

2 Definições Conceituais

Neste capítulo são definidos e/ou explicados conceitos, padrões ou tecnologias utilizadas e sua utilidade nessa dissertação.

2.1 Padrões de Metadados Utilizados

Metadados, por definição, são dados que descrevem os dados e que têm como objetivo informar-nos sobre eles para tornar mais fácil a sua organização. "Metadados são dados estruturados e codificados que descrevem as características de entidades provedoras de informação para ajudar na identificação, na descoberta, na avaliação e na gerência das entidades descritas" [8].

2.1.1 IMS LIP

O IMS *Learner Information Package* (LIP) [9] é um modelo de dados que descreve características de um estudante.

Os elementos de dados que descrevem as características de um estudante são agrupados em 11 (onze) categorias, conforme definido no IMS LIP Model Specification v1.0 (item 1.1.3):

Categoria	Descrição
<i>Identification</i>	Agrupar informações sobre dados biográficos e demográficos relevantes para o aprendizado.
<i>Goal</i>	Descreve os objetivos e aspirações pessoais do estudante. Pode ainda conter informações para monitorar o progresso na conquista desses objetivos.
<i>QCL</i>	Agrupar informações sobre as qualificações, certificações e licenças (licença para exercer medicina, advocacia, etc.) conquistadas pelo estudante.
<i>Activity</i>	Qualquer atividade relacionada ao aprendizado em andamento, por exemplo, treinamentos, estágios e serviço militar.
<i>Transcript</i>	Agrupar informações sobre o desempenho acadêmico de um estudante em uma instituição.
<i>Interest</i>	Agrupar informações sobre hobbies e outras atividades

	recreativas realizadas pelo estudante.
<i>Competency</i>	Agrupa informações sobre as habilidades adquiridas pelo estudante. Essas habilidades podem estar ou não associadas a algum treinamento formal ou trabalho realizado pelo estudante.
<i>Affiliation</i>	Armazena informações sobre as organizações profissionais das quais o estudante faz parte.
<i>Accessibility</i>	Armazena informações de acessibilidade genéricas sobre o estudante, tais como, conhecimento de línguas, deficiências e preferências cognitivas.
<i>Securitykey</i>	Armazena o conjunto de senhas e chaves de segurança do estudante para transações com sistemas de informação de estudantes e serviços.
<i>Relationship</i>	Descreve as possíveis relações existentes entre dados das outras categorias.

Tabela 2.1: Categorias do IMS LIP

O IMS LIP é utilizado juntamente com o IMS ACCLIP para descrever o perfil do usuário, o que é explicado mais adiante.

2.1.2 IMS ACCLIP e IMS ACCMD

O IMS Accessibility for LIP (ACCLIP) [10] descreve como os estudantes podem interagir com um ambiente de estudos on-line baseado nas suas necessidades e preferências.

O ACCLIP adiciona 2 novos elementos (e seus sub-elementos) ao LIP, são eles:

O **AccessForAll** que define as preferências de acessibilidade que foram deixadas como “trabalho futuro” na primeira versão do LIP. Como o próprio nome diz o elemento AccessForAll é feito para armazenar as necessidades e preferências de todas as pessoas, não somente das que possuem algum tipo de deficiência. Visa beneficiar usuários em situações de aprendizagem que requerem mídias alternativas, como em um ambiente extremamente barulhento, onde legendas são necessárias para o entendimento de um vídeo.

O **Acomodation**, que permite a descrição das acomodações⁴ que foram feitas para a interação com um objeto de aprendizagem específico (ou com um conjunto de objetos de aprendizagem).

Já a especificação IMS AccessForAll Meta-data (ACCMD) [10] foi criada com o objetivo de permitir a entrega de conteúdos que se encaixem com as necessidades ou preferências do usuário. O XSD do IMS AccessForAll Meta-data se encontra no item A.1, XSD do IMS AccessForAll.

Em conjunto, a IMS ACCLIP e a IMS ACCMD, provêm os dois lados da combinação necessária para endereçar as necessidades e preferências dos estudantes. Uma especifica o que o estudante precisa ou prefere, enquanto a outra identifica os recursos utilizando os mesmos termos.

