

O uso da abordagem sistêmica na modelagem de banco de dados

Tarcísio Ferreira Grilo Júnior
tarcisiofgjunior@gmail.com
UNPBFPB

Ana Carolina Costa de Oliveira
carolyneoliveira@gmail.com
IFPB

Maria Silene Alexandre Leite
leite@ct.ufpb.br
UFPB

Adriana Zenaide Clericuzi
adrianaclericuzi@dce.ufpb.br
UFPB

Luiz Bueno da Silva
bueno@ct.ufpb.br
UFPB

Resumo:A proposta deste artigo é demonstrar que a complexidade existente nas tarefas que envolvem uma modelagem de dados em um banco de dados relacional. Neste sentido, busca-se evidenciar a estrutura interna do banco de dados, bem como seu funcionamento, considerando os aspectos que vão além dos conceitos que envolvem a estrutura do problema e/ou do estilo cognitivo do indivíduo que está analisando a problemática. Aspectos estes presentes também no mecanismo de gerenciamento de banco de dados. Internamente os bancos de dados dispõem de mecanismos de controle, objetivando garantir a consistência e integridade dos dados neles armazenados. Mostra-se como ocorrem as interações entre estes diversos mecanismos, com o objetivo de manter a integridade relacional. Esses mecanismos estão presentes na grande maioria dos bancos de dados, que representam grande fatia de mercado, entre eles pode-se destacar: Oracle, Microsoft SQL Server, IBM DB2, Postgres, Interbase, entre outros.

Palavras Chave: Banco de dados - Complexidade - Sistema Gerenciador - -



1. INTRODUÇÃO

O estudo da Ciência da Complexidade tem sido intensificado nos últimos anos, pois apesar de se tratar de uma ciência antiga, sua aplicação é recente. As pesquisas acerca da complexidade estão trazendo novas perspectivas para a aplicação da temática nos mais diversos setores, seja acadêmico, produtivo ou econômico. Toma-se como referencial a visão sistêmica no processo de modelagem de banco de dados, bem como a utilização dos sistemas gerenciadores de banco de dados (SGBD) para a realização dessa tarefa.

Quando se estuda modelagem de bancos de dados sob a lente da abordagem sistêmica, são identificados alguns problemas ligados principalmente aos aspectos cognitivos dos membros participantes desse processo (modelagem de bancos de dados), considerando também as restrições semânticas e as várias propriedades que podem surgir no caminho, como por exemplo, as suas interações e emergências.

Diversas restrições de integridade devem ser observadas quando se realiza um estudo para a modelagem de um banco de dados, de modo a preservar a semântica e a qualidade dos dados neles armazenados. Uma vez que essas restrições tenham sido definidas e documentadas como parte integrante da modelagem conceitual, pode-se fazer com que o SGBD e a aplicação venham a efetivamente implementá-las e, assim, garantir a integridade dos seus bancos de dados com o uso de regras de integridade relacional. Visando garantir que certos tipos de restrições impostas no modelo não sejam violadas, o que poderia causar inconsistências, tanto semânticas quanto físicas. Essas restrições precisam ser consideradas na construção do esquema da aplicação, e podem ser deduzidas a partir do modelo conceitual empregado, desde que este modelo seja capaz de explicitar de maneira adequada os relacionamentos entre objetos constantes desta modelagem.

Este artigo aborda o relacionamento que existe entre a ciência computacional e a abordagem sistêmica, procurando identificar suas correlações e está dividido em três seções: a primeira aborda os conceitos de sistemas computacionais, a segunda aborda os conceitos e regras de integridade nos bancos de dados, e por fim a terceira, que aborda alguns conceitos da abordagem sistêmica e como ela se relaciona com os sistemas computacionais e de banco de dados.

