

2 Tecnologias de Banco de Dados para e-Learning

O objetivo deste capítulo é apresentar definições usuais ao contexto de sistemas de banco de dados e ao de aprendizado eletrônico, para o seu melhor entendimento quando aplicadas neste trabalho.

Inicialmente, tendo em vista a autonomia prevista para os sítios do PGL, são apresentadas noções sobre sistemas de gerência de bancos de dados heterogêneos.

A complexidade para implementar algumas tarefas é crescente quando saímos de um ambiente centralizado e passamos para um ambiente composto por diversos bancos de dados autônomos, onde inclusive os gerenciadores podem ser diferentes. As seguintes tarefas são exemplos onde tal fato acontece: execução de consultas, construção de esquemas, gestão de transações, otimização de comandos. Sua complexidade é apresentada ao serem descritas as funcionalidades disponíveis em gerenciadores de bancos de dados federados.

Datando de 1992 o conceito de objetos de aprendizado aparece hoje com diferentes versões. Vários trabalhos estão sendo conduzidos visando a melhoria da interoperabilidade de recursos produzidos por empresas e universidades. Este assunto complementa o capítulo que termina com breve resumo sobre sistemas para área de aprendizado.

2.1. Sistemas de Gerência de Banco de Dados Heterogêneos

Um SGBDH - Sistema de Gerência de Bancos de Dados Heterogêneos visa prover acesso a dados distribuídos por diversos sistemas distintos de banco de dados.

Os sistemas membros ou componentes podem ser centralizados, distribuídos e também heterogêneos. Em relação aos sistemas componentes, o sistema de bancos de dados heterogêneos pode ser analisado segundo três dimensões ortogonais: (1) distribuição dos sistemas locais, (2) sua heterogeneidade, e (3) a

autonomia dos sistemas locais [SL90, OV99]. Estas três dimensões representadas na **Figura 2** são apresentadas a seguir:

1. **Distribuição** – Os dados podem estar distribuídos por vários bancos de dados, armazenados em diferentes sistemas computacionais distribuídos geograficamente e interligados por um sistema de comunicação de dados. Em termos de representação podem ser considerados os seguintes valores: 0 – sistemas não distribuídos; 1 – sistemas cliente/servidor e 2 – comunicação ponto a ponto (*peer to peer*, em inglês);
2. **Heterogeneidade** – A **Tabela 1** apresenta exemplos que podem ser considerados para caracterizar heterogeneidades de sistemas. Na **Figura 2** considerou-se: 0 – sistemas homogêneos e 1 – sistemas heterogêneos;

Sistemas de Gerência de Banco de Dados	
Diferenças nos SGBD	
<ul style="list-style-type: none"> - Modelo de dados (estruturas, restrições, linguagens de definição, de manipulação ou de consulta) - Suporte ao nível de sistema (controle de concorrência, efetivação (<i>commit</i>), recuperação) - Heterogeneidade semântica (ocorrência de significados diferentes, distintas interpretações ou diferentes usos associados aos dados armazenados) 	
Sistemas Operacionais	Comunicação
Diferenças nos sistemas operacionais	
<ul style="list-style-type: none"> - Sistema de arquivos - Nomenclatura, tipos de arquivos, operações; - Suporte para transações - Comunicação entre processos 	
Equipamentos	
Diferenças nos equipamentos	
<ul style="list-style-type: none"> - Conjunto de instruções - Formato de dados e representação - Configuração 	

Tabela 1. – Tipos de Heterogeneidades [SL90]

3. **Autonomia** – Este item aborda o grau de independência com que os sistemas componentes podem realizar operações locais considerando sua participação no conjunto. Refere-se à distribuição do controle e existem diversas classificações como, por exemplo:

- a. **Autônomos fortemente integrados:** são sistemas caracterizados por disponibilizar uma imagem global do banco para qualquer usuário, como se os dados estivessem logicamente centralizados, mesmo que múltiplos bancos de dados estejam envolvidos;
- b. **Semi-autônomos:** consistem de sistemas de gerência de banco de dados que podem operar de forma independente e que participam de uma federação para a qual disponibilizam seus dados. Não possuem total autonomia porque precisam ser modificados de forma a permitir que seus dados sejam compartilhados por outros sistemas de gerência de banco de dados;

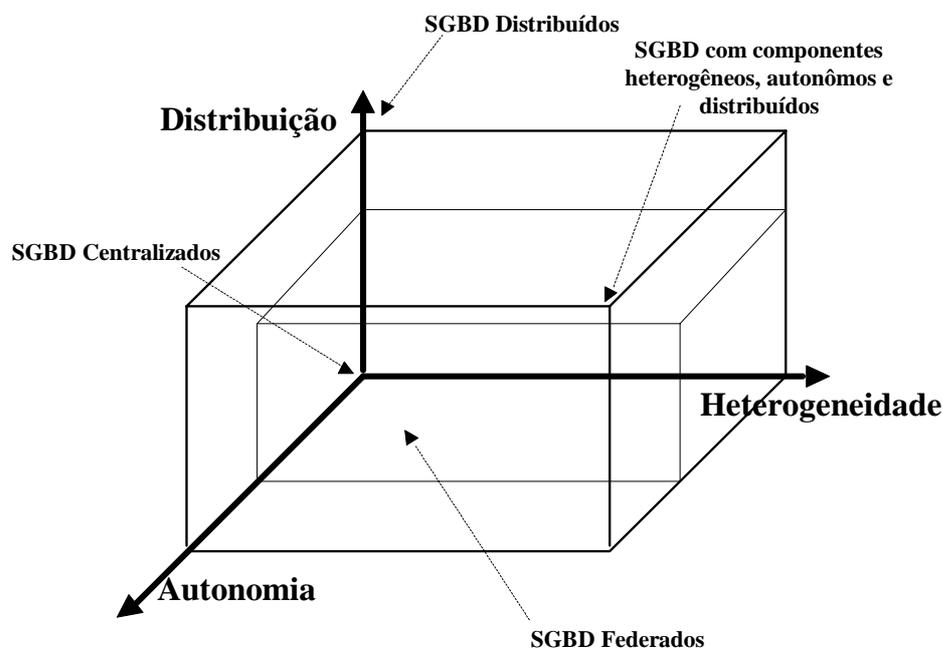


Figura 2. Dimensões de um SGBDH.

- c. **Totalmente isolados:** neste caso os sistemas de gerência individuais são auto-suficientes e não tomam conhecimento da existência de outros sistemas de gerência de banco de dados e nem se comunicam com eles.

Na **Figura 2** a autonomia foi representada considerando os seguintes valores: 0 – fortemente integrados; 1 – semi-autônomos e 2 – totalmente isolados;

Na **Figura 2** está esquematizado o espaço que representa os sistemas gerenciadores de bancos de dados federados: (A1, D0, H0), (A1, D0, H1), (A1, D1, H1), conforme a existência de distribuição e de heterogeneidade [OV99]. Esta arquitetura de sistemas será mais detalhada nos próximos itens dada a sua relevância nesta dissertação.

2.2.

Sistemas de Gerência de Bancos de Dados Federados

Um sistema de gerência de bancos de dados federado (SGBDF) é uma coleção de sistemas de banco de dados cooperativos que são autônomos e normalmente heterogêneos [SL90].

Discutidos de forma sistêmica inicialmente em 1990, uma década após, em 2001, ainda perduravam problemas e soluções eram pesquisadas frente a novas e mais complexas necessidades reveladas pelos crescentes volumes de dados armazenados e pela diversidade dos tipos de dados. Novas condições de integração, mais amplas e com suporte a maiores heterogeneidades e requisitos mais estreitos foram relatadas por Stonebraker e Hellerstein [SH01]. Algumas dificuldades persistem ainda hoje como, por exemplo, desenvolver a integração à uma base de dados de um sistema ERP (*Enterprise Resource Planning*). Criados com o propósito de reunir todos os sistemas legados das empresas, diversos fabricantes desenvolveram versões proprietárias de sistemas ERP. A dimensão do problema aparece quando constatamos que as bases dados de tais sistemas são compostas por milhares de relações. A título de exemplo, o SAP R/3, um sistema ERP bastante difundido, na sua versão 4.5 contém 13.365 tabelas.

