

Modelo Relacional para Base de Dados Científica com uso em Biblioteca Acadêmica

Jader G. T. L. Pedro, Kesley R. G. Alves, Marcio J. Romero, Luis A. da Silva

Curso de Tecnologia em Banco de Dados - Faculdade de Tecnologia de Bauru (FATEC)

Rua Manoel Bento da Cruz, nº 30 Quadra 3 - Centro - 17.015-171 - Bauru, SP - Brasil

jader.gabriel@gmail.com, {kesley_10, marcinho.r}@hotmail.com,
luis.silva51@fatec.sp.gov.br

***Abstract.** This article has the objective to introduce a relational model for an academic database for supporting cataloging graduation papers of the students, as to make available all this content over the Internet, making easy the access to the contents of the database remotely at a low cost. As a differential, this paper implementates the use of binary fields for data storage for the files, directly on the database, without the need of a third party tool, in addition to presenting a simplified, but consistent model that meets the basic need of an academic system for graduation papers.*

***Resumo.** Este artigo tem como objetivo apresentar um modelo relacional para banco de dados acadêmico dando suporte ao sistema para catalogar os trabalhos de graduação produzidos pelos alunos da faculdade, bem como disponibilizar seu conteúdo produzido através da internet, facilitando assim, o acesso aos documentos de forma remota a um baixo custo de armazenamento. Como diferencial o trabalho possui a implantação ao suporte do uso de campos binários para armazenamento dos arquivos diretamente no banco de dados sem a necessidade de uma ferramenta de apoio, além de apresentar um modelo simplificado, mas consistente que atenda a necessidade básica de um sistema acadêmico baseado em artigos de trabalho de graduação.*

1. Introdução

Devido ao crescimento da demanda de material acadêmico produzido por alunos de universidades e a complexidade de se fornecer acesso a estes documentos sob demanda, enxergou-se a dificuldade de manter estes conteúdos alocados apenas na forma física, no qual se tem um alto custo de armazenamento e uma disponibilidade limitada [Santos Neto e Santos 2013]. A cada semestre é produzida uma gama de novos artigos e trabalhos de graduação, aumentando significativamente o acervo acadêmico das instituições de ensino do país, observou-se no período de 2002 a 2007 um crescimento de 49% na produção de artigos científicos [Muccioli et al. 2007]. Atualmente, os materiais de pesquisas produzidos ficam disponibilizados em bibliotecas físicas para consultas, gerando gastos para mantê-los em bom estado de conservação e disponíveis aos leitores.

A respeito do bibliotecário mediante as novas tecnologias de informação e comunicação, Silva (2014) explica que a expansão tecnológica tornou possível

disponibilizar estes conteúdos de forma ágil, simultânea e distribuída, facilitando o acesso aos dados produzidos por todos os alunos e demais interessados nos temas abordados. Dessa forma, a implantação de um banco de dados (BD) auxilia na organização de forma a possibilitar que todo o conteúdo contido nele seja facilmente transformado em informações precisas, garantindo maior agilidade nas pesquisas e apresentações de materiais acadêmicos aos usuários finais, proporcionando a expansão e disseminação do conhecimento produzido dentro do ambiente das instituições de ensino, tanto entre o polo educacional quanto outros polos, contribuindo para a propagação de conhecimento e de avanços tecnológicos [Morales e Rossetti 2007].

Atualmente já é adotada por algumas instituições de ensino superior, a prática de disponibilizar a produção acadêmica para acesso online, como é o caso da Universidade Estadual Paulista “JÚLIO DE MESQUITA FILHO” (UNESP), que utiliza um sistema denominado “Athena – banco de dados bibliográficos da UNESP”, ou a Universidade de Campinas (UNICAMP), que utiliza o Sistema de Bibliotecas Acadêmicas (SBU) como plataforma de acesso de todo seu material acadêmico.