O IMS ACCMD é utilizado para descrever as características de acessibilidade dos materiais de ensino que são reordenados e possivelmente filtrados segundo essas características.

2.1.3 Dublin Core

É um conjunto de metadados padrão para a descrição de fontes de informação de diferentes domínios, onde uma fonte de informação é definida como “qualquer coisa que possua uma identidade” [11].

Estrutura dos Metadados:

Atributo	Descrição
<i>Identifier</i>	Identificação não ambígua do recurso dentro de um dado contexto.
<i>Contributor</i>	Entidade responsável pela contribuição ao conteúdo do recurso.
<i>Coverage</i>	Extensão ou cobertura espaço-temporal do conteúdo do recurso.
<i>Creator</i>	Entidade principal responsável pela elaboração do conteúdo do recurso.
<i>Date</i>	Data associada a um evento no ciclo de vida do recurso.
<i>Description</i>	Descrição sobre o conteúdo do recurso.
<i>Format</i>	Manifestação física ou digital do recurso.
<i>Language</i>	Idioma do conteúdo intelectual do recurso.
<i>Publisher</i>	É a instituição responsável pela difusão do recurso.

⁴ Permite especificar as acomodações necessárias para que um estudante utilize um determinado material de ensino. Ex: Uma sala privativa equipada com um computador com um display de braille atualizável para que um aluno cego possa fazer uma prova.

<i>Relation</i>	Uma referência a um outro recurso que se relaciona com o recurso em questão.
<i>Rights</i>	Informações sobre os direitos do recurso e seu uso.
<i>Source</i>	Uma referência para um outro recurso que tenha dado origem ao presente recurso.
<i>Subject</i>	Assunto referente ao conteúdo do recurso.
<i>Title</i>	Título dado ao recurso.
<i>Type</i>	A natureza ou gênero do conteúdo do recurso.

Tabela 2.2: Atributos do Dublin Core

Os padrões Dublin Core, Ariadne e IEEE LOM são padrões de metadados de materiais de ensino aceitos pelo LORIS e conseqüentemente importantes no contexto desta dissertação.

2.1.4 ARIADNE

O metadado educacional ARIADNE (*Alliance of Remote Instructional Authoring and Distribution Networks for Europe*) [12] foi derivado de trabalhos e experimentos desenvolvidos desde 1995, por instituições internacionais. O Ariadne cobre as necessidades de comunidades educacionais ou de treinamento que tenham diversidade lingüística e favorece em larga escala o compartilhamento e reuso de recursos de conhecimento.

Para garantir a simplicidade, entendimento e adaptabilidade da comunidade Ariadne, os elementos de dados são agrupados em sete categorias:

Categoria	Descrição
<i>General</i>	Agrupa informações gerais que descrevem o objeto de aprendizagem.
<i>Semantics</i>	Agrupa elementos que descrevem a classificação semântica do objeto de aprendizagem.
<i>Pedagogical</i>	Agrupa elementos que descrevem características pedagógicas e educacionais do objeto de aprendizagem.
<i>Technical</i>	Agrupa elementos que descrevem os requisitos e características técnicas do objeto de aprendizagem.
<i>Conditions</i>	Agrupa elementos que descrevem os direitos de propriedade intelectual e condições de uso do objeto de aprendizagem.
<i>Meta-metadata</i>	Agrupa elementos que descrevem informações sobre a própria instância do metadado.
<i>Annotation</i>	Agrupa elementos opcionais que descrevem pessoas ou organizações que fazem anotações sobre os objetos de aprendizagem.

Tabela 2.3: Categorias do Ariadne

2.1.5 IEEE-LOM

O IEEE LOM [13] é um modelo que busca especificar a sintaxe e a semântica dos metadados de objetos de aprendizagem, definindo os atributos necessários para descrever adequadamente os objetos. No LOM, os objetos de aprendizagem são definidos como entidades digitais ou não, que podem ser usados ou referenciados em cursos que utilizam recursos tecnológicos. Alguns exemplos desses tipos de curso são: ambientes de educação a distância, *Intelligent computer-aided instruction systems*, etc. Podemos citar alguns exemplos de objetos de aprendizagem: conteúdos multimídia, material didático digitalizado, avaliações, software de ensino e ferramentas de apoio a educação.

