
**MODELANDO DADOS:
DE REGISTROS A HIPERMÍDIA (E ALÉM?)
DATA MODELING:
FROM RECORDS TO HYPERMIDIA (AND BEYOND?)**

Carlos M. TOBAR Toledo
Inf. PUCAMP
DCA-FEEC UNICAMP

ABSTRACT

Data models have been traditionally used as data base design and implementation tools. More than this, in some cases, they have been used to specify and generate the structural and access description for the data stored in data bases. Considering the goal of presenting a broader set of situations where it is possible to use data models, together with the consideration of the evolution of user and application requirements regarding data storage and retrieval, in this paper they are presented: the subject of data modeling, data model definitions under different approaches, a historical description of data model evolution, and a data model hierarchy considering abstraction levels involved with mechanisms for persistent data storage and with their development process.

Keywords: Data Model, Data Modeling, Data Management, Information Systems

RESUMO

Modelos de Dados têm sido usados tradicionalmente como ferramentas para o projeto e implementação de bases de dados. Mais que isso, em alguns casos, têm sido usados para a especificação e geração da descrição estrutural e de acesso aos dados armazenados em bases de dados. Objetivando apresentar um leque mais abrangente de situações onde se pode utilizar modelos de dados e tomando como base a evolução dos requisitos de usuários e aplicações, em relação ao armazenamento e recuperação de dados, neste trabalho são abordados: o assunto foco da modelagem de dados, as definições para modelos de dados sob diferentes abordagens, uma descrição histórica da evolução dos modelos de dados e uma hierarquia de modelos de dados, considerando níveis de abstração envolvidos com os mecanismos para armazenamento persistente de dados e com o processo de seu desenvolvimento.

Palavras Chave: Modelo de Dados, Modelagem de Dados, Gerência de Dados, Sistemas de Informação.

1. INTRODUÇÃO

Um bom *modelo*, tal qual uma boa história, é uma descrição de um mundo (ou, ao menos, de

uma porção selecionada e isolada deste mundo), que pode ser o real ou algum outro imaginário. É uma maquete, representação codificada de objetos, processos e seus inter-relacionamentos. É a

visão de um universo colocada dentro de uma xícara de café.

Modelos permitem a representação de processos, entidades ou eventos de forma *abstrata*¹, possibilitando o estudo e análise de partes de um problema ou situação, escondendo detalhes que podem não ser importantes para a maioria dos interessados. Modelos são usados em todas as áreas de conhecimento, principalmente naquelas preocupadas com a produção de bens, com o entendimento de fenômenos e com a previsão de tendências.

O real poder dos modelos reside na habilidade que seus usuários ganham para poder especular sobre as conseqüências de mudanças nas características dos elementos enfocados e modelados, tratando-os em conjunto como um sistema, com parâmetros de entrada e resultados.

Em particular, o interesse existente é com os ditos *modelos de dados*, cujo conceito, para [1], é tão genérico quanto qualquer conceito matemático que apresente os valores que objetos de dados podem assumir (parte estática) e as operações realizáveis sobre os dados (parte dinâmica). Em particular o interesse recai em prescrições, não necessariamente com base matemática, que são usadas no desenvolvimento de sistemas de informação, mais especificamente para a integração de dados e respectivas informações, tradicionalmente, contemplando sistemas que venham a ter bases de dados (BDs) ou arquivos, ou mesmo aqueles desenvolvidos a partir de uma abordagem orientada a objetos [16].

Devido à forte interdependência histórica entre modelos de dados e a arquitetura e estruturação de BDs, algumas vezes a palavra modelo é usada para significar descrição esquemática² ou física de uma BD. Essa mesma interdependência ocasiona uma dificuldade de se descrever, entender e utilizar modelos de dados sem necessariamente existir uma BD envolvida.

Assim, tanto quanto possível no que se segue, tenta-se desvencilhar modelos de dados de BDs, considerando que um modelo de dados corresponde a um mecanismo que permite especificar descrição e, em alguns casos, a interpretação de dados³. Adota-se, então, o termo *modelagem* para o resultado obtido na utilização de tal mecanismo, seja este resultado incorporado ou não a uma BD.

Inicialmente, após esta introdução, são apresentados conceitos envolvidos com o assunto foco da modelagem de dados, ao mesmo tempo em que se procura definir o que vem a ser um modelo de dados, em termos de componentes. Segue-se uma descrição da evolução dos modelos de dados, considerando-se o tipo de requisitos das aplicações a que se destinam, além da apresentação de uma hierarquia de modelos de dados, que permite visualizar o leque de situações onde se pode utilizar modelos de dados e a possibilidade de mapeamento entre modelos de diferentes níveis hierárquicos. Finalmente termina-se com algumas conclusões.

2. O ASSUNTO FOCO DA MODELAGEM DE DADOS

Em um mundo (real ou imaginário) existem objetos (entidades ou fenômenos) e fatos associados a objetos, além de processos e eventos, envolvendo fatos e objetos, sobre os quais deseja-se obter informação ou sobre os quais deseja-se passar informação, seja para uma ou várias pessoas, seja num determinado instante, seja repetidamente, em instantes imprevisíveis, e para isso tem-se a intenção de armazenar a informação⁴.

Pessoas, de forma natural, armazenam informações em suas memórias e, quando as recuperam, o fazem através de associações, sendo que a apresen-

(1) Abstração é a maneira pela qual, mentalmente, tratamos uma informação para que possa ser registrada, armazenada e recuperada a partir de uma estrutura, em diferentes níveis de detalhe e diferentes perspectivas [19]. A finalidade de uma abstração é permitir que o usuário se concentre nos detalhes relevantes, ignorando os irrelevantes (em que o significado de relevante depende do usuário e da aplicação em questão) [17].

(2) A conceituação do termo esquema é realizada no item 4, adiante.

(3) Definições para os termos descrição e interpretação de dados são apresentadas no item 2, em seguida.

(4) Conforme é visto no item 2.1, em seguida, não se consegue armazenar informação propriamente em meios sintéticos, assim, considera-se a intenção de seu armazenamento. No contexto deste trabalho, considera-se que essa intenção é dirigida pela posterior passagem da informação, independentemente do número de pessoas alvo e da quantidade de vezes que isso possa ocorrer.

tação (reprodução) é realizada, na maioria das vezes, através da utilização de linguagem natural, em termos de objetos, propriedades de objetos, relacionamentos, comportamento, restrições e instante de ocorrência. No lugar da linguagem natural pode ser considerado um outro sistema de símbolos, utilizado para a troca simbólica de informações, tal como a escrita, a pintura, o desenho entre outros.

Historicamente, quando intenciona-se armazenar informação em meios sintéticos⁵, a título de reprodução posterior, necessita-se captar, capturar e registrar aspectos da parte do mundo que interessa, denominado *domínio de interesse*, geralmente através da representação das percepções, utilizando-se de um sistema de símbolos, que também possa ser usado para posterior reprodução.

As representações de percepções, por sua vez, podem ser mais ou menos ricas de significado⁶, em relação ao mundo focado. Primeiro porque o processo de percepção é complexo, devido à quantidade de detalhes, perspectivas e contextos possíveis de existir a cada instante e, portanto, necessita ocorrer através da realização de uma síntese do que realmente pode ser e é percebido, e, segundo, porque o tipo de meio, que se usa para se representar e registrar os elementos de informação de interesse, pode não ser o mais adequado para essa representação e registro⁷ e, ainda, porque pode oferecer pouca ou nenhuma possibilidade de se incorporar significado.