2 . ASPECTOS METODOLÓGICOS

Este artigo foi construído a partir de uma revisão de literatura fundamentada nos seguintes autores: AXELROD, R. M.; COHEN, M. D (2000); BAR-YAM, Yaneer (1997); ELMASRI & NAVATHE (2011); HEUSER (2008); JOHNS, Richard (2001); LEITE(2004); MACHADO (1995); MAXIMILIANO (2002); MONTEIRO (2004); MORIN (1977); SERMANN, LUCIA IZABEL CZERWONKA (2003); SILBERSCHATZ, Abraham; KORTH, Henri F, & SUDARSHAN S. (1999); SOMMERVILLE (2003); STAIR, R. M.; REYNOLDS, G. W (2006). Neste momento, a intenção do artigo é estabelecer uma relação entre a abordagem sistêmica e a modelagem de banco de dados, mostrando que existem pontos de convergência entres as teorias. Ressalta-se que não se pretende fazer nenhuma intervenção prática neste artigo.

3. SISTEMAS COMPUTADORIZADOS

Os sistemas baseados em computador estão presentes em grande parte das atividades do cotidiano das organizações, sejam elas de pequeno, médio ou grande porte. Contudo, esses



sistemas requerem grande processamento para armazenamento e recuperação de dados. Antes de abordar os conceitos que envolvem as regras sobre os sistemas de banco de dados, busque entender os conceitos para a computação do que é um sistema computacional. Segundo Sommerville (2003) “um sistema é uma coleção significativa de componentes inter-relacionados, que trabalham em conjunto para atingir algum objetivo”. Maximiano (2002) comenta que a palavra sistema é empregada para indicar um conjunto relativamente complexo de atividades ou subprojetos.

Sommerville (2003) afirma que uma característica dos sistemas computadorizados é que as propriedades e o comportamento dos componentes dos sistemas estão intrinsecamente interligados. O funcionamento bem-sucedido de cada componente do sistema depende do funcionamento de outros componentes. Os sistemas, frequentemente, são hierárquicos, no sentido de que eles incluem outros sistemas. Estes outros sistemas são chamados de subsistemas. Uma característica dos subsistemas é que eles podem operar como sistemas independentes.

As complexas relações entre os componentes em um sistema significam que ele é mais do que simplesmente a soma de suas partes. Ele tem propriedades que são do sistema como um todo. Essas propriedades emergentes não podem ser atribuídas a nenhuma parte específica do sistema. Pelo contrário, elas emergem somente quando o sistema como um todo é considerado. Algumas dessas propriedades podem ser derivadas diretamente das propriedades comparáveis de subsistemas, mas, com maior frequência, elas resultam de complexos inter-relacionamentos de subsistemas, que não podem, na prática, ser compreendidos pela análise de componentes individuais de sistemas (SOMMERVILLE, 2003). As propriedades emergentes muitas vezes são difíceis de serem previstas com antecedência. Elas podem ser medidas apenas depois que os subsistemas foram integrados para formar o sistema completo. Na computação há dois tipos de propriedades emergentes: as funcionais e as não funcionais.

As propriedades funcionais são aquelas que aparecem quando todas as partes de um sistema trabalham em conjunto para atingir um objetivo comum, por exemplo, uma bicicleta tem uma propriedade funcional de ser um dispositivo de transporte, uma vez que está montada com todos os seus componentes. Enquanto as não funcionais, como confiabilidade, desempenho, segurança e proteção. Essas propriedades se relacionam com o comportamento do sistema em seu ambiente operacional.

3. SISTEMAS DE BANCO DE DADOS

Segundo Monteiro (2004), o banco de dados é tão importante quanto o código do sistema. É um conjunto ou uma coleção de dados que estão de alguma forma relacionados e que são utilizados durante as operações da empresa para manter seus processos operacionais.

O dicionário Aurélio (1986), conceitua banco de dados como uma coleção abrangente, organizada e inter-relacionada de dados armazenados em meio físico, com o objetivo de evitar ou minimizar duplicidade de informação, otimizar a eficácia de seu tratamento, permitindo o acesso, através de diversas formas, a uma grande variedade de informações.

