes de usabilidade [22], que mediram, com base nos usuários, a facilidade de uso da biblioteca digital. Ademais, para verificar a eficiência e o desempenho das buscas na biblioteca, foram realizados devidos experimentos que proporcionaram analisar a satisfação do usuário, quanto a tais quesitos. Para tanto, inicialmente, foi necessário fazer o armazenamento de trabalhos monográficos, após a confecção da biblioteca digital do departamento.

Antes de atingir a fase de confecção da biblioteca digital, outras etapas foram realizadas:

- (a) coleta de dados e estudo de outras bibliotecas digitais existentes, sendo uma ótima referência a Biblioteca Digital Brasileira de Computação (BDBComp) [1], apresentada formalmente na seção 1.1;
- (b) exploração do tema para agregar conhecimento e, assim, ajudar a delimitar o escopo do problema à realidade do DECOM/UFOP;
- (c) análise de documentos técnico-científicos do DECOM/UFOP a serem armazenados

na biblioteca;

- (d) projeto da estrutura da biblioteca digital propriamente dita;
- (e) implementação do banco de dados para dar suporte à biblioteca;

Com o proposto e desenvolvido trabalho, foi possível verificar que a contribuição da biblioteca digital desenvolvida para o DECOM é grande:

- (a) o texto digital dos documentos técnico-científicos pode ser facilmente atualizado, copiado e distribuído, reduzindo os custos de aquisição, armazenamento e preservação;
- (b) o acesso aos documentos é facilitado, pois a pesquisa pode ser feita por diferentes metadados que, neste caso, são os especificados pelo padrão *Dublin Core* como, por exemplo, título, autor e contribuidores;
- (c) o acesso aos documentos técnico-científicos é ininterrupto, trazendo agilidade na obtenção das informações desejadas e auxiliando na preservação dos documentos originais

O restante deste trabalho monográfico encontra-se organizado da seguinte maneira. A Seção 1 deste capítulo apresenta, brevemente, alguns trabalhos diretamente relacionados ao objetivo do trabalho monográfico. O capítulo 2 apresenta a fundamentação teórica desse trabalho: visão geral sobre bibliotecas digitais, metadados, padrão *Dublin Core*, protocolo OAI e documentos XML; e bibliotecas digitais existentes na literatura que contribuíram fortemente com o presente trabalho. O capítulo 3 descreve a nossa proposta de confecção da biblioteca digital para documentos técnico-científicos do DECOM/UFOP. O capítulo 4 discute os experimentos realizados, assim como os resultados obtidos, para verificar a eficiência e eficácia da biblioteca. O capítulo 5 apresenta a conclusão desse trabalho e as perspectivas de trabalho futuro.

1.1 Trabalhos Relacionados

A Biblioteca Digital Brasileira de Computação (BDBComp), melhor descrita na seção 2.2, disponibiliza e permite acesso aos artigos publicados nos eventos promovidos pela Sociedade Brasileira de Computação (SBC) [29]. A BDBComp foi concebida com o objetivo de arquivar, organizar, preservar dados digitais e também disseminar o conhecimento científico produzido pela comunidade brasileira de computação, auxiliando a recuperação da informação desejada pelos usuários. Atualmente, estão registrados 10217 trabalhos publicados em periódicos nacionais e anais de eventos realizados no Brasil. Além disso, conta com um serviço de auto-arquivamento para que pesquisadores, estudantes e profissionais possam submeter trabalhos científicos em seu repositório.

A Biblioteca Digital de Computação do Uni-BH (BDCUni), melhor descrita na seção 2.3, foi criada para o curso de ciência da computação do Centro Universitário de Belo Horizonte (UNI-BH), com o objetivo de efetuar o arquivamento dos trabalhos monográficos desenvolvidos na matéria “rojeto Orientado” do curso, promovendo assim, a difusão da produção científica para a comunidade de computação. A criação e consequente análise da BDCUni [10] permitiram verificar que uma biblioteca digital

é realmente uma forma eficiente de armazenamento e recuperação de documentos digitais como trabalhos monográficos. Analisando os experimentos, foi constatado que, por meio da BDCUni, conseguiu-se documentar e registrar trabalhos monográficos corretamente, facilitando a pesquisa e a recuperação das informações pesquisadas, e atendendo, de maneira satisfatória, as pretensões desejadas.

A Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) [20], promovida pela Universidade de São Paulo em 2011, tem como objetivo integrar os sistemas de informação de teses e dissertações existentes nas instituições de ensino e pesquisa brasileiras estimulando o registro e a publicação de teses e dissertações em meio eletrônico. A BDTD envolveu equipe multidisciplinar e as principais fases do processo inicial de implementação foram: prospecção tecnológica, engenharia de *software*, proposta de projeto piloto, treinamento e plano de divulgação. A BDTD tem sido atualizada de forma a acompanhar as novas tecnologias de desenvolvimento de *software* para o ambiente *web*. Como resultado, a BDTD já possui parceria com mais de 90 instituições de ensino em todo Brasil e disponibiliza mais de 14 mil trabalhos defendidos na universidade.

A Biblioteca Digital Brasileira [2] é fruto da doação do acervo pessoal do bibliógrafo José Mindlin [19] e sua esposa Guita para a Universidade de São Paulo. O objetivo de tal biblioteca é conservar, divulgar e facilitar o acesso de estudantes, pesquisadores e do público em geral a seu acervo; assim como promover a disseminação de estudos de assuntos brasileiros por meio de programas e projetos específicos, bem como, tornar disponível pela *Internet*, todo o acervo da Biblioteca Brasileira Mindlin [33], que é composta por um acervo de aproximadamente 40 mil volumes incluindo obras de literatura brasileira e portuguesa, relatos de viajantes, manuscritos históricos e literários. A Biblioteca Brasileira Mindlin é considerada a maior e mais importante coleção do gênero formada por um particular. Como resultado do trabalho, a Biblioteca Digital Brasileira já digitalizou mais de 3 mil objetos dentre livros, mapas e imagens em domínio público. A disponibilidade dos textos em edições originais e em épocas distintas permite ao leitor uma interessante viagem às origens dos textos e à História do Brasil.

A *World Digital Library Project* (WDL) [27], promovida pela UNESCO em 2007, tem como objetivos: promover o conhecimento e a conscientização internacionalmente e interculturalmente, expandir o volume e a variedade de conteúdos na *Internet* de forma a prover recursos a professores, pesquisadores e o público em geral além de capacitar as instituições parceiras de forma a reduzir a exclusão digital dentro e entre os países. O projeto prevê a digitalização de documentos, cartas, fotos, mapas e sua apresentação nas seis línguas oficiais da ONU (inglês, francês, espanhol, árabe, chinês e russo) e mais o português.