Atualmente é possível a captura, o armazenamento de registro e a reprodução de imagens (estáticas ou temporalmente contínuas), sons e, até, sensações sensoriais, em meios sintéticos, sem a necessária interferência de pessoas para a representação dos objetos de interesse. Isto significa ser possível o armazenamento do registro de cópia quase fiel dos elementos de interesse. Quase fiel porque vários elementos do domínio de interesse não são

capturados, havendo, na grande maioria dos casos, a necessidade de contextualização do que foi capturado e, assim, recai-se na situação anterior da percepção e representação. O resultado do processo de registrar, de agora em diante, compreende tanto a representação quanto a cópia “quase fiel” no domínio de interesse.

Mesmo assim, apesar da quase fidelidade, os elementos de interesse capturados podem ser reproduzidos e percebidos repetidamente, por diferentes pessoas, gerando informação sem intermediações e com distorções minimizadas⁸.

Ao processo de registrar informações podem ser somados os seguintes complicadores: um mesmo elemento de informação pode necessitar ser descrito a partir de diferentes perspectivas ou, mesmo, ser capturado de diferentes maneiras⁹, que necessitam permanecer consistentes; além disso, pode ser necessária a descrição de elementos que evoluem no tempo¹⁰, com a consistência entre as versões de sua descrição necessitando, também, ser mantida.

2.1. Informação e dado

Utiliza-se a seguinte definição para informação: é o incremento de conhecimento que pode ser inferido a partir da percepção de (parte de) um mundo. Resultados de um processo de registrar elementos de informação em meios sintéticos, a título de posterior recuperação e apresentação de informação, são denominados dados.

O dado, sujeito da modelagem de dados, corresponde, então, primeiro ao registro de um elemento de informação, que, eventualmente, por ter sido percebido e representado, se torna discreto, devido à abstração usada na sua representação; e, segundo, aos descritores¹¹ que permitem a sua manipulação.

(⁵) Meios sintéticos são aqueles diferentes e externos à memória humana, como o papel, a fotografia, etc.

(⁶) A conceituação para significado é realizada no item 2.1, em seguida.

(⁷) O termo registro é usado no sentido de ato de registrar e não deve ser confundido com o conceito de estrutura de dados (originário do primeiro). Exemplo de meio inadequado para registro é a escrita para se registrar um ruído ou uma sinfonia para se registrar um pagamento efetuado.

(⁸) Distorções são, naturalmente, produzidas pelas intenções e diferenças entre os agentes responsáveis pela percepção e representação.

(⁹) Exemplo de captura em diversas perspectivas é de uma radiografia, um eletroencefalograma e uma ecografia de uma mesma pessoa.

(¹⁰) Exemplo de elementos que evoluem no tempo é o de um produto em fase de projeto.

(¹¹) Um descritor é um mecanismo que permite a descrição de algum aspecto do dado, como, por exemplo, a sua estrutura.

Todo dado corresponde a elementos de informação, através dos quais é possível obter-se informação sobre o domínio de interesse. O que se registra e, talvez, descreve, total ou parcialmente, ou representa o elemento focado da informação é denominado predicado e constitui o que costuma-se denominar conteúdo, enquanto que o significado do predicado, também denominado asserção, ajuda na interpretação do predicado, quando da reprodução do dado.

Convencionalmente, o que existe de relativo à interpretação para dados reside nos programas¹² que os manipulam. Isto, geralmente, ocasiona, quando os dados são compartilhados por vários programas, diferentes interpretações. Assim, é interessante que as asserções, tanto quanto possível, sejam armazenadas junto aos predicados e, desta maneira, o dado passa a apresentar duas componentes complementares e a modelagem de dados deve passar a incorporar tanto predicados quanto asserções.

O armazenamento de dados exige a sua organização, ou seja, o seu agrupamento em estruturas para que possam ser manipulados. Essa estruturação, em princípio, está diretamente ligada a detalhes de funcionamento de computadores (hardware e software) e é necessária tanto para os conteúdos, que são diretamente capturados no domínio de interesse e registrados em estruturas providas automaticamente pelas ferramentas de captura, quanto para os conteúdos frutos de representação, registrados em estruturas providas por ferramentas de edição ou autoria.

O conceito de dado passa a abranger sua organização estrutural, o conteúdo ou valor contido nessa estruturação e indicações implícitas ou explícitas (para o software controlador) sobre essa organização, de modo a permitir localização, acesso e modificação. Geralmente, dados apresentam um tipo, cuja sintaxe estabelece regras de manipulação (operações e restrições aplicáveis) e cuja semântica estabelece o significado do estado pré e pós solicitação de operações. Essa semântica, via de regra, não se relaciona a domínios de interesse, mas é computacionalmente orientada e genérica.

2.2. Rumo a uma primeira definição para modelo de dados

As construções usadas para modelar estruturalmente os objetos do domínio de interesse são comumente denominadas objetos ou entidades. Objetos podem se relacionar através de categorias ou classes, se apresentam similaridades, que por sua vez são descritas como propriedades. Os aspectos que associam uma categoria de objetos a outra(s) são modeladas por construções denominadas relacionamentos.

As propriedades e relacionamentos de uma categoria, por sua vez, podem ser classificados em estáticos e dinâmicos. Os estáticos correspondem às propriedades e relacionamentos que são, relativamente, invariantes no tempo, enquanto que os dinâmicos correspondem à natureza evolutiva do domínio de interesse.

Geralmente, um domínio de interesse não é estático [38], assim, a sua modelagem necessita considerar como e em quais condições ocorre a evolução dos aspectos dinâmicos. Este "como condicional" pode ser representado através de operações¹³, que agem sobre propriedades e/ou relacionamentos e que podem ser desde operações básicas atômicas, até transações complexas.

A existência de categorias e a necessidade da incorporação de operações, para a representação mais completa do domínio de interesse, ocasiona a necessidade da especificação de regras de integridade ou *restrições*, que garantam que:

- haja controle na manipulação dos aspectos dos objetos de acordo com a(s) categoria(s)¹⁴ a que pertençam,
- os aspectos dinâmicos, uma vez que sofram mudança (evolam) ocasionada por uma ou mais operações, saiam de um estado válido para um outro estado válido, ou
- operações não ocasionem mudanças não previstas nos aspectos dinâmicos ou estáticos.

(¹²) A partir deste ponto em diante, são considerados apenas computadores e seus periféricos como os mecanismos para a captura de elementos de informação, bem como para descrição, registro, armazenamento e reprodução de dados.

(¹³) Do ponto de vista conceitual, para [38], operações podem ser consideradas objetos, até mesmo com relacionamentos entre elas.

(¹⁴) Em alguns modelos de dados é possível que um determinado objeto pertença a mais de uma categoria, seja porque existe uma hierarquia de categorias, caracterizando o relacionamento "é-um", seja porque existe hereditariedade múltipla (as propriedades e relacionamentos do objeto são descritas a partir de mais de uma categoria).

A própria noção de transação, segundo [17], é necessária porque, implicitamente representa uma restrição, em relação a execução “atômica” de todas as operações que a compõem, e talvez, na mudança de um aspecto dinâmico não seja possível ir de um estado válido para outro, sem passar por um estado não válido.

Em um primeiro momento, então, um modelo de dados compreende, na visão de [17], os seguintes componentes:

- coleção de tipos de estruturas de dados, ou seja, regras segundo as quais os dados estão estruturados (convencionalmente em tabelas ou grafos), geralmente, para a representação das categorias de objetos no domínio de interesse,
- coleção de operadores ou regras de inferência (convencionalmente de especificação), que podem ser aplicadas às descrições e registros de entidades, de acordo com os tipos de estruturas de dados,
- coleção de regras de integridade genéricas, que implícita ou explicitamente definem um conjunto de estados ou mudanças de estado consistentes e íntegros para o conteúdo e suas estruturas.