Este trabalho tem como objetivo apresentar um modelo relacional para banco de dados acadêmicos dando suporte ao sistema para catalogar todos os trabalhos dos alunos da Faculdade de Tecnologia de Bauru (FATEC-Bauru), bem como disponibilizar o conteúdo produzido pelos próprios estudantes através da internet, facilitando assim, o acesso a estes documentos de forma remota a um baixo custo de armazenamento, e ainda, ampliando a acessibilidade através de plataformas heterogêneas, tais como: dispositivos móveis, *notebooks*, *desktops*, entre outros. O trabalho tem os seguintes diferenciais:

- a) A implantação ao suporte do uso de campos binários para armazenamento dos arquivos diretamente no banco de dados;
- b) Apresentação de um modelo simplificado, mas consistente que atende a necessidade básica de um modelo acadêmico baseado em artigos de trabalho de graduação.

O presente trabalho está dividido como segue: na Seção 2 são abordados os referenciais teóricos importantes no desenvolvimento deste trabalho. A Seção 3 apresenta os materiais e métodos que foram utilizados no desenvolvimento dos experimentos. Na Seção 4 é apresentado o uso do modelo de dados criado através de um sistema, a fim de comprovar a efetividade da proposta. E por fim, na Seção 5 são apresentadas as conclusões finais.

2. Evolução da Informação

Inegavelmente a escrita foi uma das maiores invenções da humanidade. A capacidade de produzir informações para disseminar o conhecimento e até mesmo de se arquivar o conteúdo produzido, garantiu que o desenvolvimento de todos os segmentos se expandisse rapidamente, sem a necessidade de reter o conhecimento em um único indivíduo. Há milhares de anos, os egípcios já produziam escrituras, bem como sumérios, assírios e babilônicos também descreviam sua realidade em placas de argila, com riquezas de detalhes. A partir desse contexto, é possível entender a ideia de Martins (2001), quando este afirma que a existência de bibliotecas se materializou antes dos livros e até mesmo dos manuscritos. A palavra biblioteca, de origem grega, pode ser traduzida como depósito de livros. Em seu início, ela foi concebida apenas como um

armazém de pequenas ou grandiosas quantidades de livros, placas de argilas, papiros ou pergaminhos, e sua utilização era limitada a poucas pessoas, visto que naquela época muitos não tinham o domínio da língua e da interpretação de textos, sendo somente reconhecida séculos depois como disseminadora da informação. Em meados do século XIX, o conceito de livro eletrônico era tratado apenas em obras de ficção científica. Até mesmo nos anos 80, ainda parecia distante que o computador e a internet se popularizassem, facilitando a leitura e a comunicação. O texto eletrônico era visto na época como um suporte que mantinha certa distância do leitor, porém ao mesmo tempo, já considerado universal, permitindo que vários leitores tivessem acesso ao conteúdo disponibilizado simultaneamente [Nascimento, Pinto e Vale 2013].

Com o surgimento de novas ferramentas tecnológicas e a própria comunicação, houve um aumento significativo na oferta de diversos serviços da informação possibilitando a estruturação e diferentes tipos de acesso aos repositórios informacionais. Os resultados dessas combinações originaram novos conceitos tais como: os de bibliotecas virtuais, bibliotecas digitais e banco de dados científicos [Marchiori 1997]. A sociedade e a economia estão se tornando cada vez mais baseados em informações. A queda acentuada dos custos relacionados a *hardware* e *software* auxiliou e modificou este processo, sendo assim a tecnologia se tornou um acelerador de mudanças, criando novas necessidades e tendências. Desta forma, foram se alterando sólidos paradigmas estabelecidos ao longo de muitos séculos, sendo o maior impacto desta transição o fato da informação se tornar cada vez menos ligada ao objeto físico que a contém, abrindo a possibilidade de encaixar o seu conteúdo em diversas áreas, minimizando assim os seus pontos negativos e atingindo um maior número de pessoas ao longo do tempo [Levacov 1997].