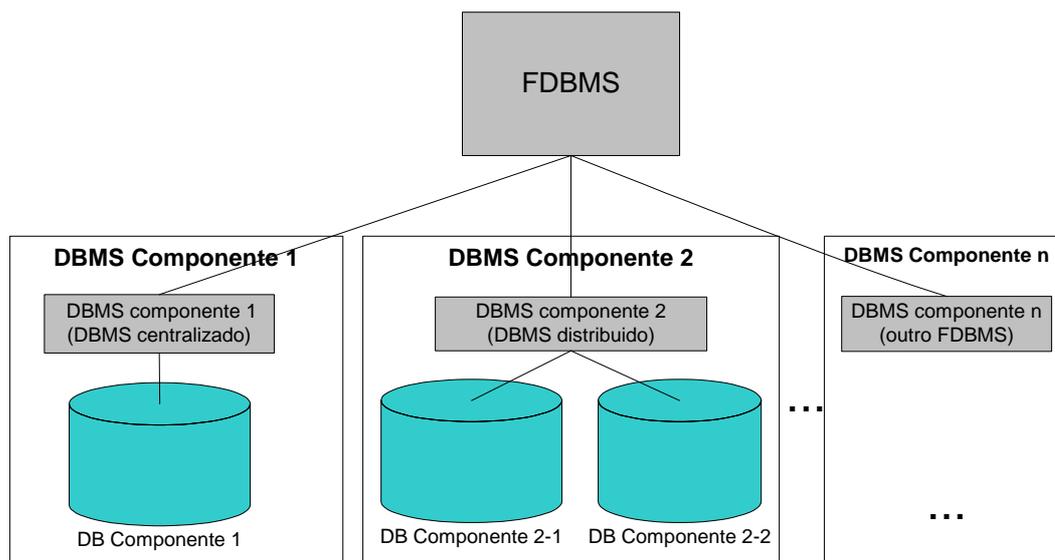


Figura 3. Aspectos gerais de um SGBDF [SL90].

As nomenclaturas *sistema de bancos de dados federados* e *arquitetura de bancos de dados federados* foram introduzidas inicialmente por Heimbigner e McLeod [1985] para significar uma “coleção de componentes unidos em uma federação com acoplamento fraco com o objetivo de compartilhar e trocar informações” e “um modelo organizacional baseado em bancos de dados iguais e autônomos com o compartilhamento controlado por interfaces explícitas”.

Bancos de dados federados podem ser entendidos como redes que integram um número de sistemas pré-existentes de bancos de dados autônomos homogêneos ou heterogêneos. Em relação aos sistemas componentes os sistemas de gerência de bancos de dados federados podem utilizar, diferentes modelos de dados, diferentes definições, diferentes manipulações, ter gerenciamento de transações e mecanismos de controle de concorrência próprios.

Os aspectos quanto à distribuição, autonomia e heterogeneidade são apresentados esquematicamente na **Figura 3**.

Nos próximos itens será apresentado um breve resumo sobre como particularidades devidas à distribuição dos dados são resolvidas.

2.2.1. Arquitetura em 5-níveis ANSI/SPARC

Nos sistemas tradicionais de banco de dados, ou seja, nos sistemas centralizados, para atender às diferentes necessidades de dados dos usuários é

utilizada a arquitetura em 3-níveis definida por ANSI/SPARC, conforme **Figura 4**.

O objetivo desta arquitetura é oferecer através dos esquemas externos uma visão dos dados que seja independente do seu tipo de armazenamento, definido no esquema interno, voltado para as necessidades dos administradores de banco de dados. Um esquema externo mapeia as necessidades específicas de usuários e aplicações e é obtido por transformações aplicadas ao esquema conceitual. O esquema conceitual representa uma visão abstrata dos dados que é independente de qualquer aplicação/usuário e da forma de armazenamento dos dados.

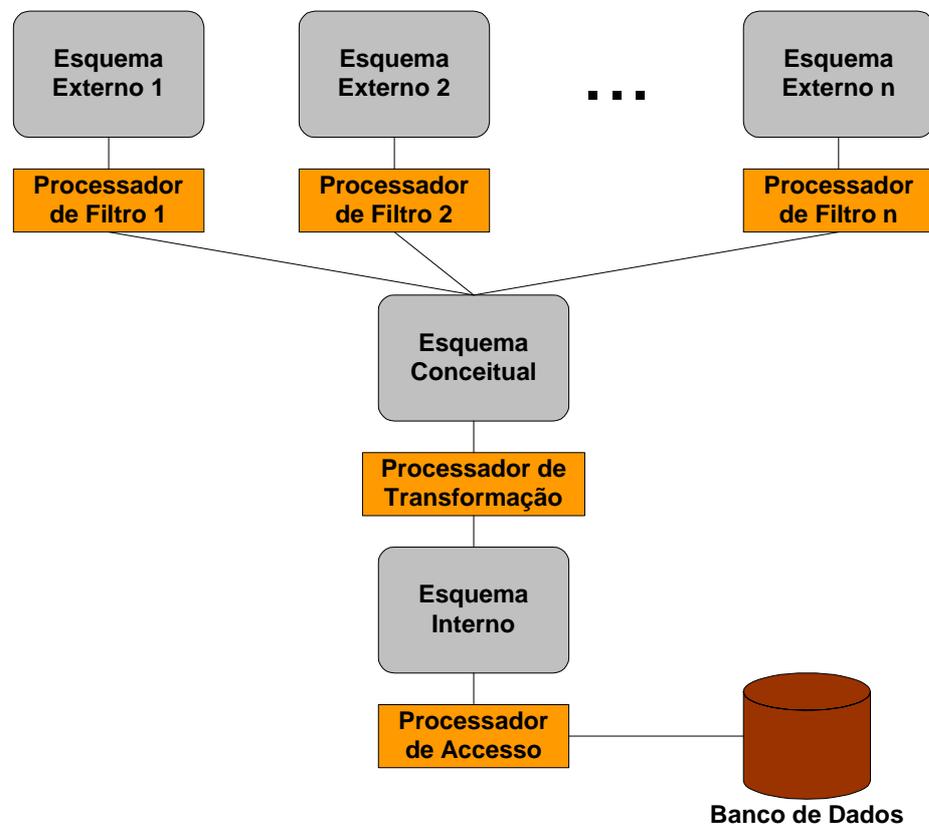


Figura 4. Arquitetura em 3-níveis de um SGBD centralizado.

Em ambientes com múltiplos bancos de dados, por exemplo, quando é considerada uma federação de bancos de dados, novas camadas são necessárias para propiciar as mesmas condições de transparência para o usuário situado no nível mais externo, ou seja, o usuário da federação. Neste caso, a arquitetura é expressa em 5-níveis, conforme **Figura 5**, onde se destacam:

- **ESQUEMA LOCAL:** correspondente ao esquema conceitual do banco de dados componente, expresso no modelo de dados nativo do sistema

de gerência de banco de dados que suporta este componente. Em uma federação é possível que existam esquemas locais expressos em modelos distintos, por exemplo, modelo relacional, modelo orientado a objeto, etc;

- **ESQUEMA COMPONENTE:** visa superar as heterogeneidades entre os diferentes modelos de dados dos sistemas de gerência dos bancos de dados componentes. Este esquema é implementado através da tradução dos esquemas locais (conceituais) para o modelo de dados comum à federação, modelo este também chamado **modelo de dados canônico** (MDC). Dentre as principais razões para redefinir esquemas componentes em modelo de dados canônico temos: (1) utilizar uma representação única para descrever esquemas distintos; (2) complementar as semânticas que não foram bem definidas nos esquemas locais;
- **ESQUEMA EXPORTAÇÃO:** tem por objetivo limitar o conjunto de dados do esquema componente que será disponibilizado para a federação e seus usuários. Um processador de filtro pode atuar para controlar o acesso e as operações permitidas;

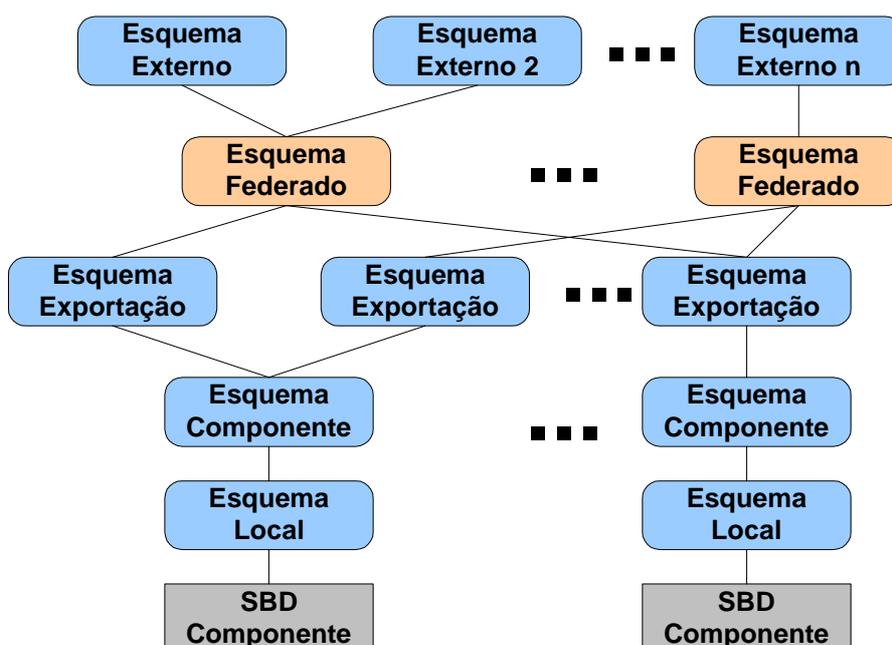


Figura 5. Arquitetura em 5-níveis de um SGBDF [SL90].