Segundo [27] a Biblioteca *The European Library*, é um serviço que propõe o acesso livre a recursos de 48 bibliotecas nacionais² em 20 línguas. Os recursos podem ser digitais ou bibliográficos (livros, posters, mapas, registros sonoros, vídeos). É uma organização não comercial, e oferece o acesso a 150 milhões de entradas de registros disponíveis. O volume de coleções digitais referenciadas estão em constante ampliação e a qualidade e confiabilidade são garantidas pelas bibliotecas participantes do projeto.

²Biblioteca Nacional é uma biblioteca especialmente estabelecida pelo país para organizar e preservar informações desse país. Normalmente são responsáveis pelo depósito legal e do centro nacional de registros bibliográficos da nação.

2 Fundamentação Teórica

Esta seção é composta pela fundamentação teórica da biblioteca digital do DECOM/UFOP. A seguir se encontra uma visão geral sobre bibliotecas digitais, que inclui: definição do que são metadados, padrão *Dublin Core* para padronização do uso de metadados, linguagem de marcação XML para transporte de dados e protocolo OAI-PMH para interoperabilidade de dados. Em seguida existem mais duas subseções que descrevem de forma detalhada a BDBComp e a BDCUni.

2.1 Visão geral sobre bibliotecas digitais

Tradicionalmente, uma biblioteca é um lugar em que livros são conservados para gerações futuras, com o objetivo de formar a memória coletiva e facilitar a localização dos mesmos pelas pessoas interessadas. Com o desenvolvimento das redes computacionais, a leitura de documentos se faz cada vez mais por intermédio da tecnologia digital [6]. Entretanto, com o passar do tempo e a popularização de tais redes, o volume de material digitalizado disponível na *internet* se tornou muito grande, dificultando a localização de documentos sobre determinado assunto desejado, já que se encontram espalhados e desorganizados.

A biblioteca digital permite o armazenamento de informação na forma eletrônica, que facilita sua disseminação, pois independe de localização física ou horário. Entretanto o aspecto fundamental para as bibliotecas digitais é a interoperabilidade, potencialidade de redes tecnológicas para trocar e compartilhar documentos, pesquisas e serviços. Para que essa interoperabilidade aconteça, é necessário caracterizar os documentos da biblioteca por meio da definição padrão de metadados. Na biblioteca digital do DECOM, foi usado o *Dublin Core* como padrão de metadados: conjunto de atributos que descrevem informações sobre o autor, ano de publicação, entre outros dados dos documentos que serão abordados mais adiante. Os metadados seguem um protocolo chamado *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) que possui a finalidade de facilitar o compartilhamento de metadados, ou seja, a recuperação e agregação desses dados para poderem ser utilizados em novos serviços.

Em uma biblioteca digital, pode-se encontrar as seguintes características [8]:

- (a) acesso remoto pelo usuário, por meio de um computador conectado a uma rede;
- (b) utilização simultânea do mesmo documento por duas ou mais pessoas;
- (c) inclusão de produtos e serviços de uma biblioteca ou centro de informação;
- (d) existência de coleções de documentos correntes onde se pode acessar não-somente a referência bibliográfica, mas também o seu texto completo;
- (e) acesso a outras fontes externas de informação (bibliotecas, museus, bancos de dados, instituições públicas e privadas);
- (f) busca de informações é facilitada por sistemas inteligentes que recuperam dados mais relevantes;
- (g) utilização de diversos suportes de registro da informação tais como texto, som, imagem e números;

- (h) existência de unidade de gerenciamento do conhecimento, que inclui sistema inteligente para ajudar na recuperação de informação mais relevante.

O conceito biblioteca digital aparenta algo revolucionário, mas na verdade, é resultado de um processo gradual e evolutivo [8]. Desde os anos 70 muitas bibliotecas implementaram catálogos em linha, passaram a acessar bancos de dados, iniciaram o uso regular de *cd-rom* para recuperar referências bibliográficas e textos completos de artigos, periódicos, verbetes de enciclopédias e itens de outras fontes de referência. A partir de 1994, com o fenomenal crescimento da *Internet*, as possibilidades de acessar e recuperar informações aumentaram de forma nunca antes imaginada. No que se refere a projetos de automação de bibliotecas [28], “a maior ênfase foi sempre para os mecanismos de descrição bibliográfica (catálogos e índices), e não no armazenamento e recuperação do próprio conteúdo dos periódicos, livros e outros documentos contidos no acervo. Admite-se que um número crescente de bases de dados legíveis por máquina incorpora os conteúdos completos de livros, artigos de periódicos, relatórios técnicos e outros materiais; muitas bibliotecas, entretanto, continuam a manter tais materiais em suas próprias coleções sob as formas de papel. Os conceitos e as tecnologias da biblioteca digital provêm um enfoque completamente computadorizado ao armazenamento e à recuperação dos materiais bibliotecários”.

2.1.1 Metadados

De acordo com [9] o conceito de metadados antecede a *Internet* e pode ser pensado como dados sobre outros dados. Os metadados estão presentes desde que o primeiro bibliotecário fez uma lista dos itens em uma prateleira de pergaminhos escritos à mão. Metadado é o termo usado na era da *Internet* para obter informações que os bibliotecários têm tradicionalmente colocado em catálogos e, mais comumente, refere-se à informação descritiva sobre recursos *web*. Um registro de metadados consiste em um conjunto de atributos, ou elementos, necessários para descrever o recurso em questão. Por exemplo, um sistema comum de metadados em bibliotecas, consiste no catálogo da biblioteca, que contém um conjunto de registros com os elementos que descrevem um livro ou outro item: autor, título, data de publicação ou assunto.

A maioria dos usuários de tecnologia da informação, provavelmente, já utilizou algum tipo de metadado, mesmo não tendo conhecimento do seu significado e até mesmo do seu uso. Isto é normal, já que a sua própria definição não é um consenso, gerando inúmeras definições de diferentes autores. Abaixo seguem algumas definições:

- (a) Metadados [4] são dados que descrevem atributos de um recurso. Ele suporta um número de funções: localização, descoberta, documentação, avaliação, seleção, etc;
- (b) Metadados [24] são dados associados com objetos que ajudam seus usuários potenciais a terem vantagem completa do conhecimento da sua existência ou características;
- (c) Metadados [26] é o instrumental para transformar dados brutos em conhecimento;
- (d) Segundo [3] “os metadados são conjuntos de atributos, mais especificamente dados referenciais, que representam o conteúdo informacional de um recurso que pode

estar em meio eletrônico ou não. Já os formatos de metadados, também chamados de padrões de metadados, são estruturas padronizadas para a representação do conteúdo informacional que será representado pelo conjunto de dados-atributos. Os metadados são elementos que compõem estruturas para dar tratamento às informações, tendo em vista o gerenciamento informacional”.