Blattmann (2001) afirmou que a função da biblioteca acadêmica é reter e ao mesmo tempo disponibilizar a informação, transcendendo o conceito de não ser mais valorizada pela informação que possui, como no passado, mas ser reconhecida pela informação que é capaz de disponibilizar para um maior número de usuários que podem estar localizados em diferentes áreas geográficas. A disponibilização deste serviço via internet permite atender a demanda de forma instantânea, mas devem existir processos e respectiva organização da informação que garanta a qualidade dos produtos e serviços *online*.

2.1 Informação Digital

No mundo da informação digital é possível modelar a informação, seja ela provida de livros ou até mesmo de outros documentos eletrônicos. Ao disponibilizar alguma informação na rede, o texto ganha proporções inimagináveis como a de sua propagação e a possibilidade de atingir vários territórios [Blattmann 2001]. Segundo Negroponte (1995) a cópia digital em alguns casos é tão perfeita ou melhor do que a original. O remanuseio da informação se torna uma tarefa bem mais rápida e fácil, se comparada com o momento anterior no qual o reprocessamento da informação em papel tende a ser mais lento. Já, Rodrigues (1995) considera que a possibilidade do acesso remoto à informação, possibilita a manipulação e edição de seu conteúdo e aumenta, pelo menos, em hipótese, o poder de escolha de seus utilizadores. Neste consenso, a internet pode ser considerada como uma valiosa fonte de informação para o uso de bibliotecários e pesquisadores.

Através da interface de rede é possível criar, alterar e até mesmo excluir sites inteiros na Internet, estabelecendo políticas sobre o acesso, desenvolvimento de coleções, armazenamento e a própria distribuição da informação digital dentro e fora das organizações. Essas questões, provocam constantes debates sobre o gerenciamento da informação digital e possibilitam inovações na gestão da tecnologia da informação nas bibliotecas. Com a migração de tecnologias ocorrida em razão da sua constante evolução, novas tendências são criadas durante seu uso e treinamentos são elaborados para satisfazer as necessidades dos usuários [Blattmann, Fachin e Rados 2000]. Ainda de acordo com os autores, o tratamento da informação gerada, bem como a forma de oferecer estes conteúdos na rede de computadores, provoca mudanças significativas nos perfis dos profissionais que interagem de forma a organizar, recuperar e disponibilizar a informação.

Todo texto digital em ambiente *Web* tem a flexibilidade, a velocidade e a quebra de espaços geográficos, requerendo profissionais preparados para catalogar o seu conteúdo e disponibiliza-los de forma adequada aos leitores.

2.2 Bibliotecas Digitais e Virtuais

As bibliotecas digitais surgiram para suprir a necessidade de grandes quantidades de arquivos a serem armazenados, como também a comodidade de acesso a estes documentos, além de proporcionar um custo relativamente baixo. Marchiori (1997) afirma que este modelo de biblioteca agrega o acervo de forma digital, seja por meio de armazenamento em discos magnéticos, óticos entre outros. Ou seja, o termo biblioteca digital está diretamente ligado ao conceito de mecanismos eletrônicos que facilitam o armazenamento e a localização da demanda informacional, proporcionando a singular possibilidade para a concepção de novos serviços de informação a partir da integração de objetos digitais heterogêneos.

Esse contexto de rápidas transformações oferece as condições primordiais para o estabelecimento de uma infra-estrutura técnica que viabiliza o surgimento de diversas atividades centradas no conhecimento e na informação globalmente distribuídos. Para Marchiori (1997), trata-se do relacionamento com o acesso por meio de redes e sistemas de informações computadorizados, possibilitando o acesso à informação em formato eletrônico, evitando uma visita a uma biblioteca física.

Apesar do termo biblioteca digital e virtual serem comumente utilizados para definir o mesmo conceito, existem suas diferenças. De forma resumida, biblioteca digital está ligada ao conceito de armazenamento de material em meios eletrônicos (mídias, discos magnéticos e demais tecnologias), enquanto biblioteca virtual está relacionada ao tipo de acesso não físico e a disponibilidade por meio de redes de computadores. [Marchiori 1997].