- **ESQUEMA FEDERADO:** corresponde a integração dos múltiplos esquemas exportação, bem como, disponibiliza as informações sobre a distribuição dos dados. Conceitos similares a esquema federado são expressos pelos termos **esquema importação** (*import schema*) [Heimbigner e McLeod, 1985], **esquema global** (*global schema*) [Landers e Rosemberg, 1982] e **esquema conceitual global** (*global conceptual schema*) [Litwin et al. 1982], **esquema unificado** (*unified schema*) e **esquema empresa** (*enterprise schema*), embora os termos diferentes de **esquema importação** sejam normalmente utilizados quando existir apenas um único esquema deste tipo no sistema [SL90];
- **ESQUEMA EXTERNO:** tem por objetivo prover informações para um usuário ou aplicação ou para classes de usuários ou de aplicações. A razão para sua utilização é minimizar a complexidade dos esquemas federados, que normalmente são muito grandes, e adicionar restrições e controles de acesso específicos.

A construção destes esquemas está relacionada com uma característica importante de bancos de dados federados, o tipo de acoplamento que existe entre os componentes e a federação.

2.2.2. Acoplamento

O relacionamento que existe entre uma federação e os bancos de dados componentes é denominado acoplamento. Este acoplamento pode ser forte (*tightly coupled*) ou fraco (*loosely coupled*) e é definido em função da gestão da federação e do processo de integração de novos componentes.

No primeiro caso, federação com acoplamento forte, as decisões são tomadas no nível da federação. Os bancos de dados componentes são integrados através de um ou mais esquemas globais. São responsabilidades da federação e de seu administrador a criação, gestão da federação e o controle de acesso aos bancos de dados componentes.

No segundo caso, considerando uma federação com acoplamento fraco, não existe um esquema global ou federado. São responsabilidades particulares de cada usuário a criação e a manutenção de um “esquema global”. Este esquema global corresponde a sua visão sobre o conjunto dos dados, através da sua interpretação da semântica destes dados. Outros termos utilizados para denominar sistemas federados com acoplamento fraco são: **sistemas de bancos de dados interoperáveis** (*interoperable database system*) e **sistemas de bancos de dados múltiplos** (*multidatabase system*) [SL90].

Na **Tabela 2** é apresentado um resumo do comportamento de algumas funcionalidades de sistemas de bancos de dados federados em função do tipo de acoplamento.

	Acoplamento Forte	Acoplamento Fraco
Gerência da federação	Administrador federativo	Usuário federação
Integração de componentes	Administrador federativo	Usuário federação
Atualizações	sim	Normalmente apenas leitura
Transações	Modelo convencional	Algum relaxamento

Tabela 2. - Características de acoplamento.

Nos próximos itens serão abordados maiores detalhes sobre estas características.

2.2.2.1. Acoplamento Forte

Em uma federação com acoplamento forte cabe ao administrador de banco de dados no nível da federação decidir como será incluído cada novo banco de dados componente. Isto é realizado através da negociação com os administradores dos bancos de dados componentes. Nesta negociação são definidos quais os dados que serão disponibilizados para o nível da federação e são também resolvidas questões semânticas.

Observe-se que apesar de dispor de um administrador de banco de dados no nível da federação, cada administrador de um banco de dados componente permanece com certo grau de autonomia. Competem a ele as decisões sobre quais

dados disponibilizar para a federação e como atender as requisições da federação em relação às requisições locais.

A elaboração dos esquemas federados e dos esquemas externos é de responsabilidade do administrador federativo.

Quanto ao suporte para o usuário federativo, o acoplamento forte permite que seja normalmente disponibilizada uma linguagem especializada que permita a formulação de requisições sem a necessidade do domínio as linguagens específicas dos sistemas gerenciadores dos bancos de dados componentes.

2.2.2.2. Acoplamento Fraco

Quando o acoplamento é fraco, cada usuário torna-se o administrador de banco de dados de seu esquema federado, e por consequência tem as seguintes atribuições e necessidades:

1. Analisar os esquemas exportação dos bancos de dados componentes e decidir sobre como compor o “esquema federado”;
2. Interpretar a semântica dos bancos de dados componentes e resolver as heterogeneidades;
3. Dispor de plenos poderes para alterar o esquema federado independente de outros usuários.

Uma federação cujo acoplamento fraco é normalmente utilizada para leitura de dados. A integridade do sistema pode ser prejudicada durante modificações de dados, uma vez que um usuário pode atualizar um conjunto de dados, sobre os quais sua visão particular não seja igual à dos demais usuários.

A **Tabela 3** apresenta um exemplo de múltipla semântica em um ambiente com acoplamento fraco.

S1.Sapato {marrom, canela, creme, branco, preto}
S2.Sapato {marrom, canela, branco, preto}
Usuário1: federação.Creme = {creme.S1 e canela.S2}
Usuário2: federação.Creme = {[canela,creme].S1 e [canela, branco].S2}

Tabela 3. - Exemplo de múltipla semântica.

Neste exemplo são apresentadas as definições semânticas para as cores de sapatos definidas em dois locais de processamento de dados, e a sua interpretação por dois usuários. Como os conjuntos de valores válidos são distintos, cada usuário no exemplo faz a sua combinação para interpretar o que irá considerar como cor *creme*.

A existência de usuários especializados é uma necessidade normalmente presente em sistemas de acoplamento fraco. Estes usuários devem ter a capacidade para formular requisições utilizando linguagens específicas dos sistemas de gerência dos bancos de dados componentes, bem como, ter domínio sobre a distribuição dos dados. Desta forma as suas requisições são endereçadas diretamente para os sítios específicos.

2.2.3.

Linguagem de Definição de Dados e Linguagem de Comandos

Em sistemas de gerência de bancos de dados federados é usual que exista uma linguagem específica, **linguagem multibanco de dados** (*multidabase language*). Esta linguagem oferece normalmente funcionalidades que não estão disponíveis nos sistemas componentes. Como, por exemplo, a capacidade de definir esquemas federados como visões sobre múltiplos esquemas de exportação e a capacidade de formular comandos sobre estas visões. Estas facilidades são essenciais para prover uma visão global sobre os dados distribuídos por vários gerenciadores. Esta linguagem deve tratar dos conflitos oriundos de diferenças entre:

- Nomes de dados;
- Estruturas de dados;
- Tipos de dados;
- Unidades dos dados.

A falta de uma linguagem que permita elaborar esquemas sobre múltiplos bancos de dados implicará ou na necessidade de usuários especializados ou no desenvolvimento de ferramentas para esta finalidade.

2.2.4. Otimização de Consultas

Em uma federação de bancos de dados o trabalho de otimização dos comandos de consulta submetidos por usuários da federação pode ser extenso. Isto acontece porque um comando ao ser submetido no nível da federação será decomposto em comandos que serão executados remotamente pelos sistemas de gerência de banco de dados componentes. Neste aspecto, algumas particularidades devem ser consideradas:

- a) Para a otimização de consultas no nível da federação o cálculo do custo para executar remotamente uma operação está relacionado com a capacidade do sistema de gerência de banco de dados componente onde está operação será realizada. Em princípio, este custo é função dos seguintes fatores: autonomia do componente, sua capacidade de executar determinadas operações, sua carga de processamento e capacidade de seu elo de conexão;
- b) Localmente cada sistema de gerência de banco de dados componente pode ter características próprias de otimização;
- c) Pode haver diferenças entre as operações do sistema operacional e de banco de dados de cada componente e as equivalentes no nível da federação.

A complexidade que está envolvida na execução de consultas em um ambiente composto por múltiplos bancos de dados é normalmente resolvida com o apoio de um catálogo global com informações sobre a distribuição dos dados e sobre as características dos sistemas componentes.

2.2.5. Gerenciamento Global das Transações

É responsabilidade do gerenciador global de transações (GTM) manter o banco de dados federado consistente enquanto permite atualizações concorrentes em múltiplos bancos de dados.

Dois tipos de transações acontecem em uma federação: as transações globais e transações locais. As primeiras são submetidas por usuários da federação e as últimas são submetidas diretamente por usuários locais ao sistema de gerência do

banco de dados componente. O principal problema é consequência do fato de que a federação desconhece o que se passa no nível local, uma vez que este nível tem seu grau de autonomia. Além deste fato básico existem outros como, por exemplo, suporte a diferentes mecanismos de concorrência e de impasses (*deadlock*).

Na falta de recurso específico de um gerenciador de bancos de dados federado, uma alternativa normalmente utilizada para solucionar este problema é limitar a concorrência em favor da autonomia do componente, podendo ser implementada através de:

- Leituras não sincronizadas
- Restrição nas atualizações simultâneas em múltiplos componentes;
- Execução de atualizações locais apenas fora de linha.