Note-se nesta primeira definição de modelo de dados a preocupação com a descrição de aspectos de elementos de informação, em detrimento do significado e, por conseguinte, da interpretação da descrição.

A ênfase descritiva se deve à (ainda atual) dificuldade dos computadores manipularem linguagem natural, ou outra forma para codificar significados de descrições (interpretações), além do (não tão atual) alto custo do armazenamento não volátil. Assim, a maioria dos modelos de dados pioneiros representam informação em estruturas de dados denominadas *registros*, pois estas construções, além de serem a base comum para se registrar dados nas aplicações então existentes, são mais amenas à manipulação e ao consumo de recursos computacionais.

Em decorrência, se estabeleceu uma (falsa) crença de que, para a representação de percepções, é necessária (não suficiente) apenas a representação

de elementos de informação atômicos¹⁵ existentes [44], no entanto, não existe mecanismo de representação conhecido que permita uma interpretação fácil de elementos complexos a partir de elementos atômicos. Além disso, passou a ser comum o descarte do estado passado dos elementos de informação, quando da evolução do domínio representado, mantendo-se armazenando apenas o último estado válido, resultante da ação das últimas operações executadas.

Em termos computacionais, neste primeiro momento, um modelo de dados define as regras gerais de acordo com as quais os dados podem ser organizados, permitindo a especificação da estrutura dos dados e, algumas vezes, das operações, que se podem realizar sobre estes, e das restrições de integridade. Quando isto ocorre, segundo [44], diz-se que se estabeleceu *aintençãodo* domínio de interesse que se representa, enquanto que os valores, que eventualmente venham a ser armazenados nessas estruturas, que sofram essas operações e obedeçam a essas restrições, constituem a *extensão*.

2.3. A necessidade de incorporação de significado aos dados

A separação da descrição de seu significado, que ocorre na maioria dos modelos de dados mais difundidos, causa dificuldades no uso dos dados. Como decorrência, passa a haver a necessidade e a preocupação com a produção de asserções, na tentativa de também capturar e codificar parte do significado intrínseco e orientado ao domínio de interesse (e à aplicação, se for o caso).

Tentativas de prover maior significado, do que aquele inerente à representação do tripé estrutura, operação e restrição, tomando tipos registro como estrutura, são descritas como a incorporação de conteúdo semântico¹⁶ às representações ou, mesmo, aos elementos de informação capturados. Na realidade, incorporação de um maior poder de expressão, que facilite a interpretação dos dados.

O enfoque inicial da captura adicional de semântica ficou centrado nas dependências interrelacionais tanto de objetos, quanto de categorias,

(¹⁵) Um elemento de informação atômico é um elemento que não apresenta decomposição.

(¹⁶) Semântica é a disciplina que lida com relacionamentos entre palavras e coisas às quais se referem essas palavras.

principalmente com a preocupação de distinguir várias formas de relacionamentos e, em decorrência, com a criação de padrões de significado para cada tipo de dependência e correspondentes operações e regras de integridade aplicáveis.

2.4. Uma definição de modelo de dados com aspectos semânticos

Enquanto os modelos baseados em registros provêm uma maneira única de se representar elementos de informação, os modelos semânticos, através de abstrações, permitem a modelagem e a visão dos elementos de informação em vários níveis.

Apesar de não se encontrar uma definição, para modelos de dados que incorporem aspectos semânticos, que seja amplamente aceita, como ocorre com a definição para os modelos de dados baseados em registros, consideram-se como fundamentais as seguintes características [38]:

- representação de objetos atômicos (tipos primitivos),
- representação de relacionamentos (como atributos, entidades, elementos independentes ou funções),
- suporte a abstrações padrão (generalização, agregação, classificação e associação)¹⁷,
- redes ou hierarquias de relacionamentos,
- derivação/hereditariedade¹⁸ (para manipulação de informações repetidas),
- restrições de inserção, remoção e modificação (para manutenção da integridade),
- grau de expressão da semântica de relacionamentos (cardinalidade, valores nulos, relacionamentos inversos, derivações, hereditariedade ou valores "default").

(¹⁷) Generalização é a maneira pela qual diferenças entre objetos similares são ignoradas para a formação de uma categoria (tipo) de maior ordem, em que as similaridades podem ser enfatizadas.

Agregação é a forma pela qual relacionamentos "ser-parte-de" entre categorias podem ser considerados através de uma categoria de nível mais alto, que reflita o significado do "todo".

Classificação é a forma de abstração em que uma coleção de objetos é considerada em um nível mais alto, representando o relacionamento "é-instância-de".

Associação é a forma de abstração em que um relacionamento entre objetos membros é considerada em um nível mais alto, significando "é-membro-de".

(¹⁸) Derivação é a maneira pela qual propriedades ou relacionamentos podem ser computados, a partir de outras propriedades ou relacionamentos.

Hereditariedade é a maneira pela qual os objetos de uma determinada categoria têm propriedades e relacionamentos definidos, a partir de categoria(s) de mais alta ordem do que a da categoria a que pertencem.

(¹⁹) A modelagem de dados original e tradicionalmente é considerada como constituída pela, só e tão somente, modelagem estática.

Note-se nesta definição que não são enfatizadas as componentes relativas a restrições de consistência genéricas e, principalmente, a operações, embora possa-se argumentar que elas estão implícitas nas demais componentes. Note-se, também, que a quantidade de significado incorporada à modelagem é muito reduzida.

Um modelo semântico, evolução do modelo baseado em registros, é composto, então, do tripé estrutura, operação e restrição, relativo à componente organizacional dos dados, em conjunto com relacionamentos, devidamente identificados de acordo com seus significados semânticos, sejam funcionais ou relacionais, além de algumas restrições de integridade e operações básicas para manipulação de propriedades dinâmicas.

2.5. A necessidade de incorporação de aspectos comportamentais aos dados

A descrição de operações e transações é denominada *modelagem dinâmica* e é o contraponto da *modelagem estática*, que abrange basicamente a descrição de propriedades de categorias e relacionamentos e é enfatizada nos modelos de dados baseados em registros e nos modelos de dados semânticos.

Da mesma maneira que, tradicionalmente, o significado dos dados residia nos programas que os manipulavam, a modelagem dinâmica era considerada e tratada como uma perspectiva à parte da modelagem de dados¹⁹. Porém, a partir do instante em que, por motivos de redundância e diferenças de interpretação, passa-se a considerar interessante que registros de elementos de informação residam juntamente com seus significados, passa-se também a considerar

interessante agregar a descrição comportamental dos elementos de informação e, assim, os dados passam a apresentar duas perspectivas de especificação: a estrutural estática e a comportamental ou dinâmica; as quais devem permitir a geração de descrições, juntamente com significados semânticos incorporados.

A abordagem da orientação a objetos para implementação de programas, oriunda de linguagens de programação, permite que os aspectos comportamentais, tanto dinâmicos quanto de integridade, sejam especificados de maneira integrada aos aspectos estruturais, através do encapsulamento²⁰ das operações e restrições (métodos) com a estrutura e conteúdo das partes dos objetos, que são por elas modificados e que necessitam permanecer íntegros. Além disso, essa abordagem permite que uma certa quantidade de aspectos semânticos seja descrita. Assim, nada mais normal do que passar a considerar modelos de dados orientados a objetos.