2.3 Base de Dados Científica

As Bases de Dados foram criadas com a intenção de disponibilizar, em um único acervo, centenas de revistas científicas, fazendo com que o pesquisador tenha um ganho de tempo ao procurar o assunto desejado. De um modo mais simplificado, pode-se dizer que as Bases de Dados científicas são um local onde encontramos centenas de artigos e revistas, sem a necessidade de se ficar navegando por diversos sites. É importante ressaltar que todas as publicações disponíveis em uma base de dados estão respaldadas

por qualidade e originalidade, devido aos criteriosos processos de seleção [Teixeira 2011].

Existem dois principais tipos de bases, as que oferecem o texto de forma integral e as bases que apenas referenciam, ou seja, que apenas indicam sobre a existência do artigo, ficando a cargo do leitor procurá-lo no site específico de origem. O site Scielo, mantenedor de uma das maiores bases científicas nacionais, teve como propósito em sua criação, permitir que as produções científicas brasileiras se tornassem cada vez mais visíveis em âmbito internacional e constituir uma base de dados sólida que fornecesse ferramentas que pudessem avaliar a produção nacional de conhecimento [Meneghini 2003]. Ainda conforme o autor, o site já possuía um acervo com mais de vinte mil artigos, cujos materiais podem ser acessados na íntegra até hoje, ficando evidente a necessidade do uso de um banco de dados para a administração e uso destes conteúdos.

2.4 Sistema de Gerenciamento de Banco de Dados (SGBD)

Juntamente ao desenvolvimento das tecnologias na área de computação e da informação, houve um grande aumento no volume de dados produzidos, principalmente pelas organizações, onde o fluxo de informação gerado se torna cada vez maior. Como soluções organizações utilizavam módulos de aplicações isoladas, desenvolvidas especificamente para uma determinada tarefa, o que acabava por gerar muita redundância de informação ou, até mesmo, isolamento de algumas tarefas. Além disso, grandes partes destes módulos não possuíam a capacidade de gerenciamento de concorrência de acesso por múltiplos usuários ou processos.

Visando este cenário, surge o conceito de Banco de Dados, que trata de uma coleção de dados que podem ser relacionados, inseridos, atualizados e recuperados. Em Banco de Dados, preza-se a minimização de inconsistências e redundância de dados, facilidade de acesso e garantia de integridade. Em teoria, um sistema de banco de dados deve ser capaz de minimizar ou, em alguns casos, até mesmo eliminar informações repetidas dentro de seu escopo, pois tais redundâncias podem gerar problemas na hora da manipulação dos dados, possibilitando sua alteração em um só local e não em todo o sistema, causando o que é chamado de inconsistência de informação. O Sistema Gerenciador de Banco de Dados (SGBD) deve preocupar-se em facilitar ao máximo o acesso aos dados, atentando-se sempre para a possibilidade de acesso concorrente, fazendo com que a mesma informação seja compartilhada por inúmeros usuários simultaneamente.

Um Sistema Gerenciador de Banco de Dados (SGBD) trata-se de um conjunto de aplicações destinadas à criação, manipulação e gerenciamento em geral de um banco de dados. Tem como funções facilitar a utilização através de sua interface, proteger contra falhas de *hardware/software* e garantir a segurança das informações contra acessos maliciosos e não identificados. Resumidamente o sistema SGBD faz a interface entre camada lógica (instâncias) e a camada física de armazenamento de dados (*storages*, discos, entre outros). Sua arquitetura é composta por algum modelo de organização, sendo os principais modelos: hierárquicos, relacional, de rede, orientado a objetos, objeto-relacional e o *Not only SQL* (NoSQL) [Cayres 2015].