2.2.6. Controle Semântico

A interpretação de esquemas e seu correto entendimento é uma tarefa constante na administração de bancos de dados. A complexidade desta tarefa cresce vertiginosamente à medida que saímos de um ambiente único centralizado e passamos a tratar uma federação, onde cada sistema componente tem o seu próprio administrador e sua própria linguagem de definição, além de outras características exclusivas.

Em um banco de dados federado, fortemente acoplado, uma interface uniforme de acesso aos dados é provida por meio do esquema federado, o qual é resultado da integração dos diferentes esquemas dos bancos de dados componentes.

Um gerenciador de banco de dados federado provê normalmente recursos que permitem o acesso aos esquemas remotos, facilitando em muito as tarefas de elaboração dos esquemas federados.

Este item sobre sistemas gerenciadores de bancos de dados federados teve por objetivo apresentar a complexidade para implementar funções como execução de consultas, construção de esquemas, gestão de transações, otimização de comandos, considerando um ambiente composto por múltiplos bancos de dados,

e, por consequência, a vantagem em utilizar gerenciadores cujas funcionalidades facilitem tais tarefas.

2.3. Objetos de Aprendizado

O termo objeto de aprendizagem (*learning object*) é creditado por muitos autores a Wayne Hodgins, que o teria utilizado pela primeira vez em 1992.

Ao preparar um novo curso, Wayne Hodgins percebeu que poderia reutilizar módulos antigos, adicionando-lhes interoperabilidade. Por este raciocínio, os módulos simples seriam agregados compondo elementos mais complexos. Este processo de composição é apresentado na **Figura 6** [Hod02a] e [Hod02b].

Nesta figura cinco níveis de agregação estão distribuídos em duas classes: a primeira reunindo conteúdos comuns aplicáveis a qualquer aplicação e a segunda na qual os conteúdos são agregados tendo em vista uma aplicação específica.

O material mais elementar composto por “**dados brutos**” e “**dados multimídia**” corresponde aos objetos sem nenhum contexto didático definido. São exemplos deste tipo de material: uma frase, um parágrafo, uma imagem, um vídeo, uma simulação, etc. Estes dados apresentam a maior possibilidade para reuso, por estarem livres de qualquer associação com objetivos de aprendizagem. Por exemplo, a imagem de uma borboleta: enquanto apenas uma imagem, ela pode ser utilizada em várias situações como, em uma aula de biologia (insetos), em uma aula de pintura (estudo de cores), em uma aula de matemática (números), em um texto sobre geografia, etc.

A reunião de dados brutos em conjuntos irá caracterizar o conceito “**Objeto de Informação Reutilizável**”. Estes extratos de informação reutilizável (OIR) irão formar sumários, sínteses, processos, conceitos ou procedimentos.

Os objetos de informação ao serem selecionados e reunidos em torno de um objetivo formam um “**Objeto de Aplicação**” ou um “**Objeto de Aprendizado**”. Um objeto de aprendizagem reutilizável (OAR) é formado pela união de vários objetos de informação reutilizáveis (OIR), selecionados para transmitir uma idéia associada a um simples objetivo de aprendizagem.

Ao reunir objetos de aprendizagem em torno de objetivos maiores serão formadas estruturas mais complexas dando origem à “**aulas**” e a “ **cursos**”. Estes

níveis de agregação crescentes ganham maiores definições de contexto e ao mesmo tempo perdem a sua capacidade de reuso. Ou seja, voltando para aquele exemplo inicialmente comentado da imagem de uma borboleta, quando esta imagem for inserida dentro de uma aula de geografia, descrevendo a floresta amazônica, esta aula de geografia não seria aplicável como uma aula de alfabetização descrevendo números.

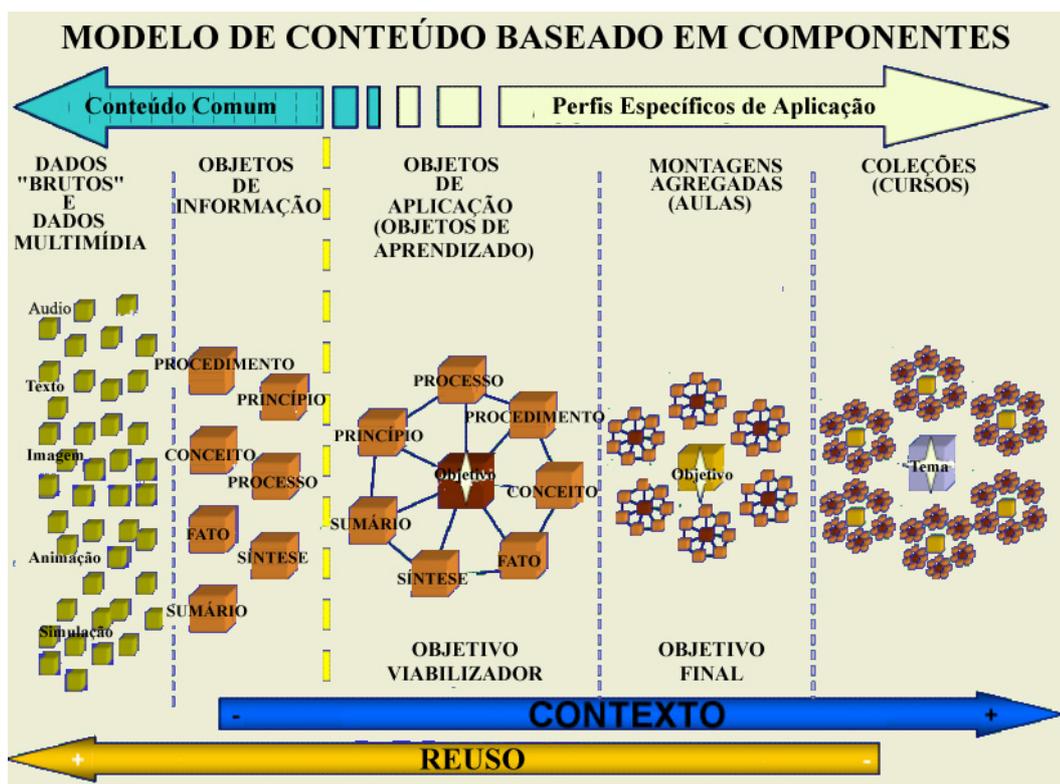


Figura 6. Modelo de composição de objetos [Hod02]

A elaboração de um conteúdo utilizando a modelagem por objeto propicia diversos benefícios, conforme foi destacado por Warren Longmire [Lon00]:

- **Valorização:** um conteúdo será mais valorizado a cada nova reutilização, ocorrendo em contrapartida uma diminuição nos custos para elaborar e produzir novos objetos de aprendizagem derivados da composição dos primeiros;
- **Flexibilidade:** um material, ao ser elaborado considerando a possibilidade para sua reutilização futura em diferentes contextos, será

mais facilmente reutilizável do que aquele, que foi feito para um contexto único, e que deverá ser reescrito a cada nova ambientação;

- **Gerenciamento de conteúdo:** o conteúdo dos objetos de aprendizado é descrito por metadados. As tarefas de organizar, identificar e localizar itens relevantes serão facilitadas recorrendo aos metadados;
- **Personalização:** a abordagem por objeto facilita adaptar o resultado segundo as necessidades de quem desenvolve através da seleção, montagem e reorganização dos componentes.

Para facilitar uma ampla aceitação do conceito de objeto de aprendizado em 1996 foi formado o IEEE/LTSC². Este comitê tem por objetivo desenvolver e promover padrões para tecnologia de aprendizado, de forma a assegurar a interoperabilidade dos recursos, especificamente objetos de aprendizado, de universidades, empresas e outras entidades.

Na mesma época, teve início um projeto semelhante chamado ARIADNE³.

Um terceiro grande projeto contemporâneo é o IMS⁴.

Cada uma destas iniciativas, assim como outras (por exemplo, ADL⁵) começaram a desenvolver padrões técnicos para suporte ao amplo uso de objetos de aprendizado.

O IEEE/LTSC apresentou a seguinte definição para objetos de aprendizado:

“Objetos de aprendizado são definidos como qualquer entidade, digitalizada ou não, que possa ser usada, reutilizada ou referenciada por tecnologia de suporte ao aprendizado. Exemplos de tecnologia de suporte ao aprendizado incluem sistemas de treinamento baseado em computadores, ambientes interativos de aprendizado, sistemas inteligentes de instrução por computador, sistemas de ensino a distância e ambientes colaborativos de aprendizado. Exemplos de objetos

² LTSC - Learning Technology Steering Committee of the IEEE - Institute of Electrical and Electronics Engineers [Lts96];

³ ARIADNE - Alliance of Remote Instructional Authoring and Distribution Networks for Europe [Ari97]

⁴ IMS - Instructional Management System Project, atualmente denominado IMS Global Learning Consortium, Inc [IMS01]

⁵ ADL - Advanced Distributed Learning [Adl03]

de aprendizado incluem conteúdos multimídia, conteúdos instrucionais, objetivos de aprendizado, programas e ferramentas instrucionais, pessoas e organizações ou eventos referenciados durante aprendizado suportado por tecnologia” [Lom00].