Normalmente um registro está associado a um conceito completo e é dividido em campos, ou atributos, que dão valores a propriedades desses conceitos. Possivelmente alguns registros podem apontar diretamente ou referenciar indiretamente outros registros, o que faz parte da caracterização do modelo adotado pelo banco de dados.

A descrição de quais são os tipos de registros existentes em um banco de dados e ainda quais são os campos de cada registro é conhecida como *esquema* do banco de dados ou esquema relacional, essa estrutura tem a forma de *tabelas*, que é composta por um conjunto



de linhas e colunas. Estritamente falando, o termo banco de dados deve ser aplicado apenas aos dados, enquanto o termo Sistema Gerenciador de Bancos de Dados (SGBD) deve ser aplicado ao software com a capacidade de manipular bancos de dados de forma geral. Porém, é comum misturar esses dois conceitos. Segundo Elmasri & Navathe (2011), vários critérios são normalmente utilizados para classificar os SGBDs, o primeiro deles é o modelo de dados no qual o SGBS é baseado. O principal modelo de dados usado atualmente em muitos SGBDs comerciais é o modelo de dados relacional. O modelo de dados de objeto foi implementado em alguns sistemas comerciais, mas não em seu uso generalizado.

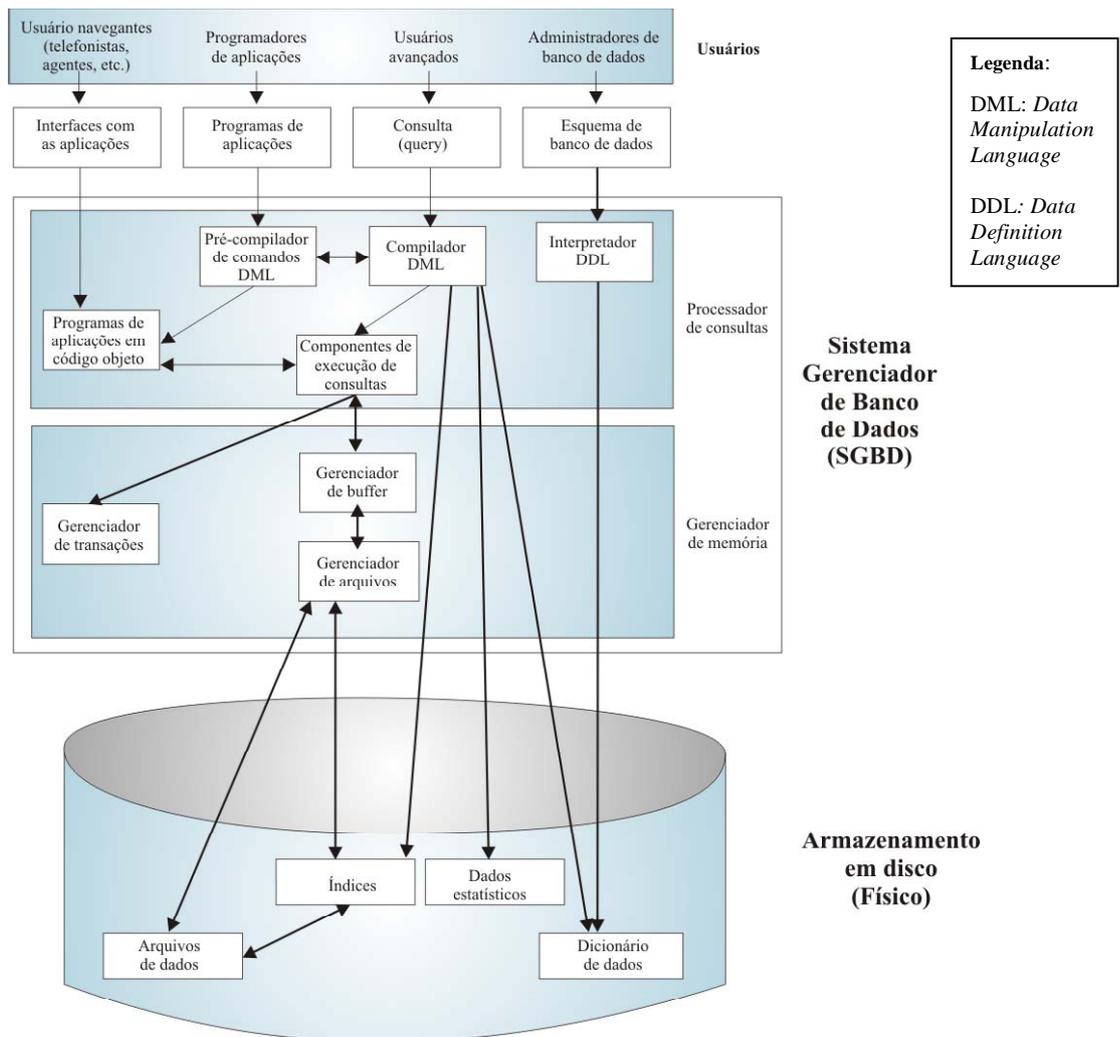


Figura 1 – Visão geral da estrutura do sistema de banco de dados.

Fonte (Silberschatz *et al.*, 1999, p. 18)

Stair (2006) define SGBD como “grupo de programas que manipulam a base de dados e proporcionam uma interface entre base de dados e o usuário da base de dados e outros programas de aplicação”. A meta básica de um SGBD é proporcionar um ambiente conveniente e eficiente para a recuperação e armazenamento de informações. Sistemas de banco de dados são projetados para gerir grandes volumes de informações. O gerenciamento de informações implica a definição de estruturas de armazenamento das informações e a definição dos mecanismos para a manipulação dessas informações. Ainda, um sistema de banco de dados deve garantir a segurança das informações armazenadas contra eventuais