2.4.1 Modelo Relacional

O modelo relacional surgiu para atender a necessidade de aumentar a independência de dados nos SGBD's, provendo funções baseadas em álgebra relacional para armazenamento e recuperação permitindo exclusividade de processamento (processos dedicados). Foi resultado de um estudo realizado na década de 70 por um pesquisador da IBM chamado Ted Codd, que produziu um artigo descrevendo a proposta de um novo modelo para recuperação e armazenamento de dados que implementa a organização por relação nas estruturas de dados, sendo necessária a imposição de restrições para evitar comportamentos indesejados, como a perda ou repetição da informação. Tais restrições são caracterizadas por junções de relações, integridade referencial e chaves de integridade [Cayres 2015]. De acordo com a Figura 1, no primeiro exemplo nota-se o relacionamento entre duas entidades com cardinalidade de 1 para 1, no segundo exemplo por se tratar de um relacionamento de N para N, cria-se uma entidade associativa que pode receber atributos independentes além de receber como chaves estrangeiras as chaves primárias das outras entidades ligadas a ela, formando uma chave primária composta.

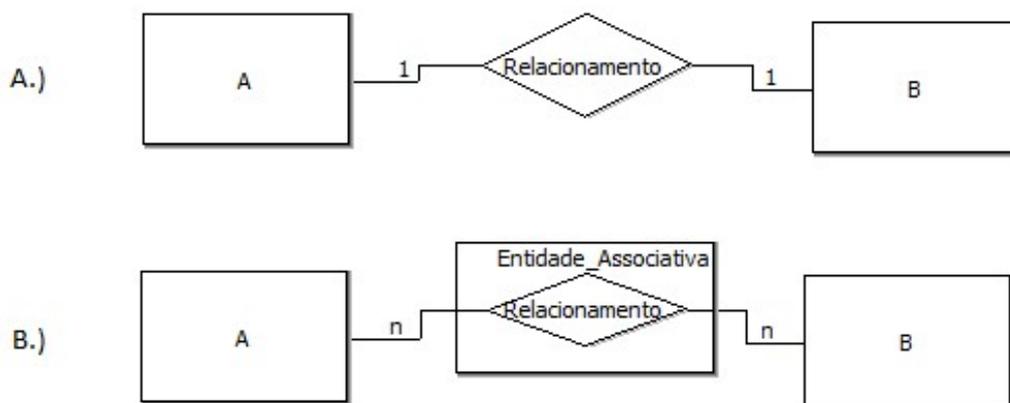


Figura1. Exemplo de Modelo Entidade Relacionamento

2.5 MySQL

MySQL é um SGBD relacional de código fonte aberto, utilizado comumente para gerir bases de aplicações livres, geralmente voltadas para a *Web*. Utiliza como linguagem de manipulação a Linguagem de Consulta Estruturada ou *Structure Query Language (SQL)*, o sistema apresenta uma fácil integração com a linguagem de programação PHP e também possui a característica de multiplataforma de Sistemas Operacionais (SO), ainda conta com desempenho de alto nível e uma boa estabilidade na gestão de seus processos [Guimarães et al. 2011]. Atualmente, vem se tornando um dos SGBDs mais utilizados em conjunto com aplicações *Web* no mundo, inclusive por corporações famosas como o Facebook, Suzuki e Google, entre outros. Essa preferência é devido a sua fácil integração com as plataformas existentes, além de ter um bom desempenho quando integrado a estas aplicações *on-line*, e de sua manipulação simplificada e de fácil entendimento [Oracle 2016].

Devido aos experimentos deste artigo ter relação direta ao ambiente *Web* este será o banco de dados utilizado, projeto que será melhor descrito nos materiais e métodos. Um recurso ainda pouco explorado nesta plataforma são os armazenamentos de arquivos diretamente nos BDs, com o uso do *Binary Large Object (BLOB)* realizando a conversão de arquivos de textos, áudios, filmes, entre outros; em códigos

binários permitindo seu armazenamento e gerenciamento pelo próprio SGBD. A Figura 2 demonstra um processo de conversão de um arquivo PDF para binário. Primeiramente o arquivo original passa por um pré-processamento que pode ocorrer em uma linguagem de programação, onde o arquivo passa pela conversão de seu conteúdo em binário, após este processo é gerado o objeto binário que é armazenado no banco de dados. Importante destacar que os bancos de dados MySQL não armazenam os arquivos originais, apenas seus objetos binários, necessitando assim um processamento inverso para converter o objeto binário no arquivo original na sua recuperação.