Esta definição foi considerada muito ampla e diversos grupos fora do IEEE/LTSC criaram novas definições, embora sempre utilizando o termo “objeto de aprendizado” [Wil02 e MHM01]:

- Cisco (2001) definiu objeto de aprendizado como “um conteúdo de informação granular, reutilizável e independente da mídia”, e utilizou termos como “objetos educacionais, objetos de conteúdo, componentes de treinamento” [MHM01];
- South e Monson (2001) definiram objetos de aprendizado como “mídias digitais elaboradas e utilizadas para fins de aprendizado, englobando mapas, vídeos e simulações interativas” [MHM01];
- Wiley (2001) definiu objetos de aprendizado como sendo “elementos de um novo conceito de instrução baseado em computadores e apoiado no paradigma da *orientação por objeto* da ciência computacional”. Estes elementos permitem que componentes pequenos em relação ao tamanho total de um curso sejam reutilizados inúmeras vezes em diferentes contextos. São normalmente vistos como entidades digitalizadas circulando pela Internet, o que significa que inúmeras pessoas podem acessá-los e utilizá-los simultaneamente [MHM01];
- Internetttime.com (2001) mencionou “pequenos conteúdos de aprendizado (grânulos ou objetos) que são rotulados de tal forma que sistemas informatizados possam tratá-los automaticamente compondo e disponibilizando experiências individualizadas de aprendizado”[Wil02];
- Asymetrix Inc (atualmente Click2learn) definiu objetos de aprendizado como sendo elementos pré-codificados visando simplificar a programação final [Wil02];
- Macromedia definiu objetos de aprendizado sendo conteúdos sonoros instrutivos, combinando possibilidades para prática, simulação, interação colaborativa e para avaliação, diretamente relacionadas ao objetivo do aprendizado [Wil02].

O tamanho dos módulos também foi objeto de debates. John Talanca [Tal02] analisando o material que encontrava disponível, em 2001, considerou os módulos excessivamente extensos para aprendizado eletrônico. Neste caso é esperado receber conhecimentos sobre um tópico específico normalmente em um intervalo de tempo reduzido. Na sua análise ele enumerou alguns exemplos:

Curso	Duração
Introdução ao Excel	2 h
Realizando uma Reunião com Sucesso	90 min
O Sistema Cardiovascular	2 h
Princípios da Negociação	60 min

Sua conclusão foi que tais módulos eram excessivamente longos para aprendizado eletrônico. A título de exemplo é citado interesse de alguém pelo assunto “Construir uma tabela virtual em Excel”. Pelos exemplos encontrados, esta pessoa deveria “assistir” ao treinamento básico, intermediário e, eventualmente, ao treinamento avançado sobre Excel para obter tal informação. Estaria totalizando algumas horas de dedicação ao aprendizado. Para contornar este problema sua conclusão foi que os “módulos” deveriam ter uma duração que não excedesse aos 10 a 12 minutos, e serem sobre assuntos bem específicos, abandonando a idéia tradicional de graduações de especialização, materializada pelos tradicionais níveis básico, intermediário e avançado.

Enquanto não existe uma definição consolidada sobre objetos de aprendizado, algumas caracterizações predominam:

- Objetos de aprendizado serão armazenados e recuperados utilizando metadados e rótulos;
- Objetos de aprendizado serão contextualizados e compostos utilizando meta arquivos. Estes meta arquivos, ao disponibilizar o significado e a aplicação dos objetos de aprendizado, facilitarão a sua composição em um novo conjunto com significado válido.

2.4. Metadados para Objetos de Aprendizado

Como vimos em item anterior existe um consenso de que a descrição das características dos objetos de aprendizagem seja feita por metadados. Porém, assim como existem diversas organizações buscando definir objetos de aprendizagem, existem também diversas linhas buscando definir padrões para os metadados.

Na **Figura 7** é apresentado o relacionamento entre os principais trabalhos existentes, selecionados principalmente pela sua inter-relação:

- LOM;
- IMS
- SCORM⁶
- ARIADNE
- DCMI⁷
- AICC⁸

As setas, nesta figura, indicam a existência de “referência ao conteúdo” ou “utilização de conteúdo”, mostrando o estreito relacionamento que existe entre os diferentes grupos.

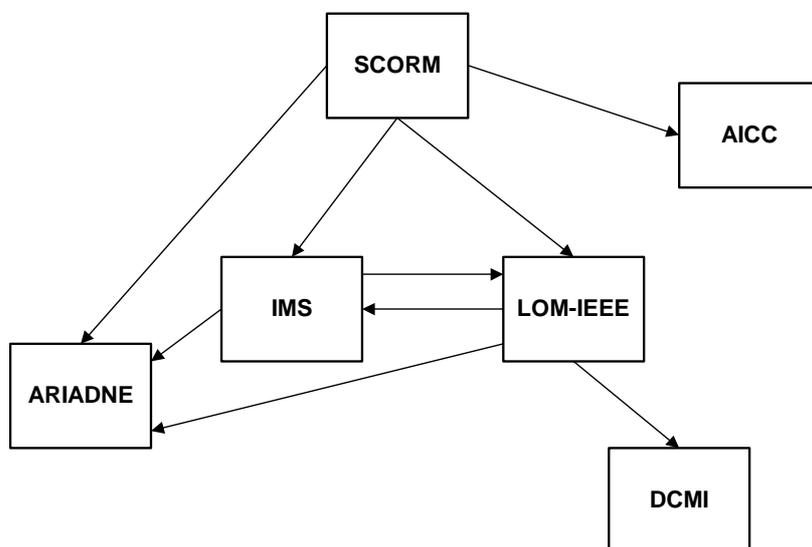


Figura 7. Relacionamento entre padrões de metadados [PPM03]

⁶ SCORM – Sharable Courseware Object Reference Model [Adl03]

⁷ DCMI – Dublin Core Metadata Initiative [DCM95]

⁸ AICC – Aviation Industry CBT Committee [AIC88]

A seguir serão apresentados alguns detalhes sobre LOM IEEE, IMS e SCORM.

2.4.1. LOM IEEE

Como vimos anteriormente o LTSC foi criado no IEEE com o objetivo de definir normas e padrões para a área de aprendizado.

Dentre os resultados produzidos, aquele voltado para descrição de metadados é denominado de Metadados para Objetos de Aprendizado (*LOM – Learning Object Metadata*).

Este trabalho ainda encontra-se em desenvolvimento, sendo que o documento consultado (P1484) e utilizado ao longo desta dissertação foi o correspondente à versão 12.1.2002 [Iee02].

Este padrão visa facilitar a utilização de objetos de aprendizado por processos automatizados e por pessoas, buscando também:

- Facilitar o compartilhamento e intercâmbio de objetos de aprendizado, através da formação de catálogos;
- Permitir ligar (*binding*) objetos de aprendizado através dos metadados;
- Permitir ordenar objetos de aprendizado por agentes de *software*.

A estrutura do LOM é definida em forma hierárquica com vários níveis de detalhamento.

Para cada elemento são indicados os seus atributos como o nome, a definição, o tipo de dado, o tamanho do campo, a multiplicidade e a estrutura hierárquica em termos de outros elementos de que pode ser composto.

Na **Figura 8**, a título de exemplo, são apresentados o primeiro nível de descrição de metadados e também os subitens de um dos elementos. No item **Apêndices - 10.5.2. Descrição em Diagramas** é apresentado o diagrama completo.