problemas como o sistema, além de impedir tentativas de acesso não autorizadas (SILBERSCHATZ *et al.*, 1999). Como mostrado na Figura 1, a estrutura de um SGBD possui mecanismos para a manipulação, armazenamento e segurança das informações nele contidas, e segundo o mesmo autor, o objetivo principal de um sistema gerenciador de banco de dados é proporcionar ao usuário uma visão abstrata dos dados, que podem ser vistas em três níveis: nível físico, nível lógico e os níveis de visão.

O nível físico descreve como esses dados estão de fato armazenados; o nível lógico descreve quais dados estão armazenados e quais os inter-relacionamentos entre eles; e o nível de visão descreve apenas parte do banco de dados através de visões diversas do mesmo banco de dados para proporcionar interações simplificadas ao usuário. Silberschatz *et al.* (1999) descreve ainda que este esquema de banco de dados é determinado por um conjunto de definições que são expressas pela *linguagem de definição de dados* (DDL - *Data Manipulation Language*), e que seus comandos são compilados e armazenados em *tabelas armazenadas* em um arquivo especial chamado *dicionário de dados*. Para habilitar o acesso e manipulação desses dados pelo usuário é utilizada a *linguagem de manipulação de dados* (DDL - *Data Definition Language*).

Segundo Machado (1995), todo o projeto de banco de dados necessita de uma modelagem do sistema, que poderá ser através da abordagem Entidade Relacionamento (E-R) ou de outro tipo de abordagem, como por exemplo, a abordagem orientada a objetos, neste artigo tratará apenas da abordagem E-R, que representa o ponto central no projeto conceitual de um sistema de dados. Ele ainda enfatiza que a utilização correta de metodologia orientada a banco de dados envolve a estruturação em três níveis de visão de dados, ou seja, etapas na execução de um projeto: conceitual, lógico e físico, conforme mostrado na Figura 2.