O conjunto de metadados proposto no LOM busca definir o mínimo de atributos necessários para permitir que os objetos de aprendizagem sejam gerenciados, avaliados e localizados.

Os elementos de dados que descrevem um objeto de aprendizagem são agrupados em 9 (nove) categorias, conforme definido no LOM v1.0 Base Schema (cláusula 6): [13]

Categoria	Descrição
<i>General</i>	Agrupar informações que descrevem características genéricas do objeto de aprendizagem como um todo.
<i>Lifecycle</i>	Agrupar informações relacionadas com a história e a situação atual do objeto. Além disso, registra todas as informações relevantes da evolução do objeto.
<i>Meta-Metadata</i>	Agrupar informações relacionadas com os próprios metadados dos objetos de aprendizagem.
<i>Technical</i>	Agrupar informações relacionadas com requisitos e características técnicas do objeto de aprendizagem.
<i>Educational</i>	Agrupar informações relacionadas com os aspectos educacionais e pedagógicos do objeto de aprendizagem.
<i>Rights</i>	Agrupar informações relacionadas com direitos de propriedade intelectual e condições de uso do objeto de aprendizagem.
<i>Relation</i>	Agrupar informações de relacionamentos semânticos entre o objeto de aprendizagem e outros objetos que estejam relacionados.
<i>Annotation</i>	Agrupar comentários relacionados com o uso educacional do objeto de aprendizagem e fornece informações dos autores e datas dos comentários.
<i>Classification</i>	Descreve o objeto de aprendizagem relacionando-o com um sistema de classificação pré-definido.