No entanto, nenhum desses conceitos é satisfatório, pois há uma pluralidade discursiva a respeito do termo nas diferentes áreas de conhecimento. Para os profissionais da Ciência da Informação [14], o termo metadados está relacionado com o tratamento da informação, mais especificamente com as formas de representação de um recurso informacional para fins de descrição, identificação, localização, busca e recuperação, ou seja, dados bibliográficos e catalográficos que servem para organizar, representar e tornar a informação identificável, localizável e acessível. Formas diferenciadas de representação para um tratamento mais acurado sobre tais informações, com a finalidade de obter uma recuperação eficiente dos recursos informacionais em ambientes digitais e que atendam aos requisitos da interoperabilidade entre os sistemas informacionais, apresentam-se como uma necessidade para a otimização de ambientes informacionais digitais. Para tanto, os metadados é a ferramenta capaz de proporcionar uma forma de representação dos recursos informacionais em meio digital.

Segundo [12] um dos objetivos do uso de metadados no contexto da *web* e das bibliotecas digitais é permitir não só descrever documentos eletrônicos e informações em geral, possibilitando sua avaliação de relevância por usuários humanos, mas também permitir agenciar computadores e programas especiais, robôs e agentes de *software*, para que eles compreendam os metadados associados a documentos e possam então recuperá-los, avaliar sua relevância e manipulá-los com mais eficiência.

Diante disso, pode-se dizer que, sem o uso de metadados [12], não haveria o estabelecimento de uma biblioteca digital de acordo com um conjunto de padrões e normas que garantam o acesso, recuperação, uso e reutilização dos recursos informacionais e que atendam às necessidades dos usuários apresentadas nas solicitações de busca.

De acordo com [30] em todo o mundo a prática de se utilizar padrões de metadados explodiu com o aumento das publicações eletrônicas, resultando em uma grande quantidade de dados digitais disponíveis *online*. A adoção de normas de descrição e práticas de recursos eletrônicos vai melhorar a recuperação de recursos relevantes em qualquer local onde a recuperação da informação é fundamental.

Um dos padrões de metadados existentes, que foi o utilizado nesse trabalho, é o *Dublin Core* e, para permitir a transferência e reunião de informações sobre documentos distribuídos em diversas partes da rede proporcionando visibilidade e integração entre informações, deve-se seguir algum tipo de protocolo como, por exemplo, o protocolo OAI (*Open Archives Initiative*), criando assim um formato padronizado para o intercâmbio de metadados. No ambiente aberto da *web*, o conceito de metadados descritivos padronizados fornece um poderoso mecanismo para melhorar a recuperação de informações. Diante dessa necessidade que surgem padrões de metadados e protocolos. Abaixo são mencionados alguns exemplos de padrão de metadados, além do *Dublin Core*:

- (a) *Resource Description Framework Metadata* (RDF metadata) [35];
- (b) *Open Digital Rights Language Metadata* (ODRL metadata) [18];

(c) *Instructional Management Systems Metadata* (IMS metadata) [7].

2.1.2 Dublin Core

Com a expansão da Internet e consequente desenvolvimento da tecnologia de redes eletrônicas, vêm se intensificando o fenômeno da explosão de documentos eletrônicos [30], ocasionando o aumento do volume de informações disponíveis. Diante dessa realidade, torna-se imprescindível o desenvolvimento de padrões que visem à descrição exata dos recursos de informação. Nesse sentido, várias iniciativas estão sendo conduzidas com o propósito de discutir a questão e propor padrões de descrição de recursos de informação, como é o caso do Dublin Core Metadata Initiative ou somente *Dublin Core*. O *Dublin Core* é a iniciativa para organização e padronização do uso de metadados, concebida em 1994 por Stuard Weisel, pesquisador da Online Computer Library Center (OCLC) [10]. O padrão Dublin Core de metadados é um elemento simples, mas eficaz para descrever uma ampla gama de recursos da rede. O Dublin Core inclui dois níveis: Simples e Qualificado. O Dublin Core Simples inclui quinze elementos, o Qualificado inclui três elementos adicionais (Audiência, Proveniência e Detentor de Direitos), assim como um grupo de refinamentos de elementos (também chamados qualificadores), que refinam a semântica dos elementos de maneira que sejam úteis na descobertas de recursos.

Dublin Core pode ser definido como o conjunto de elementos de metadados planejado para facilitar a descrição de recursos eletrônicos, usado para descrever documentos na Web, com o propósito de ser um meio de comunicação e de procura de informações. Este formato padronizado, de simples utilização, fornece recursos necessários para descrever, identificar, processar, localizar, recuperar e filtrar documentos digitais publicado em rede. É a catalogação do dado ou descrição do recurso eletrônico. A expectativa é que autores ou websites sem conhecimento de catalogação sejam capazes de usar o Dublin Core para descrição de recursos eletrônicos, tornando suas coleções mais visíveis pelos engenhos de busca e sistemas de recuperação. Na maioria dos casos, o conjunto de descritores do Dublin Core é embutido no próprio documento descrito (XML - Extensible Markup Language, HTML e outros), ou, dependendo do recurso, a meta-informação encontra-se separada do recurso catalogado. Os metadados descritos pelo Dublin Core são equivalente a uma ficha catalográfica com o objetivo de fornecer um conjunto básico de elementos de descrição que podem ser usados por catalogadores para simples descrição de recursos de informação [36].

Embora o Dublin Core tenha sido originalmente desenvolvido para descrever documentos como objetos, também pode ser aplicado a outros recursos, como imagens, vídeos, objetos multimídia. Suas principais características são simplicidade na descrição dos recursos, interoperabilidade semântica, escopo internacional e extensibilidade (o que permite sua adaptação às necessidades adicionais de descrição):

- Simplicidade de criação e manutenção

O conjunto de elementos Dublin Core tem se mantido pequeno e simples para permitir o uso por pessoas comuns, sem conhecimento prévio de todas as regras de catalogação e ainda assim consigam criar simples registros descritivos de informação e fornecer para a recuperação efetiva desses recursos no ambiente da rede.

- Interoperabilidade semântica

O Dublin Core pode ajudar o um pesquisador não-especialista a encontrar o que quer oferecendo apoio através de um conjunto comum de elementos. A simples semântica provê entendimento comum em âmbito mundial, por isso aumenta a visibilidade e acessibilidade de todos os recursos.

- Escopo internacional

Embora o Dublin Core tenha sido originalmente desenvolvido em inglês, estão sendo criadas versões em muitas outras línguas, incluindo o português. Isso gera um desenvolvimento comunitário com participação de representantes de quase todos os continentes.