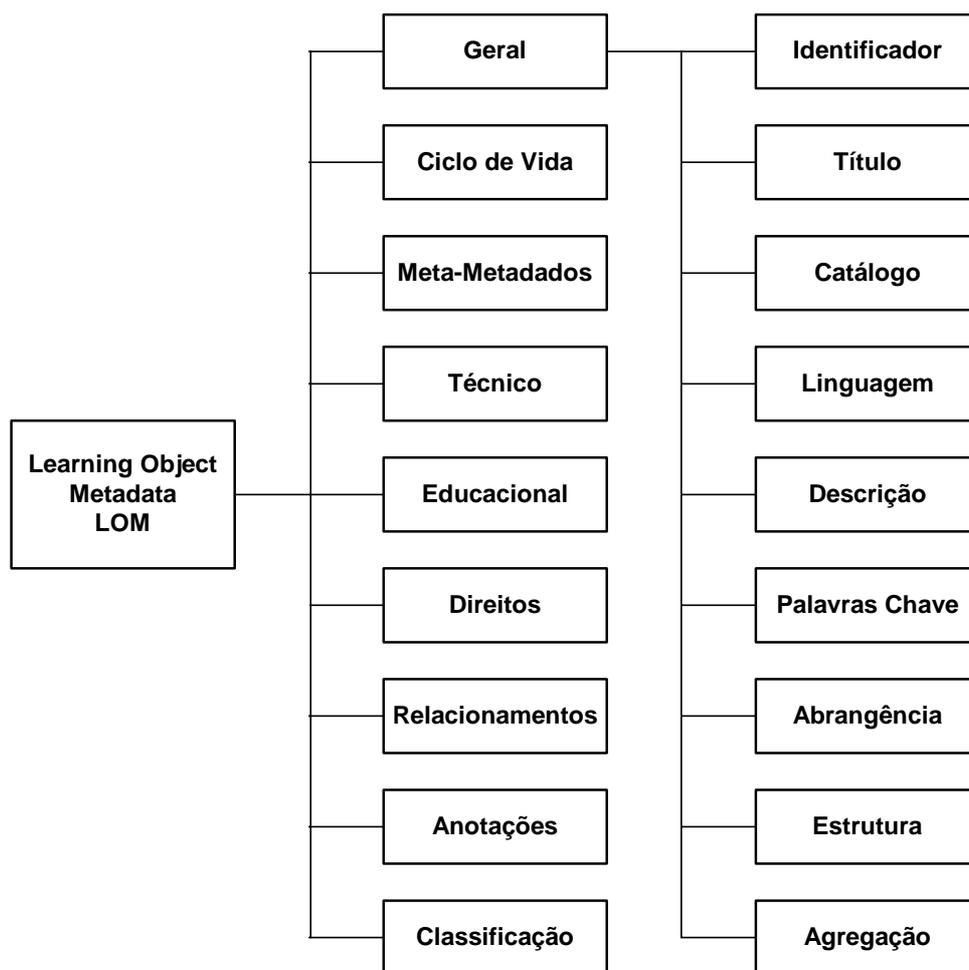


Figura 8. Primeiro nível descritivo do LOM.

Como vemos na **Figura 8**, este primeiro nível aborda os mais variados aspectos de um objeto de aprendizado. Seguem algumas explicações:

- **Geral:** reúne as características gerais sobre o objeto de aprendizado;
- **Ciclo de vida:** descreve a evolução, o estado atual, e as diversas contribuições;
- **Meta-metadados:** descreve os metadados que estão sendo utilizados;
- **Técnico:** reúne aspectos técnicos necessários para utilizar o objeto de aprendizado, bem como, suas características próprias;
- **Educacional:** descreve aspectos educacionais e pedagógicos associados;
- **Direitos:** relata condições de uso e aspectos de propriedade intelectual;

- **Relacionamentos:** descreve como este objeto de aprendizado está relacionado com outros objetos de aprendizado;
- **Anotações:** reúne comentários sobre o uso educacional do objeto de aprendizado e dados sobre a autoria dos comentários;
- **Classificação:** descreve como um objeto de aprendizado se enquadra em um sistema de classificação particular.

Todos os elementos e suas derivações são de preenchimento opcional, ressaltando-se, todavia que o preenchimento de um nível detalhado exige a presença dos níveis acima.

Além dos metadados definidos no LOM novos metadados podem ser acrescentados, sendo denominados neste caso metadados estendidos. Observe-se, porém que um metadado estendido não deve ser utilizado para substituir metadados já existentes. Outra restrição é que o procedimento de extensão não pode introduzir novos tipos de dados ou domínios para os elementos de nível superior além dos definidos no LOM.

Um resumo dos metadados LOM é apresentado no item **Apêndices - 10.5. Modelo LOM** na forma descritiva (10.5.1), em diagrama (10.5.2) e um exemplo completo de utilização (10.5.3), tendo em vista a sua relevância na proposta aqui desenvolvida. Não foi considerada necessária, todavia, a sua inclusão no corpo da dissertação.

2.4.2.

IMS – IMS Global Learning Consortium, Inc.

Criado em 1997, o projeto IMS tinha por objetivo superar três obstáculos visando prover efetivamente material e ambiente em linha para aprendizado:

- Falta de padrões para localizar e operar interativamente material de aprendizado, independente da plataforma;
- Falta de suporte para aprendizado de natureza dinâmica e colaborativa;
- Falta de incentivos e estrutura para desenvolvimento e compartilhamento de conteúdo.

Em 1999 foram publicadas as primeiras especificações de metadados.

Nesta dissertação acompanhamos especificação correspondente à versão 1.2.1 de 28 de setembro de 2001 [Ims01a, b, c].

A especificação feita pelo IMS apresenta os nomes dos metadados a serem utilizados, as definições individualizadas, a organização hierárquica, e as restrições e definições quanto ao preenchimento. O documento é composto de duas partes, na primeira é apresentado o modelo LOM IEEE/LTSC utilizado como base (normalmente não é a última versão), e na segunda, as modificações realizadas pelo IMS.

A especificação de metadados do IMS aceita extensões, assim como o LOM IEEE, desde que observadas as seguintes restrições:

1. Adição de elementos pertencentes a outros esquemas XML (namespaces);
2. Adição de elementos definidos no LOM IEEE em posição diferente da originalmente definida mantendo, todavia as demais características definidas.

2.4.3.

SCORM - Sharable Content Object Reference Model

Em 1997 foi criada a ADL com o objetivo de acelerar o desenvolvimento em grande escala de produtos de aprendizado dinâmico visando necessidades de aprendizado e de treinamento de civis e de militares. O resultado deste trabalho foi o SCORM.

O SCORM está baseado em trabalhos desenvolvidos pela AICC, IMS e IEEE, ARIADNE e outros (**Figura 9**). Entre suas definições busca atender os requisitos de alto nível para conteúdo de aprendizado do DoD⁹, tais como capacidade de reuso, durabilidade e interoperabilidade, buscando criar um modelo de referencia de especificações técnicas inter-relacionadas e manuais.

⁹ DoD – US Department of Defense

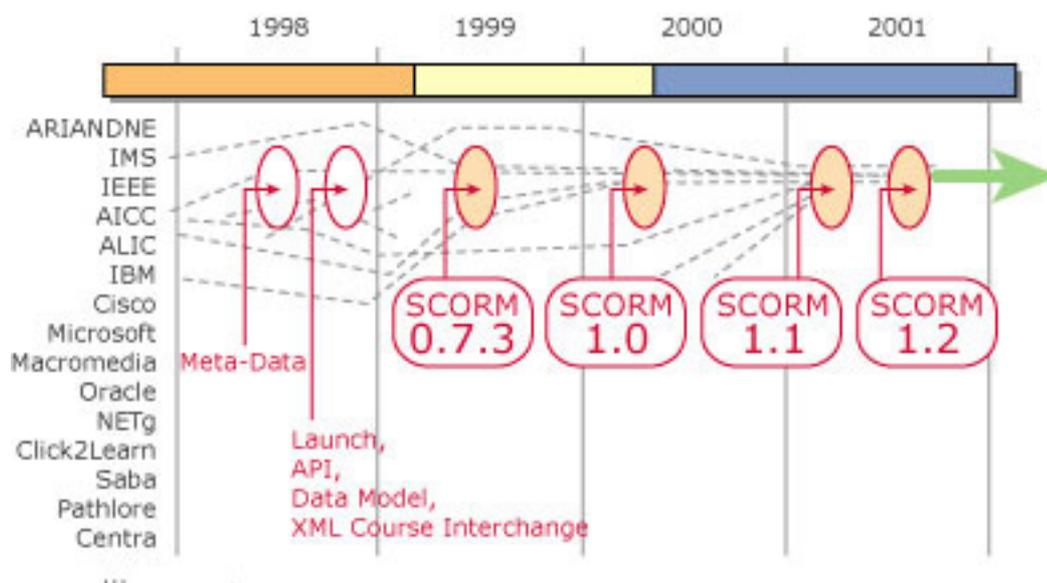


Figura 9. Evolução do SCORM [Adl03a]

A especificação SCORM inclui aspectos que afetam diversas áreas como, por exemplo, fabricantes de sistemas de gerenciamento de conteúdo de aprendizado (LCMS¹⁰, em inglês) e de ferramentas de autoria, projetistas instrucionais e de conteúdo, provedores de treinamento, entre outros. Seu foco principal é atender estes fabricantes de produtos.

Para tanto, a especificação SCORM é um conjunto de padrões técnicos que viabilizam sistemas de aprendizado baseados em WEB a encontrar, importar, compartilhar, reutilizar e exportar conteúdo de aprendizado em um modo padronizado.

A concepção do SCORM assume a existência de um conjunto de serviços que disponibilizam conteúdo de aprendizado, mantêm registro do progresso de aprendizado, controlam a ordem em que objetos de aprendizado são apresentados e acompanham o desempenho dos alunos.

Um conteúdo para WEB é normalmente constituído por ligações (*hyperlinks*) entre páginas. A especificação SCORM provê mecanismos para entregar ao instrutor o próximo material a ser utilizado quando este terminar um assunto ou uma avaliação de competência, e oferecer o conteúdo correto conforme a situação. Para um conteúdo genérico existente na WEB não é necessário ter tal

¹⁰ LCMS – Learning Content Management System

atuação, pela razão de que o usuário de um conteúdo genérico não é acompanhado e nem avaliado quanto a competências ou capacitações.