Tabela 2.4: Categorias do IEEE LOM

Árvore completa do IEEE-LOM

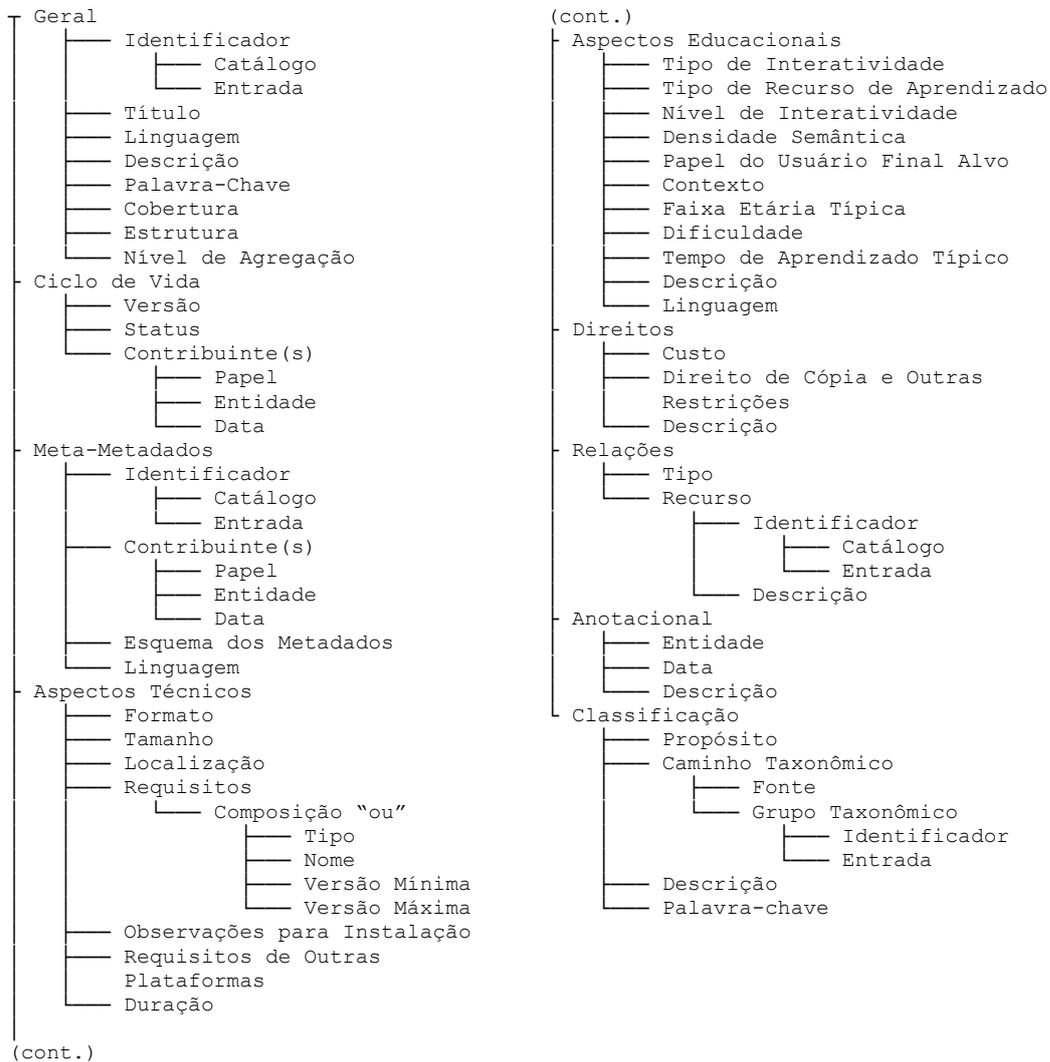


Figura 2.1: Árvore completa do IEEE-LOM.

2.2 Padrões Web

2.2.1 Serviços Web

Serviço *Web* é a mais recente evolução nos padrões de desenvolvimento de aplicações distribuídas permitindo que aplicações cooperem facilmente e compartilhem informações e dados umas com as outras [14].

Sobre a perspectiva técnica, um Serviço *Web* é nada mais do que uma coleção de uma ou mais operações que são acessíveis na rede através de uma descrição. Neste nível, o conceito Serviço *Web* não é novo. A indústria de Tecnologia de Informação (TI) vem discutindo o desafio da computação

distribuída – localizando e acessando sistemas remotos. A grande diferença é que agora a indústria está abordando este problema usando tecnologia aberta (XML e protocolos de internet) e padrões abertos gerenciados por consórcios tais como W3C (*World Wide Web Consortium*) que gerencia a evolução das especificações SOAP e WSDL.

O W3C define um Serviço *Web* como uma aplicação de software ou componente que é identificado por uma URI (*Uniform Resource Identifier*), cujas interfaces e conexões são capazes de serem definidas, descritas e descobertas em XML e apoia interações diretamente com outras aplicações de software usando mensagens codificadas em XML, via protocolos baseados na Internet. Fornecem interoperabilidade entre componentes de software, pois são baseados em mecanismos e protocolos padrão. Desta forma, Serviços *Web* são independentes de linguagem de implementação (Java, VB.Net, C#, J#, C++, JScript e Perl, entre tantas outras), modelo de objeto (EJB, COM, etc.) e plataforma (J2EE, .Net, etc.).

A utilização de Serviços *Web* na arquitetura se deu devido, como definido pelo W3C, a serem aplicações de software, cujas interfaces baseadas em XML, definem uma coleção de operações acessíveis na rede, independente da tecnologia usada para implementá-las [15].

2.2.1.1 SOAP

O SOAP (Simple Object Access Protocol) [16] é um protocolo baseado em XML para a troca de informações entre aplicações e Serviços *Web* em ambientes distribuídos, independentes de plataforma e linguagem de programação. Este protocolo encapsula as chamadas e retornos aos métodos dos Serviços web, sendo utilizado, principalmente, sobre HTTP.

2.2.1.2 WSDL

A WSDL (Web Services Description Language) [17] é baseada em XML, trafega sobre HTTP e fornece a descrição do Serviço *Web*, possibilitando a integração dele com ferramentas de desenvolvimento e de gerenciamento de componentes. Na WSDL de um Serviço *Web* estão descritos a URL para acessar o serviço, o nome dele, a descrição de cada método e a forma de fazer a requisição utilizando SOAP, GET ou POST.

2.2.2 XML

XML (Extensible Markup Language) [18] é uma linguagem de marcação definida pelo XML Working Group do W3C. XML é semelhante à Hypertext Markup Language (HTML) na medida em que XML é uma linguagem baseada em marcações, especificamente concebida para fornecer informações na Web. Porém, XML é diferente de HTML na medida em que as marcações utilizadas pela linguagem XML não são predefinidas. Em vez disso, a recomendação sobre XML do W3C especifica um conjunto de regras que têm de ser cumpridas de modo que se possa criar um conjunto de marcações funcionais próprias.