- Extensibilidade

O Dublin Core equilibra as necessidades de simplicidade em descrever recursos digitais com a necessidade de recuperação precisa. Além disso fornece um mecanismo para estender o conjunto de elementos de acordo com as necessidades adicionais de cada recurso. A Tabela 1 apresenta cada elemento e sua respectiva descrição:

Elemento	Descrição
<i>Title</i>	O nome dado ao recurso. Tipicamente, título é um nome dado pelo autor ou editor a um trabalho.
<i>Creator</i>	Uma entidade principalmente responsável por criar o conteúdo do recurso. Uma entidade pode ser uma pessoa, uma organização, ou um serviço. Tipicamente o nome do criador deve ser usado para indicar a entidade.
<i>Subject</i>	O tópico do conteúdo do recurso. Tipicamente, um assunto será expresso como palavras-chave, frases-chaves, selecionadas de um vocabulário controlado.
<i>Description</i>	Uma descrição do conteúdo do recurso. Descrição pode incluir um sumário, tabela de conteúdo, livro-texto do conteúdo ou resumo.
<i>Publisher</i>	Entidade responsável por fazer o documento disponível. Exemplos de publicador incluem uma pessoa, organização ou serviço. Tipicamente o nome de um publicador deveria ser usado para indicar a entidade.
<i>Contributor</i>	Entidade responsável por fazer contribuições ao conteúdo do recurso.
<i>Date</i>	Data da criação ou disponibilização do recurso.
<i>Type</i>	A natureza ou gênero do conteúdo do recurso. Tipo inclui condições que descrevem categorias gerais, funções, gêneros ou agregações. Existe um vocabulário padrão do Dublin Core para este elemento.
<i>Format</i>	Manifestação física ou digital do recurso. Tipicamente, formato pode incluir o tipo de mídia ou dimensão do recurso. Usado para determinar o hardware ou software ou outro equipamento necessário para exibir ou executar o recurso.
<i>Identifier</i>	Identificador único do recurso dentro de um dado contexto. Exemplos de sistemas de identificação incluem a URL ou ISBN.
<i>Source</i>	Referência do qual o presente recurso é derivado (fonte).
<i>Language</i>	Linguagem ou idioma do conteúdo intelectual do recurso.
<i>Relation</i>	Referência a um recurso relacionado.
<i>Coverage</i>	Extensão ou escopo do conteúdo do recurso. Inclui local de espaço tipicamente (um nome de lugar ou coordenadas geográficas), período temporal ou jurisdição.
<i>Rights Management</i>	Informações sobre os direitos inerentes ao recurso: informações sobre propriedade intelectual, direitos autorais e de propriedade.

Tabela 1: Elementos Dublin Core

O esquema formado somente pelos 15 elementos é chamado de Dublin Core Simples (não qualificado). É simples, pois visa possibilitar que as pessoas não especializadas em catalogação e indexação possam usa-los para descrever seus recursos. Restringe-se a elementos que possibilitam a descoberta do recurso, através de motores de busca. Entretanto existem Dublin Core Qualifiers, que refinam ou qualificam os elementos de metadados do Dublin Core, afim de tornar a descrição mais precisa. Se algum usuário não for capaz de entender o refinamento, deve ignorar o qualificador e tratar como um esquema Dublin Core Simples. No caso da biblioteca digital do DECOM foi necessária a adição de mais quatro atributos sendo esses: observação, banca, data de submissão e data de modificação: A Tabela 2 apresenta os elementos adicionais e sua respectiva descrição.

Elemento	Descrição
<i>banca</i>	Banca. Contém os nomes de quem participou da banca na defesa da monografia.
<i>observation</i>	Observação. Armazena algum dado relevante de alguma monografia. Não será necessariamente usado.
<i>modified</i>	Data de modificação. Armazena a data da última modificação.
<i>dateSubmitted</i>	Data de submissão. Armazena a data de submissão da monografia.

Tabela 2: Elementos Qualified Dublin Core

Existe uma forma de se diferenciar os atributos específicos do *Dublin Core* dos atributos qualificadores. Esta diferenciação será apresentada mais adiante na sessão sobre XML.

2.1.3 Extensible Markup Language XML

A XML [34] é um formato de texto muito simples e flexível. Originalmente concebido para enfrentar os desafios de publicação eletrônica em grande escala. XML também está desempenhando um papel cada vez mais importante na troca de uma ampla variedade de dados na Web. XML é popular por muitos motivos, mas o principal é facilitar o tratamento e troca de dados. Em XML dados e marcação são armazenados como texto que é facilmente editável. Os dados também não estão codificados e, por isso são mais acessíveis.

No XML os metadados aparecem na forma de *tag* [32], que em inglês significa etiqueta. As *tags* na internet são palavras que servem justamente como uma etiqueta e ajudam na hora de organizar informações, agrupando aquelas que receberam a mesma marcação, facilitando encontrar outras relacionadas. Esse sistema de classificação não é novo e já era utilizado em bibliotecas para organizar livros em temas semelhantes. Se alguém queria procurar livros sobre criação de gatos, por exemplo, bastava procurar por essa palavra-chave.

Existe uma enorme liberdade em XML [17], porque você define suas próprias *tags*. No entanto, à medida que mais aplicações XML foram sendo desenvolvidas, surgiu o problema de conflitos de nome de *tag*. A solução é usar *namespaces*. *Namespaces* permitem garantir que um conjunto de *tags* não poderá entrar em conflito com outro.

Namespaces permitem que você inclua um nome seguido por um sinal de dois pontos antes do nome da *tag*, evitando que entrem em conflito. A figura abaixo apresenta uma estrutura de XML:

```
1 <bookstore>
2   <book category="CHILDREN">
3     <title>Harry Potter</title>
4     <author>J K. Rowling</author>
5     <year>2005</year>
6     <price>29.99</price>
7   </book>
8   <book category="WEB">
9     <title>Learning XML</title>
10    <author>Erik T. Ray</author>
11    <year>2003</year>
12    <price>39.95</price>
13  </book>
14 </bookstore>
```

Figura 1: Exemplo de estrutura XML

2.1.4 Protocolo OAI-PMH

A falta de padrões, para disponibilização de informações na Web, motivou a criação da *Open Archives Initiative* (OAI), que [5] “desenvolve e promove soluções de interoperabilidade que visam facilitar a disseminação eficiente de conteúdo”, pois, com sua utilização, torna-se possível proporcionar simplicidade e eficiência na pretensão de unificar as consultas a bases de dados científicas e acadêmicas. A OAI [21] tem suas raízes em um esforço para melhorar o acesso a arquivos como um meio de aumentar a disponibilidade de comunicação científica. Protocolo OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*) é um protocolo desenvolvido pela Open Archives Initiative, utilizado para distribuir e coletar metadados principalmente relativos a descritores de documentos. Uma das aplicações mais comuns do OAI-PMH é a interoperabilidade entre repositórios, envolvendo um ou mais repositórios, que fornecem os seus metadados, e uma entidade receptora que recolhe esses metadados. O protocolo tem como finalidade facilitar o compartilhamento de metadados, ou seja, a recuperação e agregação desses dados para poderem ser utilizados em serviços novos, onde o mais comum é a busca federada ou metabusca.