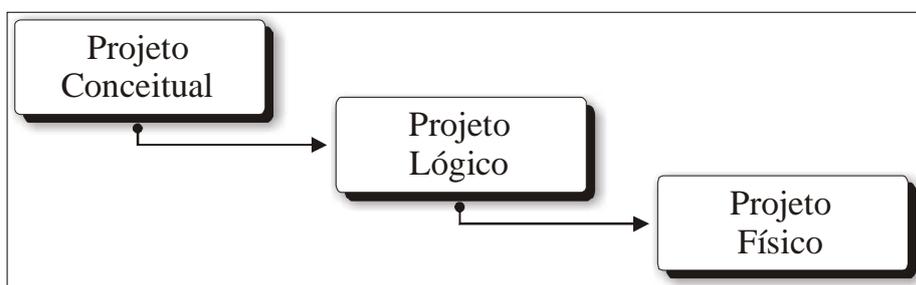


Figura 2 – Etapas na execução de um projeto.

Fonte Machado (1995, p. 29)

O modelo conceitual representa e/ou descreve a realidade do ambiente do problema, constituindo-se em uma visão global dos principais dados e relacionamentos (estruturas de informação), independente das restrições de implementação. O objetivo deste modelo é descrever as informações contidas em uma realidade, as quais irão estar armazenadas em um banco de dados.

O modelo lógico tem seu início a partir do modelo conceitual, os quatro modelos mais conhecidos são: hierárquico, em rede, relacional e orientado a objetos; existem também outros modelos, variando com o autor: o modelo relacional estendido e o modelo semi-estruturado. O modelo lógico descreve as estruturas que estão contidas no banco de dados, de acordo com as possibilidades permitidas pela abordagem, mas sem considerar, ainda, nenhuma característica específica de um SGBD, resultando em um esquema lógico de dados sob a ótica de uma das abordagens citadas.

O modelo físico irá a partir do modelo lógico descrever as estruturas físicas de armazenamento de dados, tais como: tamanho de campos, índices, tipo de preenchimento



destes campos, nomenclaturas, *etc.*, projetadas de acordo com os requisitos de processamento e uso mais econômico dos recursos computacionais.

A realidade dos negócios de uma empresa é sempre diferente da realidade de outra empresa, mesmo que se refiram a ambientes similares, existem particularidades que só dizem respeito ao funcionamento daquele ambiente específico, e devido a esta não similaridade entre ambientes de mesma natureza, será sempre necessária a criação de um modelo específico para cada nova realidade observada.

Para tanto, Silberschatz *et al.* (1999), afirma que o modelo Entidade Relacionamento (E-R) tem por base a percepção de que o mundo real é formado por um conjunto de objetos chamados *entidades* e pelo conjunto dos *relacionamentos* entre estes objetos. O modelo E-R é um dos modelos com maior capacidade semântica; os aspectos semânticos dos modelo E-R se referem à tentativa de representar o significado dos dados, sendo extremamente útil para mapear, sobre um esquema conceitual, o significado e interações das empresas reais. Uma entidade é uma “coisa” ou um “objeto” no mundo real que pode ser identificada de forma unívoca em relação a todos os outros objetos. Já um relacionamento é uma associação entre uma ou várias entidades.

Nos sistemas de banco de dados são utilizados mecanismos para permitir o controle e gerenciamento dos dados neles contidos, um deles é o conceito de *chaves*, que segundo Heuser (2008) serve para identificar linhas e estabelecer relações entre linhas de tabelas de um banco de dados relacional. Há pelo menos três tipos de chaves a considerar: a chave primária, a chave alternativa, e a chave estrangeira.

Uma chave primária é uma coluna ou combinação de colunas cujos valores distinguem uma linha das outras dentro de uma tabela. Ao definir uma chave primária não se está definindo apenas uma restrição de integridade, isto é uma regra que deve ser obedecida em todos os estados válidos de um banco de dados. No caso da chave primária, a regra é a de unicidade de valores nas colunas que compõem a chave.

A chave é uma coluna ou uma combinação de colunas, cujos valores aparecem necessariamente na chave primária de uma tabela. A chave estrangeira é o mecanismo que permite a implementação de relacionamentos em um banco de dados relacional. E a chave alternativa é uma coluna ou combinação de colunas que podem servir para distinguir uma linha das demais.