Podemos criar nossas próprias marcações para utilizar num documento XML, seguindo algumas regras simples:

- Um documento XML só pode conter um elemento raiz. O elemento raiz de um documento XML é um elemento único, que contém todo o conteúdo que faz parte do documento propriamente dito. É também conhecido como o elemento de documento.
- Todos os elementos XML precisam conter marcações de fim. Apesar das marcações de fim serem opcionais para determinados elementos de documentos HTML, todos os elementos de um documento XML precisam possuí-la.
- Os nomes das marcações de início e fim do elemento precisam ser idênticos. O XML é sensível a maiúsculas e minúsculas, logo o nome de uma marcação de fim terá de corresponder exatamente ao da marcação de início que a acompanha.
- Os nomes dos elementos XML não podem ficar sobrepostos. Se a marcação de início de um elemento aparecer em outro elemento, terá de terminar no mesmo elemento que a contém.
- Todos os valores de atributos têm de utilizar aspas. Os valores de atributos têm de estar entre aspas duplas ou aspas simples.
- Não é possível utilizar os seguintes caracteres no texto de um documento XML: '<', '>', '&'. Estes são caracteres especiais com um significado específico para analisadores de XML. Se precisar utilizar estes caracteres no texto do documento XML, deverá utilizar as referências de entidades ou caracteres predefinidos.

O cumprimento destas regras assegura que o documento XML foi bem formado, o que significa que adere à sintaxe XML tal como definida pela recomendação do W3C. Considera-se que os documentos XML têm XML válido se utilizarem um esquema XML para restringir o tipo de dados que pode ser utilizado no documento XML.

A utilização de XML na arquitetura se deu devido ao fato dos padrões utilizados serem definidos em XML, de ser utilizado pelos Serviços *Web* e de ser independente da linguagem de programação utilizada, permitindo assim que outros sistemas se comuniquem com o AccesForAll-LORIS de forma independente de linguagem.

2.2.3 XSLT

XSLT (Extensible Stylesheet Language Transformation) [19] é uma linguagem declarativa, baseada em XML, utilizada para apresentar ou transformar dados XML. A transformação de dados XML implica em utilizar o documento de origem, considerado como uma árvore de nós, e criar um segundo documento a partir dele em outro formato. A folha de estilos XSLT fornece as regras e o formato que especificam a forma como o documento de saída é criado. Se a saída tiver um formato adequado, tal como HTML, poderá ser utilizada para apresentação do documento de origem.

XPath [20] é uma linguagem utilizada para processar determinadas partes de um documento XML. A linguagem XPath também fornece funcionalidades básicas de manipulação de cadeias, números e valores booleanos. Utilizando expressões XPath, o XSLT faz componentes (ou nós) de um documento XML corresponder a modelos XSLT e, em seguida, aplica as regras de transformação e formatação aos componentes com correspondência. Os modelos XSLT são estruturas que especificam os componentes do documento XML que serão transformados, assim como o formato final nos quais os componentes com correspondência deverão ser transformados. Dado que o XSLT é representado em XML, o XSLT define um conjunto especial de elementos e atributos que podem ser utilizados para criar as transformações. O XSLT pode ser utilizado para transformar qualquer tipo de documento XML, incluindo dados e esquemas XML ou outras folhas de estilos XSLT.

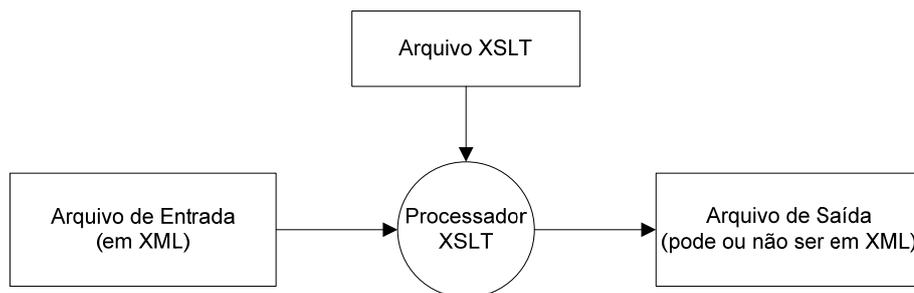


Figura 2.2: Funcionamento do processo de transformação XSLT.