Busca federada [31] caracteriza-se pela transmissão de uma pergunta a um grupo de bases de dados dispersas, agrupando os resultados coletados das bases de dados e apresentando-os em um único resultado. Pode ser definida como uma técnica pela qual as fontes de informação múltiplas podem ser pesquisadas simultaneamente, incluindo catálogos de bibliotecas, bases de dados comerciais, a web, recolhendo a informação a partir de várias fontes e fornecendo um resultado de busca integrado ao usuário. A busca federada permite aos usuários encontrar documentos, recursos, através das coleções da biblioteca, não sendo necessário conhecer uma base de dados específica, nem saber como as diferentes interfaces de busca funcionam, e utilizando uma caixa de busca única.

O usuário [23] utilizando um programa de busca federada, previamente configurado pela biblioteca, coloca sua pergunta, seu termo de busca; o programa leva a pergunta

para as diferentes fontes, como uma base de dados comercial que a biblioteca tenha assinatura, para o próprio catálogo da biblioteca e para a web, buscando em bases de dados de acesso aberto. O metabuscador recebe as respostas das diferentes fontes e as reúne para o usuário apresentando-as em uma única resposta. Os metabuscadores possuem programas que classificam o grupo de resultados obtidos estabelecendo a relevância a partir dos metadados obtidos. Os resultados podem ainda ser ordenados e filtrados por diferentes critérios, como data, tipo de material (livros, artigo, etc.) e fonte (somente itens do catálogo da biblioteca), dentre vários outros.

Um sistema, para estar em conformidade com a iniciativa OAI, deve conter dados para compartilhamento. Com os recursos oferecidos pela iniciativa, é possível melhorar significativamente a precisão das consultas eletrônicas e reduzir o tempo de procura. O protocolo OAI-PMH para a coleta de metadados fornece uma aplicação independente de estrutura de interoperabilidade com base em coleta de metadados. Há duas classes de participantes no âmbito do protocolo OAI-PMH:

1. Provedores de Dados - mantêm repositórios de documentos digitais que implementam o protocolo OAI como forma de expor os metadados de seus documentos.
2. Provedores de serviço - oferecem buscas a metadados, pois coletam, organizam e disponibilizam estes a fim de agregar valor à Iniciativa.

O protocolo OAI-PMH introduz o conceito de *Metadata Harvesting* (colheita de metadados), um processo unilateral ², onde, os provedores de serviços, a partir da lista de repositórios (provedores de dados) registrados na OAI, realizam periodicamente uma busca a estes provedores de dados, “colhendo” os metadados para exibição sob a forma de consultas efetuadas pelos usuários.



Figura 2: Metadata Harvesting [25]

O protocolo utiliza um ferramental de transações http muito simples, baseadas em request-response. A resposta deve ser no formato XML. Um harvester pode selecionar os metadados que deseja coletar de forma a pegar somente os novos ou os que foram modificados desde a última interação de coleta com o repositório. Pode também restringir os metadados que deseja coletar ao indicar qual é o conjunto de dados de interesse.

Elemento	Descrição
<i>Identify</i>	Utilizado para recuperar as informações sobre os provedores de dados. A resposta mínima esperada consiste em: nome do provedor de dados, endereço do repositório (URL), versão do protocolo e endereço eletrônico do administrador do repositório.
<i>ListMetadata-Formats</i>	Recupera os formatos de metadados disponíveis no repositório. O formato padrão é Dublin Core.
<i>GetRecord</i>	Recupera um único registro do repositório. Deve-se especificar o formato dos metadados, cujo padrão é Dublin Core, e o identificador do registro (Identifier), que é único na comunidade OAI.
<i>ListRecords</i>	Realiza a colheita dos metadados do repositório.
<i>ListIdentifiers</i>	É uma versão abreviada do verbo ListRecords, que retorna apenas os cabeçalhos dos registros.
<i>ListSets</i>	Lista a estrutura do repositório.

Tabela 3: Verbos do protocolo OAI-PMH

O protocolo OAI-PMH apresenta 6 verbos [11], utilizados para recuperar as informações, conforme a tabela abaixo 3:

O protocolo OAI-PMH, vem se consolidando como a base para a interoperabilidade entre bibliotecas/repositórios digitais acadêmicos e científicos no mundo todo. Através do OAI-PMH é possível proporcionar visibilidade e integração de informações (metadados), com custos acessíveis à realidade de países em desenvolvimento, como o Brasil.

2.2 BDBComp

A Biblioteca Digital Brasileira de Computação (BDBComp) disponibiliza e permite o acesso aos artigos publicados nos eventos promovidos pela SBC. A BDBComp, tem por finalidade prestar serviços de informações à comunidade da área de Informática e Computação apoiando as atividades de ensino, pesquisa e extensão. A BDBComp foi concebida com o objetivo de arquivar, organizar, preservar dados digitais e também disseminar o conhecimento científico produzido pela comunidade brasileira de computação, auxiliando a recuperação da informação desejada pelos usuários. A BDBComp possui atualmente 8288 trabalhos publicados em periódicos nacionais e anais de eventos realizados no Brasil. Além disso, conta com um serviço de auto-arquivamento para que pesquisadores, estudantes e profissionais possam submeter trabalhos científicos em seu repositório.

2.3 BDCUNI

A Biblioteca Digital de Computação do Uni-BH (BDCUni) foi criada para o curso de ciência da computação do Centro Universitário de Belo Horizonte (UNI-BH), com o objetivo de efetuar o arquivamento de todos os trabalhos monográficos desenvolvidos na disciplina "Projeto Orientado" do curso, promovendo assim, a difusão da produção científica para a comunidade de computação. A interface da BDCUni foi criada utili-

zando a linguagem PHP. Já os dados foram armazenados no sistema de gerenciamento de bancos de dados relacional MySQL.

3 Desenvolvimento da Biblioteca Digital

Esta seção é composta pela proposta de confecção da biblioteca digital do DECOM/UFOP para o armazenamento de trabalhos monográficos. A seguir se encontra o mapeamento de dados do padrão *Dublin Core* para a biblioteca digitaldo e, em seguida, a especificação do esquema conceitual e relacional do banco de dados e, também, um exemplo de documento XML que será utilizado na biblioteca digital.