Os modelos de dados orientados a objetos são considerados uma evolução sobre os modelos de dados semânticos, pois permitem a consideração da perspectiva comportamental [2], apesar de que, em termos de poder de expressão para abstração estrutural, costumem ser menos poderosos que os modelos semânticos. Existem, ainda, pesquisadores (por exemplo [38]) que consideram que os modelos de dados semânticos compreendem os ditos orientados a objetos, devido a existência de alguns modelos de dados híbridos.

2.6. Uma definição de modelo de dados com aspectos comportamentais

Um exemplo de modelo de dados orientado a objetos encontra-se definido no padrão ODMG-93²¹ [9], que resumidamente é constituído por:

- uma primitiva básica para modelagem, que é o objeto,
- categorização dos objetos em tipos (objetos com mesmo tipo exibem comportamento e domínio de estado comuns),

- comportamento dos objetos, que é definido por um conjunto de operações que podem ser executadas sobre uma categoria de objetos,
- estado dos objetos, que é definido pelos valores de um conjunto de propriedades (atributos ou relacionamentos com outros objetos).

Segundo [39] os três primeiros aspectos constituem a base mínima de um modelo de dados com orientação a objetos.

Os dois primeiros aspectos apresentam correspondentes nos modelos semânticos, e, de forma genérica, orientam a modelagem estrutural para os objetos de informação no domínio de interesse. Ambos os aspectos sugerem que um modelo orientado a objetos suporta primitivas para abstração semântica explícita, para representação direta de objetos, categorização de objetos e restrições de integridade, entre outras. As primitivas para modelagem semântica, mais comumente encontradas, são a classificação (categorização) e a herança.

Modelos de dados orientados a objetos podem, então, ser definidos, segundo [31], como constituídos por:

- um núcleo do modelo, que corresponde à definição de modelo orientado a objetos, vista anteriormente, mais
- restrições semânticas e de integridade, além de
- um número de relacionamentos semânticos.

Note-se que esta definição, apesar de ser uma evolução em relação aos modelos de dados existentes, continua essencialmente com a preocupação descritiva, sendo que o pouco de incorporação de significado é menor ou igual à aquela obtida através dos modelos de dados semânticos. Note-se também, que a modelagem de objetos de interesse obrigatoriamente passa pela categorização, cuja descrição compreende tanto os aspectos estruturais quanto os comportamentais.

⁽²⁰⁾ O encapsulamento de conteúdo, estrutura e operações provê as seguintes vantagens, segundo [6]: o ocultamento da estrutura dos dados e da implementação das operações, a conseqüente independência de dados e a modularidade.

⁽²¹⁾ ODMG-93 é o resultado da definição de um padrão para sistemas para gerência de bases de dados orientadas a objetos, patrocinado e com contribuições de diversas companhias de software, incluindo tanto padronização para o modelo de dados quanto para linguagens de manipulação, definição e consulta.

2.7. A necessidade de incorporar estruturação associativa e aspectos temporais

Os modelos de dados vistos até aqui (baseados em registros, semânticos e orientados a objetos) têm sua definição dominada pela perspectiva organizacional e, assim, via de regra, oferecem regras para estruturação do conteúdo dos dados expressas de maneira compreensiva. E essas regras, na maioria dos casos, se baseiam no processo de desenvolvimento da escrita, principalmente, da cultura ocidental.

Dentro deste contexto, o conceito de hipertexto, e por extensão de hipermídia²², introduz uma nova perspectiva organizacional à modelagem de dados. A hipermídia permite a personalização da organização do conteúdo dos dados, através da criação de conexões arbitrárias (ligações) entre pedaços de conteúdo de dados.

As ligações em conjunto podem ser vistas como uma estrutura ligada para o conteúdo dos dados, que tenta copiar a habilidade do cérebro de acessar informação rápida e intuitivamente, através de referências (associações). Esta estrutura associativa não é a única forma organizacional em um hiperdocumento, continua havendo ainda a necessidade da organização estrutural tradicional.

A estrutura associativa, denominada *hiperestrutura*, é tão diferente de qualquer outra forma organizacional conhecida, que é até difícil de descrevê-la como uma verdadeira estrutura. Ela não é linear, é explicitamente não seqüencial, nem é hierárquica, nem em rede. Portanto, pode parecer caótica e entrópica. Além disso, por sua própria natureza, a hipermídia está diretamente relacionada à interação, ou seja, a hiperestrutura existe para permitir a um usuário, interativamente, percorrer as ligações (navegar) e acessar o conteúdo dos dados, teoricamente, sem ordem preestabelecida. Essas peculiaridades acarretam, segundo [20], a necessidade do desenvolvimento de um novo tipo de modelo de dados específico

e, segundo [29], em particular para o aspecto da navegação.

Estruturalmente, um sistema com hipermídia é equivalente a uma rede semântica, onde nós podem ser expandidos em uma variedade de formas visuais, auditivas e, porque não, sensoriais. Mais do que uma rede de nós e ligações, é um objeto estruturado com comportamento e restrições, além, é claro, de significado. Sem um modelo de hipermídia, o desenvolvimento de sistemas com hipermídia torna-se não gerenciável conforme o número de nós e ligações cresce de maneira desestruturada [37]

Não há um consenso de como modelar hipermídia, várias são as propostas para modelagem²³ e existem, até, alguns esforços para padronização, como é o caso do MHEG²⁴ [10].

A hipermídia significa, implicitamente, uma mistura de hiperestrutura com multimídia, onde o conteúdo dos dados pode ser registrado em diferentes tipos de mídia. Mas, independentemente da estrutura associativa, a possibilidade de integrar áudio e vídeo, como meios para se registrar informações, pode requerer formas para a especificação de sincronização temporal entre conteúdos de dados, para sua apropriada apresentação²⁵.

Geralmente, um *objeto multimídia*, resultado de processos de captura ou representação na presença de multimídia, pode ser um agregado altamente estruturado de componentes mais simples, como ocorre em um vídeo e que também podem ser considerados como objetos multimídia. Cada objeto multimídia simples pode ser visto como um pedaço de dados, cujo conteúdo está relacionado a qualquer um dentre os diferentes tipos de mídia.

Componentes multimídia baseados em tempo ainda podem conter pedaços de conteúdo derivados, que são definidos em função de outros conteúdos, e gerados somente quando necessário, provendo, segundo [22], um certo tipo de independência física de

(²²) A hipermídia chega a ser considerada a ciência do relacionamento, por [3], e, nessa perspectiva, refere-se à estruturação, apresentação e possibilidade de acesso direto ao conteúdo e a interconexões, dentro de um domínio de interesse, através de associações conceituais entre pedaços de dados.

(²³) Algumas referências de propostas de modelos são apresentadas no item 3, em seguida.

(²⁴) MHEG (Multimedia and Hypermedia information coding Expert Group) é um padrão em discussão para o registro codificado de objetos multimídia ou hipermídia em sua forma final, que serão trocados entre serviços e aplicações, em qualquer forma (meio de armazenamento, LAN, rede de telecomunicação, etc.).

(²⁵) A necessidade de sincronismo entre conteúdos com características de continuidade temporal pode ser visto em um filme, seu som e falas, e destes com as legendas de tradução.

dados, havendo, então, a necessidade da descrição dessa derivação.

Os componentes (digitais ou não²⁶) de um objeto multimídia podem se relacionar de diversas maneiras, sendo que alguns desses relacionamentos, além de envolverem componentes do próprio objeto multimídia, podem envolver componentes de outros objetos multimídia.