Utilizou-se XSLT, por ser a linguagem ideal para transformar um ou mais XMLs em um outro documento, por ser amplamente utilizada, ter um excelente desempenho e sobretudo ser interpretada e não compilada, o que permite que alterações pontuais sejam feitas sem a necessidade de recompilar o programa.

2.2.4 XSD

O XSD (*XML Schema Definition*) [21] é uma recomendação do W3C que especifica como descrever formalmente os elementos de um documento XML. Essa descrição pode ser utilizada para verificar se cada item do conteúdo de um documento se adequou à descrição do elemento onde o conteúdo deveria estar.

Em linhas gerais, um esquema é uma representação abstrata das características de um objeto e de suas relações com outros objetos. Um esquema XML representa os interrelacionamentos entre os atributos e os elementos de um XML (por exemplo, um documento ou parte de um documento). Para criar um esquema para um documento é necessário analisar sua estrutura e definir cada elemento estrutural conforme encontrado.

O XSD possui diversas vantagens em relação a outras linguagens de definição de esquemas XML mais antigas, tais como DTD (*document type definition*) ou SOX (*Simple Object XML*). Diferentemente das linguagens mais antigas, o XSD é escrito em XML, o que significa que ele não necessita de um processamento intermediário.

2.3 Classificação Facetada

A classificação facetada permite que um mesmo objeto possua diversas classificações, permitindo assim a busca e a navegação por classes que formam grupos de objetos.

Segundo [22], a classificação facetada afetou o desenvolvimento dos sistemas mais conservadores na área da descrição e organização da documentação tradicional. É uma ferramenta de navegação popular para sítios de todos os tipos, ajudando a estruturar todos os tipos de objetos e informações sobre eles. Está começando a ser utilizada por pesquisadores nos campos da indexação automática e *web* semântica como uma ferramenta conceitual para auxiliar o entendimento das mais complexas relações entre objetos.

2.3.1 Facetas

Ranganathan foi o primeiro a introduzir ao mundo a palavra “faceta” na ciência da computação e na biblioteconomia e também foi o primeiro a desenvolver a teoria de análise de facetas. [23]

Originalmente facetas foram definidas como aspectos, propriedades ou características de uma classe ou tópico específico, definidas claramente e que deviam ser mutuamente exclusivas e exaustivas.

Diversos autores fizeram menção a facetas, classificação facetada e análise de facetas com diferentes conotações. Segundo [23], análise de facetas é definida como a organização de termos de um domínio de conhecimento em facetas homogêneas, mutuamente exclusivas.

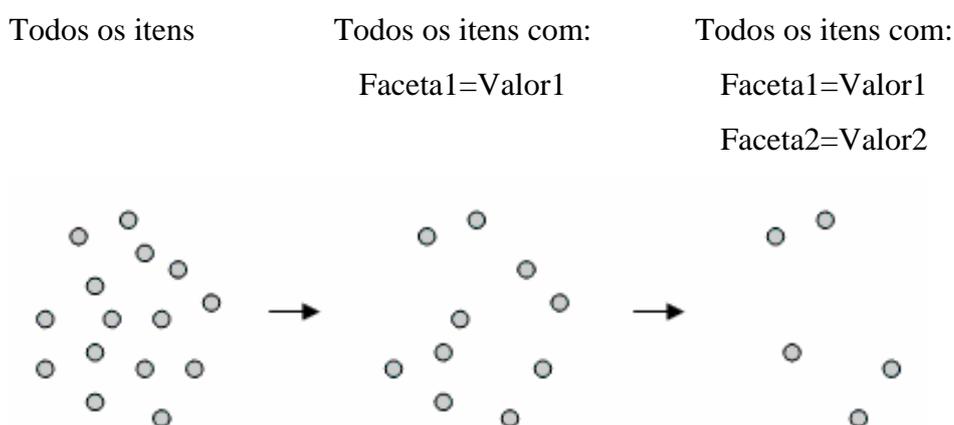
Cada faceta é derivada de um “ancestral” através de uma única característica de divisão. Cada categoria lógica deve ser isolada e cada nova característica de divisão deve ser indicada claramente. Portanto, uma faceta consiste em um grupo de termos que representa uma e somente uma característica de divisão de um tópico pertencente a um domínio.