Além da chave primária, o banco de dados possui índices, que são mecanismos de ordenação dos dados objetivando otimizar a performance no acesso as informações do sistema. Segundo Silberschatz *et al.* (1999), cada estrutura de índice pode ser associada a uma chave de procura em particular e um arquivo pode ter diversos índices com diferentes chaves de procura. Pode-se generalizar o conceito de índices como sendo um mecanismo que deve ser utilizado para localizar informações, de forma otimizada, em uma estrutura de banco de dados.

Heuser (2008) afirma que um dos objetivos primordiais de um SGBD é a integridade de dados. Para tentar garantir esta regra é oferecido um mecanismo de restrição de integridade, que nada mais é do que uma regra de consistência de dados que é garantida pelo próprio SGBD. Ainda segundo este mesmo autor, no caso da abordagem relacional, costuma-se classificar as restrições de integridade nas seguintes categorias: integridade de domínio, integridade de vazio, integridade de chave, e integridade referencial.

A integridade de domínio é uma restrição que especifica que o valor de um campo deve obedecer a definição de valores admitidas para a coluna; a integridade de vazio é uma restrição onde é especificado se um campo de uma coluna pode ou não ser vazio (se a coluna é de preenchimento obrigatório ou não); a integridade de chave é quando se define que os



valores da chave primária e alternativa devem ser únicos; e a integridade referencial é o tipo de integridade que define que os valores dos campos que aparecem em uma chave estrangeira devem parecer na chave primária da tabela referenciada.

Apesar das restrições impostas pelo SGBD para garantir a integridade dos dados nele armazenados. Há muitas outras restrições de integridade que não se encaixam em nenhuma das categorias citadas anteriormente. Essas restrições são chamadas de restrições semânticas. Um exemplo pode ser: Um empregado do departamento denominado “Finanças” não pode ter uma categoria funcional “Engenheiro”; ou o salário de um empregado não pode exceder o salário de seu supervisor. Essas restrições semânticas envolvem muitas variáveis, o que torna a tarefa de modelagem de dados bastante complexa e difícil de ser padronizada.

4. ABORDAGEM SISTÊMICA

Do ponto de vista da abordagem sistêmica, o conceito de sistema computadorizado está bem relacionado com o conceito de sistemas complexos, desta forma faz-se necessário algumas definições sobre o termo *sistema* na ótica da complexidade. Segundo Morin (1977), “um sistema é uma unidade global organizada de inter-relações entre elementos, ações e indivíduos”, ou seja, uma inter-relação de elementos constituindo uma entidade ou unidade global. Um sistema é um conjunto de unidades em inter-relações mútuas.

Segundo Leite (2004), “a lógica de funcionamento dos sistemas complexos é associar elementos diferentes, que pela interação, se organizam e evoluem para níveis mais ou menos complexos. Os sistemas complexos assumem uma modelagem sistêmica, onde partes e todo interagem constantemente. As interações fazem surgir a dinâmica do sistema. Da interação pode surgir um comportamento imprevisível e da interação emerge um todo organizado”.

O dicionário Aurélio (1986), conceitua subsistema com sendo um conjunto de partes inter-relacionadas integrante de um sistema mais amplo.

Portanto podemos considerar que cada sistema por sua vez pode ser o subsistema de um todo ainda maior, remetendo a disposição das partes em relação ao todo, que segundo Morin (1977), “*o todo é mais do que a soma das partes*”. A idéia de sistema remete à unidade complexa do todo *inter-relacionado*, suas características e propriedades fenomenais (LEITE, 2004).

Os mecanismos de banco de dados são compostos por subsistemas que interagem entre si, compondo o todo, como pôde ser observado na Figura 1, possuem também vários níveis de interação, demonstrado nos conceitos de *chaves* e *relacionamentos*. As chaves são as responsáveis por manter uma estrutura única dentro de cada tabela de dados. É através dos relacionamentos entre as tabelas que se buscam garantir a integridade dos dados nelas contidos, mas, mesmo com todos estes mecanismos presentes nos SGBD's, podem surgir emergências oriundas das restrições semânticas, que se mostram difíceis de serem catalogadas e/ou padronizadas. Através da abordagem Entidade Relacionamento (E-R) fica fácil de identificar os níveis hierárquicos existentes no processo de modelagem em um banco de dados. A Figura 3 mostra estes diversos níveis no modelo apresentado.