É possível que um determinado objeto multimídia dê origem a diferentes *objetos de exibição*, que correspondem rudemente ao que é exibido ao usuário, havendo a preservação de parte do conteúdo dos dados registrados, devido a fatores como: segurança, perfil e preferência do usuário, e recursos disponíveis, entre outros. Além disso, durante a autoria de um objeto multimídia, é permitida a criação de diversas visões do mesmo, denominadas *perspectivas*, cada uma das quais responsável pela origem de um objeto de exibição, que, por sua vez, pode necessitar sofrer adaptações provenientes do comportamento dinâmico do ambiente de transmissão e exibição.

Cada perspectiva deve permitir a identificação dos componentes que realmente serão (ou poderão ser) exibidos e seus relacionamentos, juntamente com a informação computacional sobre formato de codificação, duração e localização.

Um objeto de exibição é um todo, do ponto de vista de sua manipulação, deve ser lido, ouvido, visto, sentido e/ou consumido, podendo ocorrer a interação do usuário com o veículo exibidor. Esta interação ocasiona início de atividade computacional, que pode resultar em: mudança no objeto em exibição (da parte sendo exibida, da forma, velocidade ou qualidade de exibição); mudança do contexto em exibição (para um novo objeto ou para outro nível de representação²⁷); ou atualização de dados (modificação, inserção ou remoção).

Como pode ser percebido, um outro nível de modelo de dados é necessário, com vistas a facilitar o

processo de captura e registro de elementos de informação, considerando os requisitos de agregação de interpretação ao conteúdo, a existência de várias mídias e a possibilidade de acesso não seqüencial. Esse novo tipo de modelo de dados permite a obtenção de modelagens de dados em níveis de abstração mais altos, que apesar de poderem sofrer influência do distante computador, devem considerar vários fatores impactantes no resultado ou no processo de realização da modelagem, segundo [23], e apresentados a seguir. Os quatro primeiros deles diretamente relacionados ao domínio de interesse e os dois últimos diretamente geradores de interferência no (posterior) processo de desenvolvimento de programas ou BDs:

- formas de acesso aos dados,
- relacionamentos entre diferentes dados,
- disponibilização de diferentes perspectivas de visualização de dados,
- segurança de acesso aos dados,
- recuperação a partir do armazenamento e facilidades suportadas para acesso aos dados,
- reuso e facilidades para manutenção dos dados.

Um modelo de dados para hipermídia requer, além dos mecanismos para a especificação estrutural-semântica (tradicional), comportamental e de integridade, mais:

- modelos para percepção, que englobam modelagem de distribuição espacial, temporal, seletividade e apresentação²⁸, relativos à parte multimídia;
- modelo hiperestrutural ou de navegação²⁹, relativo à parte associativa;
- modelo para a parte comportamental e de integridade da interação homem-computador;

⁽²⁶⁾ Pode-se manipular a partir do computador conteúdos de dados registrados em forma análogica e armazenados externamente, havendo necessidade de algum tipo de integração entre o mecanismo externo e o computador para localização e disponibilização do conteúdo.

⁽²⁷⁾ Exemplo de mudança de contexto é a que ocorre quando o usuário solicita informações sobre o objeto cujo conteúdo encontra-se em exibição.

⁽²⁸⁾ Distribuição espacial ou como os componentes ocupam o espaço de exibição. Continuidade temporal ou como dois ou mais componentes necessitam ser exibidos como um todo ou dependem um do outro em relação ao tempo de exibição. Seletividade ou qual de dois ou mais componentes deve ser considerado para exibição. Apresentação que compreende aspectos diretamente vinculados à apresentação de conteúdo, tal como tamanho de símbolos, cor, contraste, tamanho de janela, etc.

⁽²⁹⁾ Mais do que apenas a especificação de uma rede de nós e ligações, é necessário, também, especificar como os dados são acessados e navegados, isto porque podem existir tipos específicos de ligações que implicam em comportamento específico (interno ou externo ao contexto em que o usuário se encontra no instante da navegação).

- incorporação de significado a todos os objetos resultantes das diversas modelagens, permitindo a agregação de asserções não relacionadas e não cobertas pelos modelos que descrevem estrutura, comportamento, percepção, interação ou navegação, provendo auxílio para a interpretação dos dados, caso esta seja necessária.

Além disso, dois outros conceitos relacionados a tempo podem necessitar ser especificados: o primeiro relativo à evolução, ou seja, estabelece a seqüência em que os dados sofrem mudanças ou são criados ou removidos; o segundo relativo à validade, que engloba o instante da evolução e tempo de validade (duração da utilidade dos dados).

3. EVOLUÇÃO DOS MODELOS DE DADOS

Modelos de dados são, desde seus primórdios, fortemente relacionados à atividade de projeto de BDs, pois o projeto de uma BD nada mais é do que a construção de um modelo de uma parte de um mundo real ou imaginário, de maneira que possa ser entendido por várias pessoas, os usuários da BD.

As tarefas de projeto e desenvolvimento de BDs, tradicionalmente desenvolvidas artesanalmente, muitas vezes de maneira inconsistente, ganham muito com a utilização de modelos de dados, que, se bem usados, permitem o aparecimento de um processo estruturado e racional, baseado em métodos bem definidos de projeto [46].

O registro é uma das estruturas de dados mais básicas que existem e, tradicionalmente, tem sido usado como a unidade lógica para armazenamento não volátil de conteúdo de dados, tanto em arquivos, quanto em BDs. Os campos, que um registro pode conter, podem ser definidos em termos de seu tipo, estabelecendo sua sintaxe e sua semântica, parcialmente relacionada ao seu significado na aplicação.

Como não poderia deixar de ser, originalmente, os modelos de dados sofreram o impacto da cultura do registro, ou seja, foram criados de maneira a estarem

sempre direcionados à forma como se registrava informação em computadores³⁰, para aplicações bancárias, folhas de pagamento, controle de inventários e similares. Surgiram, então, o que denominamos modelos de dados baseados em registros, dos quais salientamos três enfoques básicos e diferentes:

- o relacional [14],
- em rede [41] [13] e
- o hierárquico [43].

Esses modelos de dados, utilizam a estrutura de registros como unidade básica para representar e registrar conteúdo de dados, além disso, utilizam as abstrações da relação ou tabela³¹ e de relacionamentos.

Não se pode proclamar que, em suas conversações e processos de pensamento, as pessoas entendam informação apenas como tabelas ou grafos³². De fato, parece que a maneira natural de representação parece ser o diálogo ou o texto em linguagem natural. Além disso, foi rapidamente percebido que a organização inicial dos dados, por si, não determinava uma estrutura ótima para uma BD. Ao contrário, uma BD deveria ser projetada de acordo com a funcionalidade e uso que dela se desejava.

Assim sendo, nos últimos tempos têm havido uma preocupação de se criar modelos de dados orientados aos usuários, para que seja possível o entendimento de uma modelagem por leigos, através de, preferencialmente, meios naturais (convencionais) de comunicação. Uma modelagem de dados obtida através de modelos orientados a usuários é uma representação abstrata do domínio de interesse.

Uma conseqüência da existência de modelos orientados aos usuários é a necessidade de um mapeamento, primeiro dos aspectos do domínio de interesse para uma descrição fruto desse tipo de modelagem, inteligível aos usuários, denominado a parte *infológica* da modelagem. Em segundo lugar, dos conceitos infológicos básicos em representações computacionais, denominado a parte *datalógica* da modelagem [34].

(³⁰) O processo de registrar informação é tradicionalmente realizado de forma textual, em campos com características pré-estabelecidas, geralmente, tamanho pequeno e tipo registro.