Resumindo, a idéia é dividir um tópico em facetas de alto nível ou categorias fundamentais. Cada faceta de alto nível é subdividida em componentes, as subfacetas, que representam diferentes características.

2.3.2 Navegação Facetada

A navegação facetada permite que usuários naveguem pela informação selecionando progressivamente as facetas dos itens de informação. [24]

Inicialmente o usuário escolhe uma das facetas dentre o conjunto de facetas disponível. Posteriormente o usuário poderá escolher, sucessivamente, outras facetas dentre o conjunto de facetas disponível após a filtragem executada pelas facetas por ele escolhidas anteriormente. O conjunto de facetas que o usuário terá disponível para navegar é dependente do caminho por ele escolhido.



Optou-se pela utilização de navegação faceta para a exibição dos materiais de ensino retornados pela consulta executada pelo usuário, pois com ela o usuário pode aplicar múltiplos filtros aos resultados da consulta sem a necessidade de submeter uma nova consulta, diminuindo-se assim o tempo necessário para que o usuário localize aquilo que está procurando.

2.4 Tratamento de Versões de Objetos

Versionamento de objetos é o conceito que um único objeto pode ser representado por múltiplas versões (ex: instâncias de um objeto) de uma única vez. Duas diferentes formas de versionamento serão explicadas abaixo: [25]

- Versionamento linear é o conceito de armazenar versões anteriores dos objetos conforme eles são modificados. No versionamento linear, somente uma nova versão pode ser criada de cada versão existente de um objeto.

- Versionamento ramificado é um conceito de versionamento que suporta políticas concorrentes onde múltiplos usuários podem atualizar os mesmos dados concorrentemente. O trabalho de cada usuário é baseado em uma versão base consistente. Cada usuário pode modificar sua versão de um objeto. Em algum ponto do futuro, sob o suporte de uma aplicação ou de um usuário, os múltiplos ramos de uma versão são unificados para formar uma única versão do objeto. No versionamento ramificado podem ser criadas múltiplas versões novas de um objeto.

O versionamento de materiais de ensino é previsto no IEEE LOM através da categoria *Relation*. No campo “recurso” encontra-se o identificador para o outro material de ensino referenciado pelo atual e o no campo “tipo” define-se qual é a relação entre o material de ensino atual e o que está sendo referenciado no campo “recurso”. Esse campo pode ser preenchido com os seguintes valores:

Tipo	Descrição
<i>isversionof</i> e <i>hasversion</i>	O material de ensino descrito é uma versão, edição ou adaptação do material de ensino referenciado. Modificações na versão implicam em mudanças substantivas de conteúdo e não de formato.
<i>ispartof</i> e <i>haspart</i>	O material de ensino descrito é uma parte física ou lógica do material de ensino referenciado.
<i>isformatof</i> e <i>hasformat</i>	O material de ensino descrito possui o mesmo conteúdo que o material de ensino referenciado, porém apresentado em outro formato.
<i>references</i> e <i>isreferencedby</i>	O material de ensino descrito é referenciado, citado ou apontado de outra forma pelo material de ensino referenciado.
<i>isbasion</i> e <i>isbasisfor</i>	O material de ensino descrito é base do material de ensino referenciado.
<i>requires</i> e <i>isrequiredby</i>	O material de ensino descrito é requerido pelo material de ensino referenciado física ou logicamente.

Tabela 2.5: Possíveis valores para o campo “tipo” da categoria *Relation* do IEEE LOM.

No contexto desta dissertação a questão do versionamento torna-se mais complexa, pois deve ser capaz de lidar com materiais de ensino heterogêneos e armazenados em diversos repositórios que não necessariamente conhecem um ao outro, portanto, os relacionamentos entre seus materiais de ensino não está definido.

Surge então uma necessidade de considerar uma semântica de identidade e similaridade de materiais de ensino ao detectar versões.