2.5.1 Binary Large Object

O MySQL é preparado para lidar com grande quantidade de dados, o que torna viável a sua utilização na proposta de realização da construção de uma biblioteca acadêmica, garantindo assim a integridade dos arquivos e um acesso rápido aos conteúdos sem prejudicar a performance do banco de dados. Em sua manipulação é possível trabalhar com tipo de dado BLOB que realiza o armazenamento de arquivos convertidos em dados binários, podendo estes ser imagens, arquivos: PDF, Word, entre outros. Dentre as vantagens para utilização deste tipo de dado se destacam: integridade referencial, baixa quantidade de conflito de nomes de arquivos, possibilidade de replicação em vários servidores, além de, o tipo de dado BLOB pode ser utilizado como uma alternativa à necessidade de um repositório de arquivos dedicado para armazenamento de imagens, pois, com a utilização dele é possível armazenar no próprio servidor MySQL os arquivos de mídia. Um campo de uma tabela formatado para este tipo de dado é na realidade uma coleção de dados binários armazenados como uma única entidade em um SGBD. [Dionísio 2015].

A capacidade do BLOB para armazenamento equivale a 16 bits. Em seguida, temos o MEDIUMBLOB, este possui a capacidade de armazenamento equivalente à 24 bits. Como qualquer recurso de banco de dados, os campos do tipo BLOB possuem suas particularidades, dentre as quais podemos destacar que não podem ser atribuídos como chaves primárias, além do fato de que, para estes campos, não podemos utilizar os comandos GROUP e SORT na realização de consultas [Dionísio 2015].



Figura 2. Processo de conversão e armazenamento no SGBD

3 Materiais e Métodos

Para a realização dos experimentos foi utilizado um ambiente virtual configurado através do programa *VirtualBox*, com uma máquina virtual com 4GB de RAM e um espaço em disco dedicado de 25GB, a qual recebeu a instalação do banco de dados MySQL, sobre o sistema operacional Windows Server-2008.

Com os estudos realizados chegou-se a definição das seguintes entidades para um modelo ideal para a proposta apresentada: *Autores*, *Cursos*, *Eventos*, *Trabalhos*, *Palavras_Chaves*, *Apresentacao*, *Areas*, *Temas*, *Autores_Trabalhos* e *Trabalhos_PalavrasChaves*, onde a entidade *Trabalhos* apresentada na Figura 3 é a principal do modelo, pois será responsável por armazenar os trabalhos recebidos e referenciar a busca dos conteúdos. A seguir um breve detalhamento dos campos que compõem sua estrutura: *IdTrabalhos*, campo da tabela para leitura de números inteiros, sendo chave primária de identificação com incremento automático na geração de valores. *Data_publicacao*, este campo armazena a data de publicação do trabalho. *Referencia_CD*, tipo de dado texto onde será armazenada as referências da mídia entregue pelos autores, também usado para uma busca caso seja necessária a recuperação física do arquivo em possíveis casos de falhas do sistema, o que serve como medida de backup. *idAreas*, campo no formato de chave estrangeira, recebe o código da sequência de informação da tabela de *Areas*. *idTemas*, também chave estrangeira, recebe o código da sequência de informação da tabela de *Temas*, e por fim o campo *Trabalho*, este no formato MEDIUMLOB, utilizado para armazenar os arquivos postados em formato binário, esse tipo foi escolhido pela sua grande capacidade, dando possibilidade de inserir arquivos de ordem de até 16MB.