3.1 Modelagem de Dados

A partir de reuniões com futuros usuários da biblioteca digital do DECOM/UFOP, foi possível detectar quais informações devem ser armazenadas sobre trabalhos monográficos. De acordo com esses encontros conseguimos concluir que é necessário armazenar o seguinte conjunto de dados: título, nome completo dos autores, palavras-chave, descrição, pessoa/entidade responsável por publicar o trabalho monográfico, nome completo do professor orientador, data de realização da defesa do trabalho monográfico, tipo do conteúdo (monografia, artigo), formato de divulgação, identificador único, idioma, áreas de abrangência do trabalho monográfico, referências bibliográficas, co-orientadores, data de submissão, data da última modificação, observações e composição da banca avaliadora.

A partir das necessidades da biblioteca digital do DECOM foi possível mapear os metadados do padrão *Dublin Core* e fazer com que atendessem a essas necessidades. Além do *Dublin Core* simples, foi necessária a utilização de qualificadores. Na biblioteca digital do DECOM/UFOP são utilizados os seguintes qualificadores: *coContributor*, *Modified*, *DateSubmitted*, *Obsrvation* e *Banca*. O padrão *Dublin Core* tem como vantagem a extensibilidade e, por isso é recomendado como padrão de uso para a internet, pois reflete necessidades específicas da *web*. A tabela 4 apresenta o mapeamento dos dados do padrão *Dublin Core* para as necessidades da biblioteca digital do DECOM/UFOP. Traçando um paralelo de como cada elemento do *Dublin Core* foi adaptado para a realidade da biblioteca digital do DECOM/UFOP. A coluna nomeada de categorização informa se o metadado é simples ou composto e multivalorado ou monovalorado.

Alguns dos atributos especificados para a biblioteca digital do DECOM têm especificidades que devem ser ressaltadas:

- (a) Os atributos *creator* e *contributor* do padrão *Dublin Core* são representados como multivalorados, apesar da disciplina monografia permitir, apenas, um orientador e um autor. Isto aconteceu, pois haverá demanda em trabalhos futuros.
- (b) O atributo *identifier* será montado a partir da concatenação das seguintes informações: número de matrícula + /DECOM.UFOP. + *dateSubmitted*. Um exemplo seria: 0714160/DECOM.UFOP.2012.11.10.
- (c) Devem ser armazenados os seguintes dados dos componentes da banca: nome, graduação e instituição.

Dublin Core	Monografia	Categorização
<i>title</i>	Título	Monovalorado/Simples
<i>creator</i>	Autor	Multivalorado/Simples
<i>subject</i>	Palavras-chave	Multivalorado/Simples
<i>description</i>	Descrição	Monovalorado/Simples
<i>publisher</i>	Pessoa/entidade responsável por divulgar o trabalho monográfico	Monovalorado/Simples
<i>contributor</i>	Orientador	Multivalorado/Simples
<i>date</i>	Data de realização da defesa	Monovalorado/Simples
<i>type</i>	Tipo do conteúdo, que pode ser: <i>text</i> , <i>image</i>	Monovalorado/Simples
<i>format</i>	Formato de divulgação, que pode ser: PDF, HTML	Multivalorado/Simples
<i>identifier</i>	Identificador único	Monovalorado/Simples
<i>language</i>	Idioma	Monovalorado/Simples
<i>relation</i>	Áreas de abrangência do trabalho monográfico	Multivalorado/Simples
<i>source</i>	URL	Monovalorado/Simples
<i>coContributor</i>	Co-orientadores	Multivalorado/Simples
<i>modified</i>	Data da última modificação	Monovalorado/Simples
<i>dateSubmitted</i>	Data de submissão	Monovalorado/Simples
<i>observation</i>	Observações específicas de cada monografia	Monovalorado/Simples
<i>banca</i>	Banca avaliadora do trabalho monográfico	Multivalorado/Simples
<i>category</i>	Categoria do trabalho, que pode ser: monografia, relatório técnico-científico	Monovalorado/Simples
<i>references</i>	Referências bibliográficas	Multivalorado/Simples

Tabela 4: Mapeamento do *Dublin Core* para a biblioteca digital do DECOM

Alguns atributos de metadados do *Dublin Core* não foram incluídos por não se aplicarem à proposta da biblioteca digital do DECOM/UFOP. São eles:

- (a) *Coverage*: área geográfica que o trabalho monográfico engloba.
- (b) *Rights Management*: direitos ou outras propriedades intelectuais, especificando as condições através das quais o trabalho monográfico pode ou não ser usado.

De posse do modelo conceitual da biblioteca digital do DECOM/UFOP e do mapeamento de metadados seguindo o padrão *Dublin Core* foi possível construir o modelo relacional. A figura 3 apresenta o modelo Entidade-Relacionameto da biblioteca digital do DECOM/UFOP:

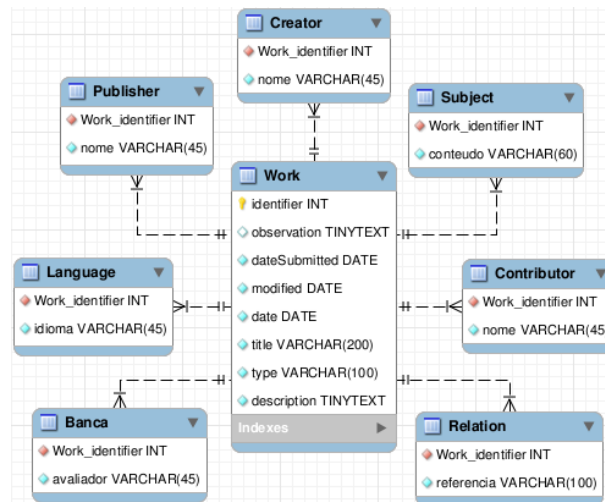


Figura 3: Modelagem de dados usando o modelo Entidade-Relacionamento

A partir dos dados inseridos no banco de dados é possível construir a estrutura do XML da biblioteca digital do DECOM/UFOP. Os atributos que definem o escopo do *Dublin Core* Simple são diferenciados dos qualificadores pelos *namespaces*. A figura 4 contém o exemplo de XML da biblioteca digital do DECOM/UFOP:

```

1 <?xml version="1.0"? encoding="iso-8859-1">
2 <metadata
3   xmlns="http://www.decom.ufop.br/bddecom/"
4   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
5   xsi:schemaLocation="http://www.decom.ufop.br/bddecom/schema.xsd"
6   xmlns:dc="http://purl.org/dc/elements/1.1/"
7   xmlns:dcterms="http://purl.org/dc/terms/">
8
9   <dc:title>Proposta de uma Biblioteca Digital para Documentos Técnico-científicos do Departamento
10  de Computação da Universidade Federal de Ouro Preto
11  </dc:title>
12  <dc:creator>Gustavo Henrique Braz da Silva</dc:creator>
13  <dc:subject>metadado</dc:subject>
14  <dc:description>Este trabalho tem como objetivo desenvolver uma biblioteca digital para armazenar
15  os trabalhos monográficos do DECOM/UFOP
16  </dc:description>
17  <dc:publisher>DECOM</dc:publisher>
18  <dc:contributor>Guilherme Tavares de Assis</dc:contributor>
19  <dc:date>2012-10-30</dc:date>
20  <dc:type>text</dc:type>
21  <dc:format>pdf</dc:format>
22  <dc:identifier>0714160/DECOM.UFOP.2012.11.06</dc:identifier>
23  <dc:language>português</dc:language>
24  <dc:relation>Recuperação de informação</dc:relation>
25  <dc:source>decom.ufop.br/monografia/0714160/DECOM.UFOP.2012.11.06.pdf</dc:source>
26  <dcterms:coContributor>João da Silva</dcterms:coContributor>
27  <dcterms:modified xsi:type="dcterms:W3CDTF">2012-10-30</dcterms:modified>
28  <dcterms:dateSubmitted xsi:type="dcterms:W3CDTF">2012-11-06</dcterms:dateSubmitted>
29  <dcterms:observation>Monografia 1. Depende da implementação em monografia 2.</dcterms:observation>
30  <dcterms:banca>João, Maria, José</dcterms:banca>
31  <dcterms:category>Monografia</dcterms:category>
32  <dcterms:references>http://dublincore.org/documents/dc-xml-guidelines/#DCTERMS</dcterms:references>
33 </metadata>
  
```

Figura 4: XML biblioteca digital DECOM/UFOP

4 Anexo - Script do Banco de Dados

Programa 1: ListaVazia

```
SET @OLD_UNIQUE_CHECKS=@@UNIQUE_CHECKS, UNIQUE_CHECKS=0;
SET @OLD_FOREIGN_KEY_CHECKS=@@FOREIGN_KEY_CHECKS, FOREIGN_KEY_CHECKS=0;
SET @OLD_SQL_MODE=@@SQL_MODE, SQL_MODE='TRADITIONAL,ALLOW_INVALID_DATES';
```

```
CREATE SCHEMA IF NOT EXISTS 'bddecom' DEFAULT CHARACTER SET latin1 ;
USE 'bddecom' ;
```

— Table 'bddecom'.'Work'

```
CREATE TABLE IF NOT EXISTS 'bddecom'.'Work' (
  'identifier' INT NOT NULL ,
  'observation' TINYTEXT NULL ,
  'dateSubmitted' DATE NOT NULL ,
  'modified' DATE NOT NULL ,
  'date' DATE NOT NULL ,
  'title' VARCHAR(200) NOT NULL ,
  'type' VARCHAR(100) NOT NULL ,
  'description' TINYTEXT NOT NULL ,
  PRIMARY KEY ('identifier') )
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8
COLLATE = utf8_unicode_ci;
```

— Table 'bddecom'.'Subject'

```
CREATE TABLE IF NOT EXISTS 'bddecom'.'Subject' (
  'Work_identifier' INT NOT NULL ,
  'conteudo' VARCHAR(60) NOT NULL ,
  CONSTRAINT 'fk_Subject_Work1'
    FOREIGN KEY ('Work_identifier' )
    REFERENCES 'bddecom'.'Work' ('identifier' )
    ON DELETE CASCADE
    ON UPDATE CASCADE)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8
COLLATE = utf8_unicode_ci;
```

— Table 'bddecom'.'Relation'

```

CREATE TABLE IF NOT EXISTS 'bddecom'.'. 'Relation' (
    'Work_identifier' INT NOT NULL ,
    'referencia' VARCHAR(100) NOT NULL ,
    CONSTRAINT 'fk_Relation_Work1'
        FOREIGN KEY ('Work_identifier' )
        REFERENCES 'bddecom'.'. 'Work' ('identifier' )
        ON DELETE CASCADE
        ON UPDATE CASCADE)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8
COLLATE = utf8_unicode_ci;

```

— Table 'bddecom'.'. 'Publisher'

```

CREATE TABLE IF NOT EXISTS 'bddecom'.'. 'Publisher' (
    'Work_identifier' INT NOT NULL ,
    'nome' VARCHAR(45) NOT NULL ,
    CONSTRAINT 'fk_Publisher_Work1'
        FOREIGN KEY ('Work_identifier' )
        REFERENCES 'bddecom'.'. 'Work' ('identifier' )
        ON DELETE CASCADE
        ON UPDATE CASCADE)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8
COLLATE = utf8_unicode_ci;

```

— Table 'bddecom'.'. 'Language'

```

CREATE TABLE IF NOT EXISTS 'bddecom'.'. 'Language' (
    'Work_identifier' INT NOT NULL ,
    'idioma' VARCHAR(45) NOT NULL ,
    CONSTRAINT 'fk_Language_Work1'
        FOREIGN KEY ('Work_identifier' )
        REFERENCES 'bddecom'.'. 'Work' ('identifier' )
        ON DELETE CASCADE
        ON UPDATE CASCADE)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8
COLLATE = utf8_unicode_ci;

```

— Table 'bddecom'.'. 'Creator'

```

CREATE TABLE IF NOT EXISTS 'bddecom'.'Creator' (
  'Work_identifier' INT NOT NULL ,
  'nome' VARCHAR(45) NOT NULL ,
  CONSTRAINT 'fk_Creator_Work'
    FOREIGN KEY ('Work_identifier' )
    REFERENCES 'bddecom'.'Work' ('identifier' )
    ON DELETE CASCADE
    ON UPDATE CASCADE)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8
COLLATE = utf8_unicode_ci;

```

— Table 'bddecom'.'Contributor'

```

CREATE TABLE IF NOT EXISTS 'bddecom'.'Contributor' (
  'Work_identifier' INT NOT NULL ,
  'nome' VARCHAR(45) NOT NULL ,
  CONSTRAINT 'fk_Contributor_Work1'
    FOREIGN KEY ('Work_identifier' )
    REFERENCES 'bddecom'.'Work' ('identifier' )
    ON DELETE CASCADE
    ON UPDATE CASCADE)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8
COLLATE = utf8_unicode_ci;