(³¹) Uma relação representa uma categoria de entidades, onde cada tupla contém os dados de uma entidade específica e, em conjunto, podem ser representadas como uma tabela.

(³²) Os modelos de dados utilizam grafos para representar categorias (nós) e relacionamentos (arestas) e permitir uma visualização mais amigável ao usuário.

A parte infológica é totalmente independente dos mecanismos computacionais ou atributos de dispositivos físicos, que podem ser utilizados para a armazenagem dos dados, e é usada para orientar e documentar o processo de elucidação dos requisitos de informação, além de prover um meio de comunicação de alto nível destes requisitos a usuários e implementadores.

Os modelos baseados em registros enfatizam a parte datalógica da modelagem, principalmente devido a sua origem histórica. Enquanto os modelos de dados hierárquico e o em rede evoluíram a partir do processamento de arquivos e sistemas de geração de relatórios, o modelo de dados relacional, por outro lado, é baseado mais em fundamentos teóricos do que em experiência prática [44], mas, mesmo assim, considerando a então maneira como se realizavam projeto e implementação de aplicações.

Em seguida ao aparecimento dos modelos de dados baseados em registros, percebeu-se a necessidade de desenvolvimento de modelos de dados infológicos, que permitissem uma maior facilidade para se capturar e codificar significado no projeto, tanto de aplicações convencionais, quanto de outras novas aplicações³³.

Os aspectos semânticos relativos à organização estrutural dos dados é o foco do que pode ser considerado uma nova geração de modelos de dados, os modelos de dados semânticos. Nestes a evolução, em relação à geração anterior, passa pela possibilidade de uso e visão de várias abstrações estruturais estáticas, em níveis variados, da parte do domínio de interesse, que se quer representar.

Alguns modelos de dados semânticos interessantes são o Modelo Entidade-Relacionamento [11], conhecido por MER; o modelo suportado pela linguagem TAXIS [4]; o Modelo de Dados Semântico [25], SDM; o modelo RM/T [15], extensão do modelo de

dados relacional criada pelo próprio mentor do modelo relacional; e o Modelo de Eventos [32].

Apesar da distinção entre modelos de dados semânticos e orientados a objetos não estar bem definida, pois existe uma sinergia entre eles, a não existência ou a pouca ênfase encontrada em vários modelos de dados semânticos em relação a aspectos comportamentais, permite o estabelecimento de uma terceira geração de modelos de dados preocupados em estender os modelos de dados semânticos pela incorporação da perspectiva comportamental ou funcional à perspectiva estrutural, incorporando operações em categorias³⁴.

O suporte de mecanismos para a definição de novos tipos estruturais e comportamentais, tem sido utilizado como justificativa para a utilização de modelos de dados orientados a objetos para responder pelas necessidades impostas por aplicações nas áreas de CAD, CAM, CASE, OAS e MM³⁵, entre outras. Porém, tem sido notada a ausência de suporte direto a novas perspectivas abstrações para um eficiente e adequado uso desse tipo de modelo de dados para algumas dessas aplicações, como é o caso com a hipermídia. Esta falha de suporte parece indicar a possibilidade de surgimento de uma nova geração de modelos de dados.

O paradigma da orientação a objetos, oriundo das áreas da Inteligência Artificial e Linguagens de Programação, foi descoberto e incorporado por pesquisadores atuantes nas áreas de Banco de Dados [31] [9] e Engenharia de Software [30] [16], criando uma espécie de congruência entre a modelagem de dados, com ênfase estrutural, desenvolvida pelos primeiros, e a modelagem de processos, com ênfase funcional, desenvolvida pelos últimos.

As idiosincrasias da aplicação da hipermídia já recebem suporte de modelos de dados semânticos e orientados a objetos estendidos, alguns deles consi-

⁽³³⁾ As novas aplicações são aplicações não convencionais para as quais as estruturas de dados apresentam tamanhos grandes e organizações complexas, refletindo hierarquias ou outras organizações mais genéricas, e, portanto, deficientemente suportadas pelos modelos de dados com tendências datalógicas.

⁽³⁴⁾ A incorporação se realiza utilizando o que se convencionou chamar tipos abstratos de dados, na área de linguagens de programação.

⁽³⁵⁾ CAD - Computer Aided Design,
CAM - Computer Aided Manufacturing,
CASE - Computer Aided Software Engineering,
OAS - Office Automation System,
MM - MultiMedia.

derados modelos de referência formais por [40]. Exemplos interessantes desses modelos são o modelo Dexter [24], utilizado primariamente para comparação e avaliação de ferramentas para construção de sistemas com hipermídia; o modelo de Lange [35], que procura explorar diversas abstrações relacionadas com a navegação de hiperdocumentos; o modelo Amsterdam [26], preocupado em combinar o modelo Dexter com a modelagem de aspectos temporalmente contínuos; e o modelo Tower [5], que considera várias das perspectivas encontradas na hipermídia como níveis de descrição.

4. A HIERARQUIA DE MODELOS DE DADOS

A categorização dos modelos de dados em infológicos (preocupados com a representação abstrata³⁶) e datalógicos (preocupados com a representação computacional ou física) pode ocasionar a necessidade de um mapeamento das modelagens da primeira categoria em modelagens da segunda, quando se deseja a implementação de um sistema de informação. E esse mapeamento pode, ainda, ser realizado em etapas.

Um projetista de BD, durante a fase preliminar de projeto, tem por objetivo transformar as descrições de requisitos de informação e requisitos de processamento em uma ou mais visões, posteriormente integradas, utilizando-se de um modelo de dados abstrato. O modelo de dados utilizado é, geralmente, distinto do modelo de dados base para a descrição estrutural (e algumas vezes de integridade) da BD. A descrição da BD ocorre durante a fase de projeto esquemático, que segue o projeto abstrato, e consiste (pode consistir) de uma tradução entre modelos de dados. Seguem, ainda, as fases de projeto de implementação e projeto físico, onde utilizam-se ferramentas de um Sistema para gerência de Bases de Dados (SBD) para a especificação da implementação propriamente dita.

Uma proposta consagrada para descrição de uma hierarquia de modelos de dados, denominada

arquitetura ANSI/SPARC [42] [8], estabelece os seguintes três níveis para modelagem:

- o nível abstrato ou externo, que representa uma visão abstrata para os dados que serão armazenados, enfocando os elementos de informação que são relevantes aos usuários, evitando aspectos relacionados com armazenamento, estratégias de recuperação ou desempenho;
- o nível esquemático ou intermediário³⁷, que permite a descrição de conteúdo e estrutura, considerando ou não a existência de uma modelagem abstrata, e utiliza mecanismos como tipos dos dados, relacionamentos entre tipos e regras de consistência para os mesmos, a partir do ponto de vista do projetista da BD;
- o nível físico ou interno, no qual estão descritas as estruturas físicas da BD, considerando sua implementação e sua gerência, da maneira como são vistas pelo S.O., incluindo tamanho e layout de campos e registros, blocos, buffers, etc.

Modelos de dados de alto nível ou abstratos estão mais próximos do usuário e de como este percebe o mundo e determinam, entre outras coisas, a organização dos dados em categorias e o meio em que se realizará o registro de dados (texto, áudio, vídeo ou gráfico, por exemplo), permitindo a especificação de abstrações de uma imagem parcial de um mundo real ou imaginário, focadas nos elementos de informação considerados relevantes, que são mais fáceis de entender do que conceitos computacionais, permitindo a captura dos requisitos de informação de uma aplicação.

