

Universidade Tecnológica Federal do Paraná - Campus CT

Aluna: Amanda Jury Nakamura RA2582686 Bacharelado em Sistemas de Informação

Disciplina: Introdução a Banco de Dados Profº. Leandro Batista

Relatório - Modelagem de Bancos de Dados: Organização de artigos científicos

Através do enunciado, um primeiro passo foi a definição das entidades principais necessárias para obter um banco de dados com as funcionalidades especificadas.

As entidades principais foram então:

- 1) Autor (ID, nome, email e afiliação), com os dados dos pesquisadores que escreveram cada artigo;
- 2) Artigo (ID, título, resumo, número de páginas, ano de publicação, tipo), com os dados dos trabalhos científicos;
- 3) Keyword (ID, e a palavra chave em si), que armazena as palavra-chaves de cada artigo para ajudar a identificá-los;
- 4) Coleção (ID, nome, descrição), com dados e características em comum para agrupar artigos.

A partir disso, foi necessário definir como essas entidades se relacionam entre si. Os tipos de relacionamentos estudados foram: Um para um, um para muitos e muitos para muitos. Como forma de revisar cada um deles:

O primeiro (1:1), trata-se de um relacionamento em que uma tabela se relaciona com apenas um registro de uma outra. No segundo (1:N), um registro se relaciona com vários de uma outra tabela, mas o contrário não ocorre, sendo necessário que cada registro da segunda tabela se relacione com apenas um da primeira tabela. No último caso (N:M), vários registros de uma tabela se relacionam com vários de outra e vice-versa.

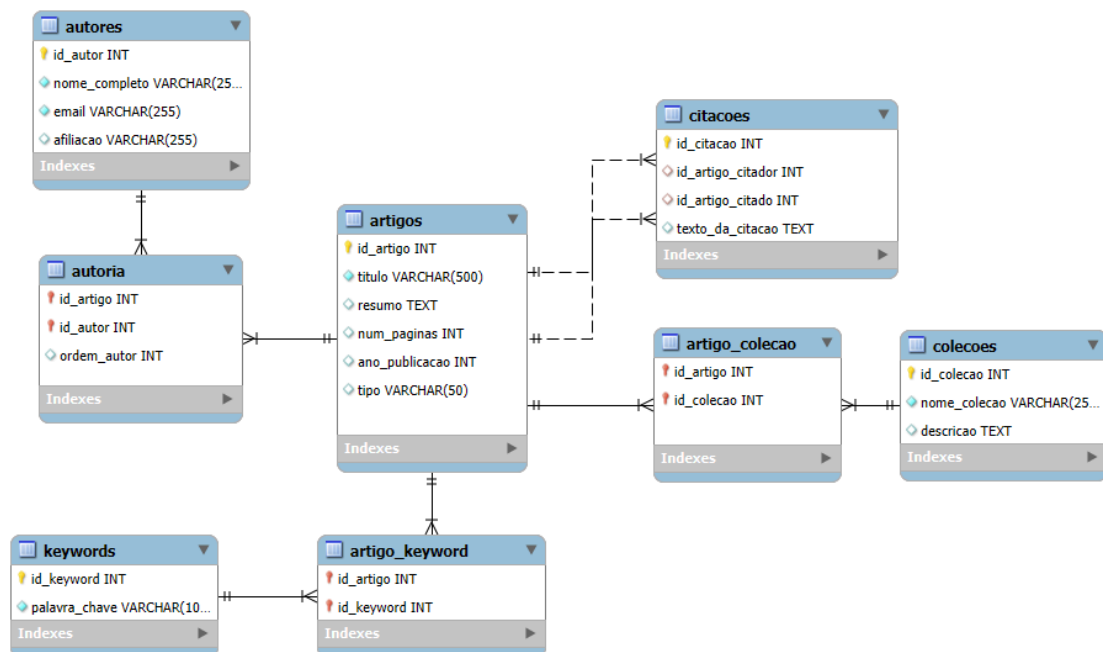
Neste trabalho, as relações que defini e as novas tabelas geradas foram:

- Artigo~Autores, relacionamento N:M
 - Um artigo pode ter vários autores, e um autor pode escrever vários artigos.
 - tabela Autoria - ID do autor(FK), ID do artigo(FK)
- Artigo~Keyword, relacionamento N:M
 - Um artigo tem várias keywords, e cada keyword pode aparecer em mais de um artigo.
 - tabela artigo_keyword - ID do artigo(FK) e ID da keyword(FK)
- Artigo~Artigo, relacionamento N:M
 - Um artigo pode citar vários outros e também pode ser citado em vários artigos.
 - tabela Citação - ID da citação(PK), ID do artigo citador(FK), ID do artigo que é citado(FK), texto da citação em si.

- Artigo~Coleção, relacionamento N:M
 - Um artigo pode estar em várias coleções, e cada coleção pode ter vários artigos.
 - tabela artigo_coleção - ID do artigo(FK) e ID da coleção (FK)

Tendo tudo isso definido, criei o diagrama fazendo cada tabela separadamente, e ao final apliquei o Forward Engineer e assim obtive a modelagem lógica e física do banco de dados gerado. Junto ao envio deste relatório, está o arquivo SQL text file.

Segue a figura exportada do projeto no MySQL Workbench:



Além disso, o enunciado também solicita uma análise do banco de dados gerado sob a perspectiva das Formas Normais.

- Em relação a Primeira Forma Normal (1FN):
 - A regra de que cada tabela deve conter um valor atômico (único) é seguida. Isso é visualizado na tabela 'autores', por exemplo, que tem colunas separadas para 'nome_completo', 'email' e 'afiliacao'. Além disso, não há uma coluna 'dados_autor' que armazena tudo junto. A tabela 'keywords' armazena uma única palavra-chave por linha (`id_keyword`), pois existe uma tabela formada pela junção, 'artigo_keyword'.
 - Em relação à regra das chaves primárias, todas as tabelas possuem uma chave primária claramente definida, por exemplo, todas as tabelas de entidade

principais como ‘autores’, ‘artigos’, ‘colecoes’ e ‘keywords’ usam têm a PK dada como o título somado de “ID”.

- Em relação à Segunda Forma Normal (2FN):
 - As tabelas ‘autores’, ‘artigos’, ‘colecoes’ e ‘keywords’ tem apenas uma coluna, por isso não há "parte da chave" da qual um atributo possa depender, dessa forma a 2FN é satisfeita
 - Para as demais tabelas:
 - Na tabela ‘autoria’, ‘ordem_autor’ não depende apenas do id_artigo nem apenas do id_autor, portanto, depende da chave completa.
 - Na tabela ‘citações’, o atributo não-chave ‘texto_da_citacao’ depende de qual artigo está citando qual outro artigo, ou seja, depende da relação completa entre os dois, satisfazendo a 2FN.
- Em relação à Terceira Forma Normal (3FN):
 - A 3FN diz que não deve haver dependências transitivas. Um atributo não-chave não pode depender de outro atributo não-chave. Na tabela ‘autores’, Os atributos ‘nome_completo’, ‘email’ e ‘afiliacao’ dependem apenas do id_autor; na tabela ‘artigos’, tudo depende do ‘id_artigo’; e em ‘coleções’, ‘nome_colecao’ e ‘descricao’ dependem diretamente do ‘id_colecao’. Dessa forma, está tudo conforme a 3FN.