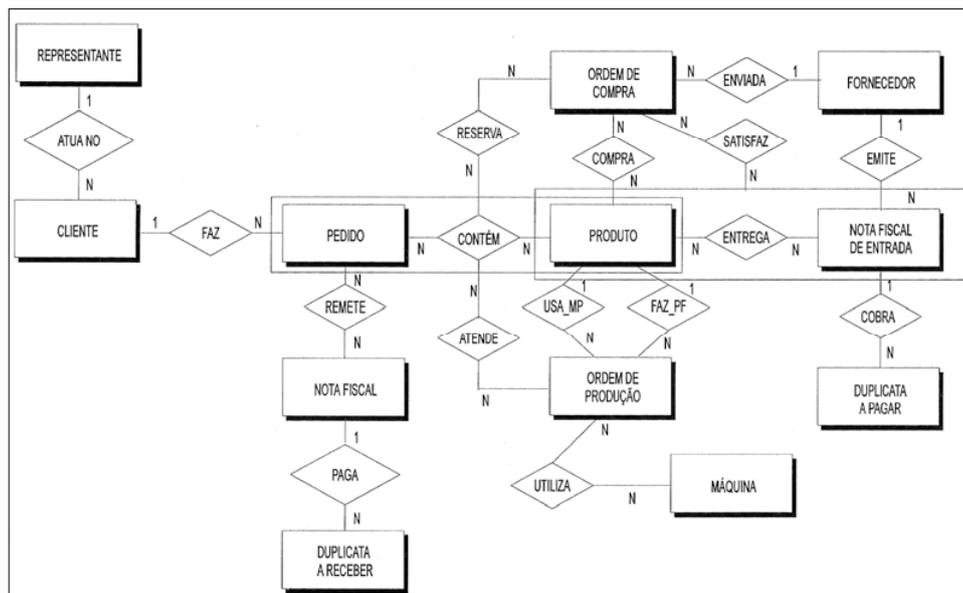


Figura 3 – Representação de um modelo E-R.

Fonte Machado (1995, p. 33)

Quando se faz uma análise para implementação de um banco de dados em uma empresa, os analistas de sistemas encontram dificuldades em criar um modelo genérico, que possa ser reaproveitado em modelagens futuras. Estes modelos diferem um dos outros de acordo com o estilo cognitivo de cada analista, pois cada um tem uma formação profissional ou acadêmica diferente, sem contar também com os níveis de interação presentes na dinâmica existente entre os membros integrantes de uma mesma equipe ou mesmo com seus entrevistados (clientes ou potenciais clientes).

Para Axerold & Cohen (2000), a “complexidade” não significa, simplesmente, muitas partes em movimento. O conceito desse termo pressupõe que o sistema consiste de partes que interagem, influenciando fortemente as probabilidades de ocorrência de eventos futuros. Desta forma as interações entre as diversas tabelas contidas em um banco de dados e suas restrições constituí um sistema complexo. Johns (2001) afirma que a complexidade de um objeto é a quantidade de informação necessária para especificá-lo.

Bar-Yam (1997) entende que um sistema complexo é um sistema formado por muitos componentes cujo comportamento é emergente, ou seja, o comportamento do sistema não pode ser simplesmente deduzido do comportamento dos seus componentes. A quantidade de informações necessárias para descrever o comportamento de um sistema deste tipo é uma medida de sua complexidade, e que desta forma a complexidade é a quantidade de informações necessárias para descrever um sistema.

Morin (1977) aponta que uma das características de um sistema complexo são as emergências que surgem e são decorrentes das inter-relações entre as partes que o compõem.