O SCORM é descrito normalmente, em sentido figurado, como sendo uma estante com vários livros. Estes livros são originários de outras organizações como ARIADNE, AICC, IMS e IEEE. Observe-se que estas especificações foram estendidas sendo acrescentados detalhes, manuais de implementação e material para teste e verificação.

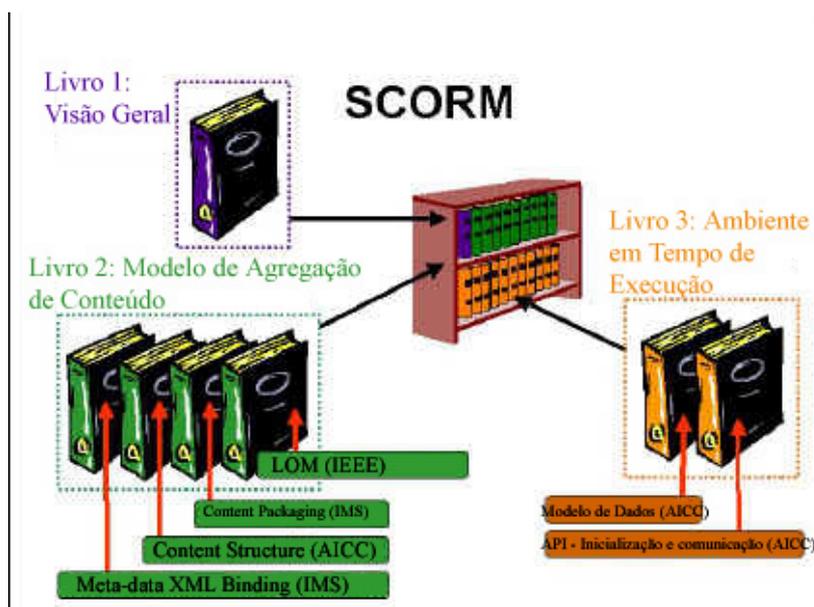


Figura 10. Estante SCORM [RHA03]

O SCORM é composto por três partes:

- **Visão Geral**¹¹: onde são descritos o modelo, os conceitos fundamentais e apresentados sumários sobre as demais seções;
- **Modelo de Agregação de Conteúdo - CAM**¹²: onde é descrito como juntar conteúdos de aprendizado de tal forma que possam ser movidos e reutilizados;
- **Ambiente em Tempo de Execução**¹³: descreve como iniciar um conteúdo, como acompanhar o progresso do instrutor e como reportá-lo.

¹¹ *Overview*, em inglês

¹² *CAM - Content Aggregation Model*, em inglês

¹³ *Run Time Environment*, em inglês

2.4.3.1. Modelo de Agregação de Conteúdo

O CAM é composto de várias partes:

- A primeira especificação corresponde ao LOM, baseado nas especificações existentes do IEEE, ARIADNE, IMS e Dublin Core. Corresponde a um dicionário de rótulos (*tags*, em inglês) utilizados para descrever um conteúdo de aprendizado sob vários focos. Estes metadados servem para descrever o que o conteúdo vem a ser, a quem pertence, custos para utilização, requisitos técnicos, contexto educacional, etc.
- A segunda especificação é a que define como conectar (*binding*, em inglês) rótulos XML dos metadados;
- A terceira especificação é a IMS Content Packing, que define como empacotar uma coleção de objetos de aprendizado, seus metadados e informações sobre como esta matéria deve ser entregue para o usuário. Este empacotamento define como um conteúdo de aprendizado de qualquer tipo pode ser transferido de modo padronizado entre sistemas. O empacotamento produz a compressão de todos os arquivos relevantes juntamente com um manifesto XML que define todo o conteúdo e o inter-relacionamento. O manifesto mantém também informações sobre o comportamento esperado do conteúdo e esta informação é interpretada pelos sistemas de gerencia de aprendizado (LMS¹⁴). Os LMS usam esta informação para determinar o que deve ser apresentado e quando.

O empacotamento é uma das características mais importantes do SCORM.

¹⁴ LMS – Learning Management System

2.4.3.2. Ambiente em Tempo de Execução

Durante a evolução do SCORM foi necessário um modo padronizado para o intercâmbio de informações entre o instrutor e o LMS.

ADL e AICC definiram para esta finalidade uma API¹⁵ que permite a comunicação com o LMS, não importando quais ferramentas tenham sido utilizadas para elaborar o conteúdo de aprendizado.

Assim que a interface fora definida, o problema passou para a definição do que deveria ser comunicado. O conjunto inicial de elementos de comunicação foi derivado a partir do modelo de dados AICC CMI¹⁶, padronizando como os sistemas LMS devem rastrear o progresso dos instrutores.

2.4.3.3. Complementação

Ao longo dos últimos anos percebeu-se que o processo de criar comportamentos complexos não é totalmente suportado pela especificação CAM.

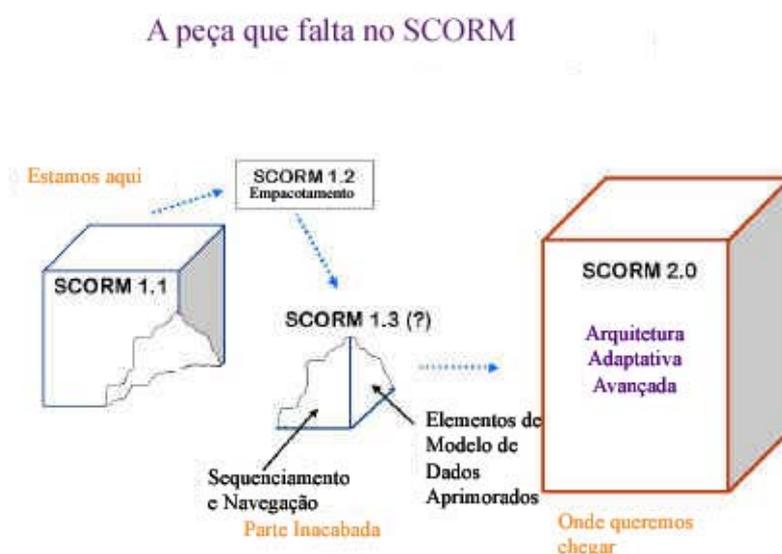


Figura 11. Complementação SCORM [RHA03]

¹⁵ API – Application Program Interface

¹⁶ CMI – Computer Managed Instruction

Desde 2001 ADL tem trabalhado juntamente com IMS focando estas limitações.

Foi colhido um conjunto de casos e teve início um processo para representar seqüências e navegações complexas por objetos de conteúdo. Este trabalho irá complementar as informações de organização existentes no CAM.

Todavia, com o modelo SCORM no seu estágio de atual de definição não é possível implementar seqüências complexas de objetos de aprendizado.

A ausência atual de informações suficientes para implementar seqüências complexas de objetos de aprendizado, resultantes da interação com um instrutor esteja explorando um conteúdo de aprendizado, leva-nos ao próximo item onde serão revistos os relacionamentos atualmente previstos entre os objetos de aprendizado.

2.5. Associações entre Objetos de Aprendizado

Neste tópico serão apresentadas as definições de relacionamentos previstas entre os objetos de aprendizado, definidas pelos itens **1.8 - Estrutura** e **7 - Relações** no LOM IEEE.

Estas definições podem ser classificadas como: (1) caráter estrutural ao definir o tipo de composição e (2) caráter semântico ao estabelecer o relacionamento.

Segundo o LOM IEEE [Ims00], a estrutura dos objetos de aprendizado pode ser:

- **Atômica** (*atomic*): o objeto de aprendizado é indivisível;
- **Coleção** (*collection*): o objeto de aprendizado é uma biblioteca ou um repositório de recursos. Ele é composto por uma seqüência não ordenada de objetos de aprendizado e, portanto sem noção de navegação entre os componentes, sendo que os componentes podem ser acessados diretamente. Por exemplo: uma coleção de imagens;
- **Mista** (*mixed*): o objeto de aprendizado pode ser utilizado de vários modos, sem que nenhum seja efetivamente dominante;
- **Linear**: o objeto de aprendizado é composto por uma seqüência ordenada de objetos de aprendizado;

- **Hierárquica** (*hierarchical*): o objeto de aprendizado é composto por uma estrutura em árvore. Os desvios podem ser feitos segundo um processo educacional, como por exemplo, em resposta a um questionário. A navegação é iniciada por um ponto único de entrada. A estrutura hierárquica implica em dispor de níveis bem definidos de importância, granularidade ou de especificidade;
- **Rede** (*networked*): o objeto de aprendizado é composto por objetos de aprendizado ligados entre si, não exatamente na forma hierárquica e nem exatamente na forma linear. O início da navegação pode ser em qualquer ponto e pode estender-se além do objeto em questão. Nós individuais podem ser construídos com o conteúdo de nós que estejam várias ligações adiante;
- **Ramificada** (*branched*): caracterizada pela existência de pontos de decisão;
- **Parcelada** (*parceled*): dividida em componentes ou domínios distintos, que podem ser considerados como sala de interesse;