Modelos de dados de baixo nível ou físicos provêm conceitos para a descrição de detalhes de como os dados ficarão armazenados, podendo ser inicialmente obtidos a partir de mapeamentos do nível externo, através de um mapeamento em modelos de dados esquemáticos, que escondem, ainda, alguns detalhes de armazenamento, mas que, às vezes,

⁽³⁶⁾ O termo abstrato é aqui utilizado para se referir aos subprodutos das atividades de coleta, elucidação e especificação de requisitos, que em alguns autores (por exemplo [18]) aparece como conceitual, ou (por exemplo [12] ou [17]), como lógico.

O termo esquemático, por sua vez, é aqui utilizado para se refletir aos subprodutos das atividades de projeto (preliminar e detalhado), que (por exemplo em [17]) denomina-se conceitual e que também são associados ao termo lógico em outros autores (por exemplo [37]). Note-se, ainda, que alguns autores utilizam os dois termos de maneira intercambiável, hora com o nosso significado de abstrato (por exemplo [38]), hora com o nosso significado de esquemático (por exemplo [38]), hora com nosso significado de abstrato (por exemplo [27]).

⁽³⁷⁾ Os modelos baseados em registros, o relacional, o em rede e o hierárquico, são considerados modelos esquemáticos.

podem ser implementados de uma forma, mais ou menos, direta.

A arquitetura ANSI/SPARC permite a obtenção de dois tipos de independência de dados em relação a programas:

- a independência esquemática, que permite a mudança do esquema³⁸ intermediário, sem necessariamente envolver mudança dos esquemas externos ou das aplicações, e
- a independência física, que permite a mudança do esquema interno, sem necessariamente envolver mudança do esquema intermediário.

Além disso, outro tipo de independência de dados possível é o de diferentes usuários terem visões diferentes dos mesmos dados.

A hierarquia estabelecida pelos mapeamentos entre modelos de dados, na realidade, pode representar uma estratégia de desenvolvimento de uma BD³⁹ para um domínio de interesse.

5. CONCLUSÕES

Dados devem ser organizados para serem úteis. Modelos de dados permitem especificar como essa organização pode ser especificada e quais as perspectivas⁴⁰ que devem ser consideradas para isso.

Um modelo de dados é uma ferramenta, que permite uma interpretação e uma visão suficientemente poderosas, para algum entendimento sobre como os dados podem ser organizados, considerando o significado daquilo que, como descrições ou registros, parcialmente representam, em um mundo real ou imaginário.

Entende-se, em um primeiro momento, por modelo de dados a um conjunto de regras e construções, que podem ser usados para a descrição da organização estrutural abstrata, esquemática ou física dos

dados relativos a um domínio de interesse, incluindo-se a isso, a preocupação com a descrição das operações, que permitam a sua manipulação, e as restrições, para manter a sua integridade. Pode-se estender esta definição para qualquer aplicação que requeira "estruturação de dados"⁴¹.

Grande parte dos conceitos e entendimento envolvidos com modelos de dados é originária da época em que apareceram os modelos de dados baseados em registros. Assim⁴², os modelos de dados que se seguiram apresentam um arcabouço oriundo dos modelos pioneiros.

Por um lado, o papel de um modelo de dados, que originariamente era apenas o de especificar as propriedades de uma futura BD, evoluiu para servir como um meio de comunicação, uma notação para pessoas [44], utilizando conceitos abstratos. Apesar disso, a maneira pela qual a maioria dos modelos de dados em uso expressam requisitos de informação, geralmente, não corresponde às maneiras como as pessoas percebem a informação na aplicação, uma vez que esta esteja desenvolvida, pois utilizam-se construções artificiais. Esta situação ocorre devido à necessidade de se representar as construções do modelo de dados em termos de construções computacionais.

Por outro lado, os modelos de dados têm evoluído, principalmente, para atender as necessidades das aplicações que surgem em cada época, decorrentes do aparecimento de novas tecnologias. Desta maneira pode-se, a partir dos modelos de dados existentes e de suas épocas de criação, traçar um perfil das aplicações por eles atendidas.

Modelos de dados baseados em registros, responderam, e bem, às necessidades das aplicações tipicamente comerciais e bancárias surgidas entre as décadas de 50 a 70, mas, segundo [45], não provêm suporte para a descrição de todos os aspectos envolvidos com o domínio de interesse, em especial com operações e relacionamentos.

⁽³⁸⁾ O termo esquema está sendo usado para significar o conjunto de informações computacionais que permitem inferir a estrutura de uma BD (tipos de dados, relacionamentos e restrições), embora esse mesmo termo seja usado para denotar o resultado de uma modelagem abstrata para um domínio de interesse em particular.

⁽³⁹⁾ Estratégias similares, em termos de organização, são encontrados na engenharia de software e são conhecidas como metodologias de desenvolvimento "passo a passo" e "top down".

⁽⁴⁰⁾ A extensão e quantidade das perspectivas que devem ser consideradas são questões abertas e em contínua mudança.

⁽⁴¹⁾ Entenda-se aqui por estruturação de dados à modelagem estrutural dos aspectos estáticos do domínio de interesse, nos termos tradicionalmente usados para o desenvolvimento de aplicações em computadores.

⁽⁴²⁾ Considerando-se que na área da Ciência da Computação, é melhor evoluir do que revolucionar.

A falta de expressividade nos modelos de dados convencionais⁴³ levou à extensão e à integração de conceitos de outras áreas da computação a esses modelos. Mesmo assim, segundo [45], há, ainda, uma distância muito grande entre a semântica das aplicações e a sua representação em uma BD.

Modelos de dados semânticos foram criados para resolver o problema do entendimento limitado do significado dos dados em uma BD, através da agregação de maior poder semântico. Estes modelos sofreram influência de aplicações em inteligência artificial entre outras, mas não abrangem, nas suas origens, aspectos comportamentais, que são compreendidos nos modelos de dados orientados a objetos. Porém, ambos, em essência, não comportam mecanismos para o suporte de hipermídia, além de outros tipos de aplicação.

Como resultado dessa inadequação, modelos para estruturação de documentos⁴⁴ passaram a ser usados como paliativos, mesmo não suportando, de maneira adequada, multimídia e computação, em especial, na visão de [47], não permitindo que o usuário especifique e controle a semântica para navegação de documentos. Além disso, na visão de [7], esforço adicional é necessário para caracterizar a estrutura lógica de documentos multimídia.

Um novo modelo de dados tem que ser definido para o suporte de multimídia, porém sendo orientado a objetos não é suficiente. Modelos formais, que permitam a especificação semântica para dados multimídia, devem ser ricos em capacidades para abstração de informação e captura de semântica, além de providenciar uma forma canônica para a representação de imagens, cenas e eventos complexos, em termos de objetos e de seu comportamento espacial e temporal [21].