Como as emergências são as qualidades ou propriedades de um sistema, que se apresentam como novidade, no processo de modelagem de um banco de dados, muitos aspectos fogem ao conhecimento de quem está modelando, e podem não ser incluídos no modelo proposto pela equipe que analisam a problemática. Podem surgir, no decorrer do tempo de vida do sistema, várias emergências que não puderam ser identificadas nessa fase inicial, mas que devem ser solucionadas quando emergirem durante o tempo de vida, no qual o sistema está em uso.



5. CONSIDERAÇÕES FINAIS

Considerando o objetivo do presente artigo, a exposição dos fundamentos teóricos que abordam o processo de armazenamento de dados através de sistemas computadorizados, foi possível identificar que a abordagem sistêmica está presente em todas as etapas de criação, modelagem e armazenamento de dados em um SGBD, especificamente com relação as emergências e as interações.

Ressaltando ao mesmo tempo a influência do ambiente externo sobre os processos de modelagem do banco de dados e da importância da utilização dos conhecimentos baseados na abordagem sistêmica para elaborar uma estrutura de armazenamento de dados que seja adequada e duradoura para as organizações estudadas, e que possa ser capaz de emitir sinais de alerta quando algo estiver desviando a organização dos seus reais objetivos.

Desta forma, se a abordagem sistêmica for considerada durante o processo de modelagem de um banco de dados, os analistas envolvidos nessa tarefa, podem visualizar suas diversas interações e restrições, inclusive as restrições semânticas, por esta nova abordagem, que vem se mostrando um instrumento eficaz na identificação e solução para as problemáticas existentes dentro das organizações.

6. REFERÊNCIAS BIBLIOGRÁFICAS

AXELROD, R. M.; COHEN, M. D. Harnessing Complexity: organizational implications of a scientific frontier. *New York: The Free Press, 2000.*

BAR-YAM, Yaneer. Dynamics of Complex Systems. Westview Press. USA. 1997.

ELMASRI, Ramez & NAVATHE, Shamkant B. Sistemas de Banco de Dados. 6ª edição, São Paulo-SP: Pearson Addison Wesley, 2011.

FERREIRA, Aurélio Buarque de H. “Novo Dicionário da Língua Portuguesa”, 2ª edição, Rio de Janeiro: Nova Fronteira, 1996.

HEUSER, Carlos Alberto. Projeto de banco de dados. Porto Alegre: Artmed, 2008.

JOHNS, Richard. Dynamical Complexity and Regularity. University of British Columbia. 2001.

LEITE, Maria Silene Alexandre. Proposta de uma modelagem de referência para representar sistemas complexos. Tese (Doutorado) – Universidade Federal de Santa Catarina. 2004.

MACHADO, Felipe Nery R. & ABREU, Maurício. Projeto de Banco de Dados. Editora Érica, São Paulo, 1995.

MAXIMILIANO, Antonio Cesar Amaru. Administração de projetos: como transformar idéias em resultados. 2ª ed. – São Paulo: Atlas, 2002.

MONTEIRO, Emiliano Soares. Projeto de sistemas e bancos de dados. Rio de Janeiro: Brasport, 2004.

MORIN, Edgar. O método 1: da natureza da natureza; Tradução: Ilana Heineberg. 2ª edição. Porto Alegre: v. 1, Sulina, 1977.

SERMANN, LUCIA IZABEL CZERWONKA. Da estabilidade racional à complexidade substantiva: a busca da estabilidade em uma universidade. Tese (Doutorado). Universidade Federal de Santa Catarina. 2003

SILBERSCHATZ, Abraham; KORTH, Henri F, & SUDARSHAN S. Sistemas de Banco de Dados. São Paulo: MAKRON BOOKS, 1999.

SOMMERVILLE, Ian. Engenharia de Software. 6ª edição, tradução: André Maurício de Andrade Ribeiro; São Paulo: Addison Wesley, 2003.

STAIR, R. M.; REYNOLDS, G. W. Princípios de Sistemas de Informação. São Paulo: Pioneira, 2006.