A associação semântica enumerada no item 7.- Relações do LOM é originária dos termos utilizados pelo Dublin Core:

- **Relação todo/parte**: é uma relação de composição informando se um objeto de aprendizado é parte de (*IsPartOf*) ou tem parte (*HasPartOf*) de outro objeto de aprendizado;
- **Relação de versão**: quando é relatada a evolução de um objeto de aprendizado em relação a outro, utilizando os termos de ser uma versão (*IsVersionOf*) ou possuir uma versão (*HasVersion*);
- **Relação de transformação de formato**: quando é relatada a evolução de um objeto de aprendizado em relação a outro em termos de representação, utilizando os termos de ter o formato (*IsFormatOf*) ou estar no formato (*HasFormat*);
- **Relação de referência**: quando um objeto de aprendizado menciona ou é mencionado por um outro conteúdo de objeto de aprendizado, utilizando os termos (*References*) ou (*IsReferencedBy*);

- **Relação de criação:** quando é relatada a evolução de um objeto de aprendizado a partir de outro, utilizando os termos de ser uma versão (*IsBasedOf*) ou compor uma versão (*IsBasisFor*);
- **Relação de dependência:** quando é relatada a dependência de um objeto de aprendizado em relação a outro, utilizando os termos de haver necessidade (*Requires*) ou suprir uma necessidade (*IsRequiredFor*);

Este tópico teve por objetivo expor a estrutura estática dos relacionamentos e a ausência de interação com fatores externos durante a sua navegação, exatamente como assinalado no último item do SCORM.

No próximo item será apresentada uma breve noção validação de dados XML contra uma estrutura proposta.

2.6. Validação de Metadados

As descrições dos objetos de aprendizado contendo metadados são implementadas em XML e, como todo documento XML, podem ser validadas quanto à correção do preenchimento e da estrutura.

Existem duas formas para validar dados em XML: (1) definição de documento (DTD¹⁷) e (2) esquema XML. Nos dois casos são definidos os nomes dos elementos e os atributos de cada elemento (tipo de dado, valores aceitáveis para preenchimento, número de ocorrências, etc.). A opção esquema XML oferece maiores possibilidades de extensão com adição de informações de herança e regras de apresentação.

Este conceito sobre validação de dados XML será utilizado posteriormente na implementação de soluções.

Com estes tópicos pretendemos ter apresentado uma visão dos padrões atualmente existentes para a descrição dos objetos de aprendizado e das diversas tendências existentes para o seu compartilhamento.

No próximo item será apresentada uma breve noção sobre sistemas de aprendizado complementando assim os assuntos considerados relevantes para este trabalho.

¹⁷ DTD – Document Type Definition

2.7. Sistemas para Área de Aprendizado

Os sistemas existentes voltados para o atendimento da área de aprendizagem são enquadrados em duas categorias:

- a. Sistema para Gerenciamento de Aprendizado (LMS)
- b. Sistema para Gerenciamento de Conteúdo de Aprendizado (LCMS).

Conforme definições apresentadas no LCMS Council [Lcm03] um sistema LCMS tem as seguintes características:

- "LCMS é um ambiente de múltiplos usuários onde equipes de desenvolvimento de material de aprendizagem podem criar, armazenar, reutilizar, gerir e disponibilizar conteúdo didático digitalizado a partir de um repositório de objetos centralizado" [BH03];
- Deve prover quatro características [BFA01]:
 - a) **Repositório de Objetos de Aprendizado:** um banco de dados central que armazene e gere conteúdo de aprendizagem criado por múltiplos autores. Objetos de aprendizagem deste repositório podem ser disponibilizados para educadores individualmente, agregados em módulos maiores ou mesmo em cursos inteiros;
 - b) **Ferramentas de Autoria:** compreende aplicações automatizadas de autoria utilizadas para criar objetos de aprendizagem reutilizáveis que serão geridos pelo repositório. Modelos de formato podem ser disponibilizados para incorporar características padronizadas;
 - c) **Interface de Recuperação:** utilizada para disponibilizar objetos de aprendizagem baseada no perfil do educador ou outra forma de pesquisa. Normalmente provê também rastreamento e ligações para fontes relacionadas de informação, bem como suporte para múltiplos recursos de avaliação e de retorno para os usuários;
 - d) **Aplicação Administrativa:** utilizada para gerenciar registros de estudantes, ativar cursos a partir de catálogo, rastrear e reportar progresso dos alunos e prover outras funções administrativas.

A **Figura 12** [Don02] apresenta a idéia geral sobre LCMS.



Figura 12. Componentes de um LCMS [Don02]

Na **Figura 13** [Nich01] é apresentada uma visão esquemática sobre um sistema de gerência de aprendizado.

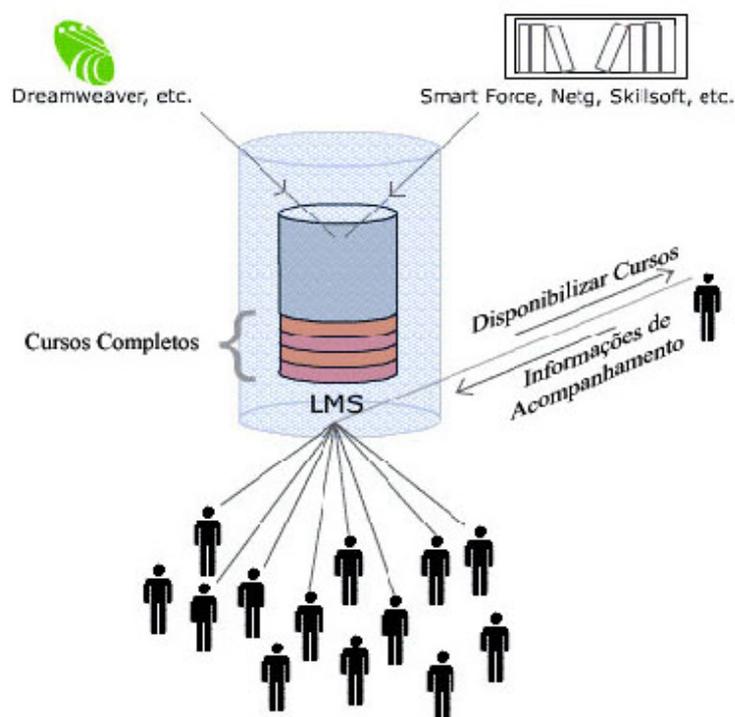


Figura 13. Esquema de um LMS [Nich01]

Um sistema de gerenciamento de conteúdo de aprendizado é diferente de um sistema de gerência de aprendizado. Enquanto este tem por foco principal o aluno, o primeiro focaliza principalmente o conteúdo de aprendizado.

A maioria dos LMS implementa as características abaixo:

- a. Colaboração (bate-papo, quadro de avisos);
- b. Gerência de salas de aula
- c. Gerência da base de dados de conhecimento

Na **Figura 14** [Nich01] é apresentada uma visão esquemática sobre um sistema de gerência de conteúdo de aprendizado

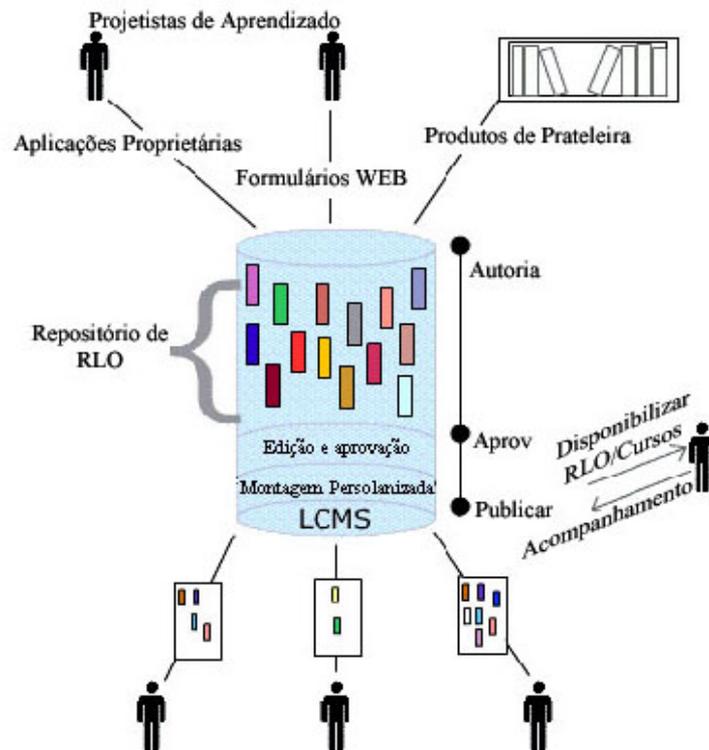


Figura 14. Esquema de um LCMS [Nich01]