```

— Table 'bddecom'.'Banca'

```

CREATE TABLE IF NOT EXISTS 'bddecom'.'Banca' (
  'Work_identifier' INT NOT NULL ,
  'avaliador' VARCHAR(45) NOT NULL ,
  CONSTRAINT 'fk_Banca_Work1'
    FOREIGN KEY ('Work_identifier' )
    REFERENCES 'bddecom'.'Work' ('identifier' )
    ON DELETE CASCADE
    ON UPDATE CASCADE)
ENGINE = InnoDB
DEFAULT CHARACTER SET = utf8
COLLATE = utf8_unicode_ci;

```

```
SET SQL_MODE=@OLD_SQL_MODE;  
SET FOREIGN_KEY_CHECKS=@OLD_FOREIGN_KEY_CHECKS;  
SET UNIQUE_CHECKS=@OLD_UNIQUE_CHECKS;
```

Referências

- [1] Pablo A. Robert. *The Web-DL Environment for Building Digital Libraries from the Web*. artigo, publicado na Joint Conference on Digital Libraries, 2003.
- [2] Gabriela Agustini. Edson gommi detalha a infraestrutura usada pela biblioteca digital brasileira. <http://culturadigital.br/simposioacervosdigitais/2010/04/08/edson-gomi-fala-sobre-a-experiencia-da-biblioteca-digital-brasiliana/>, abril 2010.
- [3] R.C.V. Alves. Web semântica: uma análise focada no uso de metadados. *Dissertação (Mestrado em Ciência da Informação) Universidade Estadual Paulista*, 2005.
- [4] Monty Bieber. Data warehousing in government. *DM Review Magazine*, may 1998.
- [5] Lagoze C. and Van de Sompel H. The open archives initiative: Building a low-barrier interoperability framework. *Proceeding of the ACM/IEEE Joint Conference on Digital Libraries, Roanoke VA*, jun 2001.
- [6] Gary Cleveland. Digital libraries: Definitions, issues and challenges. 1998.
- [7] IMS Global Learning Consortium. Instructional management systems (ims). <http://www.imsglobal.org/>.
- [8] Murilo Bastos da Cunha. Desafios na construção de uma biblioteca digital. *Universidade de Brasilia*, dez 1999.
- [9] Rafael Port da Rocha. Metadados, websemântica, categorização automática: combinando esforços humanos e computacionais para a descoberta e uso dos recursos da web. *UFRGS - Porto Alegre*, junho 2004.
- [10] Celeste Magela Campos da Silva e Guilherme Tavarez de Assis. Proposta de uma biblioteca digital para trabalhos monográficos. 2006.
- [11] Patrícia de Andrade Bueno Garcia and Marcos Sfair Sunye. O protocolo oai-pmh para interoperabilidade em bibliotecas digitais. *Universidade Federal do Paraná*.
- [12] Fabiano Ferreira de Castro and Plácio Leopoldina V.A. da Costa Santos. Os metadados como instrumentos tecnológicos na padronização e potencialização dos recursos informacionais no Âmbito das bibliotecas digitais na era da web semântica. *Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista-UNESP*, agosto 2007.
- [13] E. A. FOX and G. MARCHIONINI. *Digital libraries*. Communications of the ACM, May 2001.
- [14] A. Gilliland-Swetland. Defining metadata. *Introduction to Metadata: Pathways to Digital Information, Los Angeles: Getty Information Institute*, 1998.

- [15] Marcelo Santana Graça Pimentel, Liliane Bernardes. *Biblioteca Escolar - Pró funcionário: Curso tecnico de formação para os funcionários da educação*. Universidade de Brasília - UnB, 2007.
- [16] Diane Hillmann. Using dublin core. novembro,. <http://dublincore.org/documents/2005/11/07/usageguide/>.
- [17] Steven Holzner. *Desvendando XML*. Editora Campus, 2001.
- [18] ODRL Initiative. Open digital rights language (odrl). <http://www.w3.org/community/odrl/>.
- [19] Academia brasileira de letras. <http://www.academia.org.br>, 2012.
- [20] Rogerio Toshiaki Kondo, Maria de Lourdes Rebucci Lirani, and Caetano Traina jr. Bibliotecas digitais: a experiência da usp. *Revista USP, São Paulo*, fevereiro 2009.
- [21] Carl Lagoze, Herbert Van de Sompel, Michael Nelson, and Simeon Warner. The open archives initiative protocol for metadata harvesting. <http://www.openarchives.org/OAI/openarchivesprotocol.html>, dezembro 2008.
- [22] Jakob Nielsen Hoa Loranger. *Usabilidade na Web - Projetando Websites com qualidade*. Editora CAMPUS, 2007.
- [23] Ana Maria Neves Maranhão. Dos catálogos aos metabuscadores e serviços de descoberta na internet. *XXIV Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação - Maceió Alagoas*, agosto 2011.
- [24] Terry Moriaty. What is metadata? database programming and desing. *San Mateo*, july 1997.
- [25] Garcia P.A.B. Provedores de dados de baixo custo: Publicação digital ao alcance de todos. *Dissertação de Mestrado. Curitiba. Setor de Ciências Exatas. Universidade Federal do Paraná*, 2003.
- [26] Gilberto Pessanha Ribeiro. Metadados geoespaciais digitais. *Workshop de Banco de Dados Não Convencionais*, 1995.
- [27] Marcia Rosetto. Bibliotecas digitais - cenario e perspectivas. *Revista Brasileira de Biblioteconomia e Documentação*, julho 2008.
- [28] William Saffady. Digital libraries. *Library Technology Report*, jun 1995.
- [29] SBC. Sociedade brasileira de computação. <http://www.sbc.org.br/>.
- [30] Marcia Izabel Fugisawa Souza, Laurimar Gonçalves Vendrusculo, and Geane Cristina Melo. Metadados para a descrição de recursos de informação eletrônica: utilização do padrão dublin core. *Embrapa Informática Agropecuária*, abril 2000.
- [31] Ferreira Sueli Mara S.P. and Souto Leonardo Fernandes. Dos sistemas de informação federados à federação de bibliotecas digitais. *Revista Brasileira de Biblioteconomia e Documentação, São Paulo*, jun 2006.

- [32] Tecmundo. O que é tag? <http://www.tecmundo.com.br/navegador/2051-o-que-e-tag-.htm>.
- [33] Brasiliana USP. Biblioteca brasileira guita e josé mindlin. <http://www.brasiliana.usp.br/node/504>.
- [34] World Wide Web Consortium (W3C). Extensible markup language (xml). <http://www.w3.org/XML/>.
- [35] World Wide Web Consortium (W3C). Resource description framework (rdf). <http://www.w3.org/RDF/>.
- [36] S. Weibel. The dublin core: a simple content description model for electronic resources. *Bulletin of the American Society for Information Science*, Nov 1997.