A abordagem ideal, para a próxima geração de modelos de dados, parece ser orientada a objetos e estrutural, que integre tanto técnicas comportamentais como estruturais. O termo "hipermodelagem orientada a objetos" (object-oriented hypermodeling) é utilizada por [33] para designar o novo tipo de modelo de dados necessário para suportar aspectos de hipermídia em sistemas de informação.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] A.V. Aho and J.D. Ullman
Foundations of Computer Science
Computer Science Press, 1992
ISBN 0.7167.8233.2
- [2] E. Bertino and L. Martino
Object-Oriented Database Systems: Concepts and Architectures
Addison- Wesley, 1994
ISBN 0.201.62439.7
- [3] M. Bieber and T. Isakowitz
Designing Hypermedia Applications
CACM, 38(8):26-29, aug/95
- [4] A. Borgida and J. Mylopoulos and H.K.T. Wong
Generalization/specialization as a basis for software specification
In On Conceptual Modelling, Perspectives from Artificial Intelligence, Databases, and Programming Languages
M.L. Brodie, J. Mylopoulos and J.W. Schmidt, Eds.
Springer-Verlag, 1984
- [5] P. Bra and G. Houben and Y. Kornatzky
An Extensible Data Model for Hyperdocuments
4th ACM Conference on Hypertext, dec/92, pp. 222-231
<http://www.win.tue.nl/~debra/echt92/final.ps>
- [6] A.W. Brown
From semantic data models to object orientation in design databases
Info. and Soft. Technology, 31(1):39-46, jan/feb/89
- [7] J. F. K. Buford
Multimedia File Systems and Information Models in Multimedia Systems, pp: 265-283
Addison-Wesley, 1994
ISBN 0.201.53258.1
- [8] T. Burns and E. Fong and E. Jefferson and R. Knox and L. Mark and C. Reedy and L. Reich and N. Roussopoulos and N. Truszowski
Reference Model for DBMS Standardization
Database Architecture Framework Task Group of the ANSI/X3/SPARC Database System Study Group
ACM Sigmod Rec., 15(1):19-58, 1986
- [9] R. G. G. Cattell
The Object Database Standard: ODMG-93
Morgan Kauffman, 1994
ISBN 1.55860.302.6

⁽⁴³⁾ Modelos baseados em registros são ditos convencionais pois são comumente utilizados para o projeto esquemático de BDs.

⁽⁴⁴⁾ Modelos para estruturação de documentos como o ODA [28] e o HyTime [36] podem ser considerados modelos de dados, porém não foram discutidos neste trabalho.

- [10] Committee Draft International Standard MHEG Part I — Base Notation (ASN.1), ISO/IEC CD13522-1, 1993
- [11] P. Chen
The Entity-Relationship Model: toward a unified view of data
ACM Trans. Database Sys., 1(1):9-36, mar/76
- [12] D.N. Chorafas
Intelligent Multimedia Databases
Prentice-Hall, 1994
ISBN 0.13.031188-X
- [13] Codasyl
Data Description Language Journal of Development Material Data Management Branch
Dep. of Supply and Services, Ottawa
- [14] E.F. Codd
A relational model of data for large shared data banks
CACM, 13, pp. 377-387
- [15] E.F. Codd
Extending the database relational model to capture more meaning
ACM Trans. Database Sys., 4(4):397-434, dec/79
- [16] D. Coleman and P. Arnold and S. Bodoff and C. Dollin and H. Gilchrist and F. Hayes and P. Jeremaes
Object-oriented development: the fusion method
Prentice-Hall, 1994
ISBN 0.13.338823.9
- [17] C.J. Date
An Introduction to Database Systems, vol. 2, 1983
Addison-Wesley
ISBN 85.7001.494.5
- [18] R. Elmasri and S.B. Navathe
Fundamentals of Database Systems
The Benjamin/Cummings Pub. Co., 1989
ISBN 0.8053.0145.3
- [19] P. K. Garg
Abstraction Mechanisms in Hypertext
CACM, 31(7):862-870, jul/88
- [20] F. Garzotto and P. Paolini and D. Schwabe
HDM - A Model-Based Approach to Hypertext Application Design
ACM ToIS, 11(1):1-26, jan/93
- [21] A. Ghafoor
Multimedia Database Management Systems
ACM Computing Surveys, 27(4):593-598, dec/95
- [22] S. Gibbs and C. Breiteneder and D. Tschirtz
Data Modeling of Time-Based Media
Proceedings of the 1994 ACM SIGMOD International Conference on Management of Data, pp:91-102, jun/94
ISBN 0-89791-639-5
- [23] A. Ginige and D.B. Lowe and J. Robertson
Hypermedia Authoring
IEEE Multimedia, winter, 1995, pp. 24-35
- [24] F. Halasz and M. Schwartz
The Dexter Hypertext Reference Model
CACM, 37(2):30-39, fev/94
- [25] M. Hammer and D. McLeod
Database description with SDM: A semantic data model
ACM Trans. Database Sys., 6(3):351-386, sep/81
- [26] L. Hardman and D. Bulterman and G. van Rossun
The Amsterdam Hypermedia Model: Adding Time and Context to the Dexter Model
CACM, 37(2):50-62, fev/94
- [27] W. Horak
Office Document Architecture and Office Document Interchange Formats: Current Status of International Standardization
Computer, 18(10):50-60, oct/85
- [28] R. Hunter and P. Kaijser and F. Nielsen
ODA: a document architecture for open systems
Comput. Commun., 12(2):69-79, apr/89
- [29] T. Isakowitz and E. A. Stohr and P. Balasubramanian
RMM: A Methodology for Structured Hypermedia Design
CACM, 38(8):34-44, aug/95
- [30] I. Jacobson
Object-Oriented Software Engineering: A Use Case Driven Approach
Addison-Wesley, 1992
ISBN 0.201.54435.0
- [31] W. Kim
Introduction to Object-Oriented Databases
The MIT-Press, 1990
ISBN 0.262.11124.1
- [32] R. King and D. McLeod
A unified model and methodology for conceptual database design
In On Conceptual Modelling, Perspectives from Artificial Intelligence, Databases, and Programming Languages
M.L. Brodie, J. Mylopoulos and J.W. Schmidt, Eds.
Springer-Verlag, 1984

- [33] P. Kuijpers
Object-Oriented Databases and Hypermedia
www.wis.cs.utwente.nl:8080/~kuijpers/hyper/dbhyp.htm
- [34] B. Langefors
Information Systems Theory
Information Systems, 2, pp. 207-219
- [35] D.B. Lange
A Formal Model for Hypertext
Proc. NIST Hypertext Standardization Workshop, fb/
90, pp. 145-166
- [36] S.R. Newcomb
Multimedia Interchange Using SGML/Hytlme
IEEE Multimedia, fall/95, pp.60-64
- [37] K. Parsaye and M. Chignell and D. Khoshafian and
H. Wong
Intelligent Databases
John Wiley and Sons, 1989
ISBN 0.471.50346.0
- [38] J. Peckham and F. Maryanski
Semantic Data Models
ACM Computing Surveys, 20(3):153-189, sep/88
- [39] N. Rishe
Database Design: The Semantic Modeling Approach
McGraw-Hill, 1992
ISBN 0.07.052955.8
- [40] J. L. Schnase and J. J. Leggett and D. L. Hicks and
R. L. Szabo
Semantic Data Modeling of Hypermedia Associations
ACM ToIS, 11(1):27-50, jan/93
- [41] R.W. Taylor and R.L. Frank
CODASYL data-base management systems
ACM Comp. Surv., 8(1):67-103, mar/76
- [42] D.C. Tsichritzis and A. Klug
The ANSI/X3/SPARC DBMS Framework: Report of the
Study Group on Data Base Management Systems
Information Systems, 3, 1978
- [43] D. C. Tsichritzis and F. H. Lochovsky
Hierarchical database management: a survey
ACM Comp. Surv., 8(1):105-123, mar/76
- [44] D. C. Tsichritzis and F. H. Lochovsky
Data Models
Prentice-Hall, 1982
ISBN 0.13.196428.3
- [45] G. Vossen
Data Models, Database Languages and Database
Management Systems
Addison-Wesley, 1991
ISBN 0.201.41604.2
- [46] G. Wiederhold
Database Design
McGraw Hill, 1983
- [47] P. T. Zellweger
Toward a Model for Active Multimedia Documents
in Blattner and Dannenberg
Multimedia Interface Design
ACM Press, 1992
ISBN 0.201.54981.6