Trabalhos	
	idTrabalhos: INTEGER
	Data_Publicação: DATE
	referencia_CD: VARCHAR(10)
	idAreas: INTEGER
	idTemas: INTEGER
	Trabalho: MEDIUMBLOB

Figura 3. Entidade Trabalhos do modelo relacional

Na figura 4 temos a entidade *Autores* responsável pelo armazenamento dos dados de todos os autores dos trabalhos postados, na qual foram identificados os seguintes elementos: *idAutores*, campo da tabela para leitura de números inteiros, sendo chave primária de identificação com incremento automático na geração de valores. *RA*, campo para leitura de números inteiros, será armazenado informações do código institucional de cada autor. *E-mail*, campo do tipo texto, armazena os e-mails dos autores dos trabalhos. *CPF*, campo em formato de texto, armazena o documento de identificação. *Nome*, campo do tipo texto, armazena nome completo dos autores.

Autores	
	idAutores: INTEGER
	RA: NUMERIC(13)
	Email: VARCHAR(80)
	CPF: VARCHAR(11)
	Nome: VARCHAR(80)

Figura 4. Entidade Autores do modelo relacional

Para a entidade *Cursos* responsável pelo armazenamento de todos os cursos oferecidos pela instituição, como mostra a Figura 5 estão relacionados os atributos abaixo:

idCursos, campo da tabela para leitura de números inteiros, sendo chave primária de identificação com incremento automático na geração de valores. Chave primária de identificação. *Curso*, campo no formato texto para armazenar o nome do curso relacionado ao trabalho.

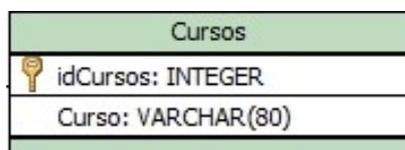


Figura 5. Entidade Cursos do modelo relacional

De acordo com a Figura 6, a entidade associativa *Cursos_Autores_Trabalhos* é uma entidade associativa devido ao relacionamento entre as entidades *Trabalhos*, *Cursos* e *Autores*, sua responsabilidade é de relacionar as três entidades e fazer um *check* na informação classificando se o trabalho foi feito por um orientador ou por aluno. Esta entidade atende o modelo com os seguintes atributos: *IdTrabalhos*, campo da tabela para leitura de números inteiros, sendo chave composta de identificação. *idAutores*, campo da tabela para leitura de números inteiros, sendo chave primária de identificação com incremento automático na geração de valores. *idCursos*, campo da tabela para leitura de números inteiros, sendo chave primária de identificação com incremento automático na geração de valores. *Orientador*, campo no formato texto para checagem se o trabalho for de algum professor orientador. *AutorPrincipal*, campo no formato texto para checagem se o trabalho pertence a algum aluno da faculdade.

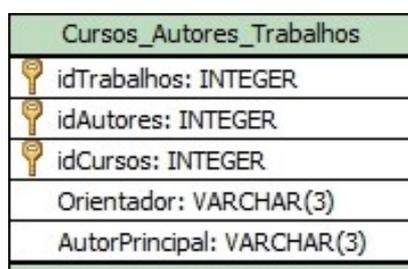


Figura 6. Entidade Cursos_Autores_Trabalhos do modelo relacional

Na entidade *Eventos* cuja é representada pela Figura 7, armazena os eventos para os quais o trabalho foi publicado, nela estão relacionados os elementos: *idEventos*, campo da tabela para leitura de números inteiros, sendo chave primária de identificação com incremento automático na geração de valores. *NomeEvento*, campo no formato texto, armazena informações do Evento onde o trabalho foi publicado.

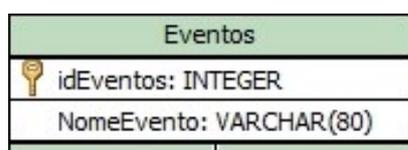


Figura 7. Entidade Eventos do modelo relacional

A entidade associativa *Apresentacao* representada pela Figura 8, é responsável pelo armazenamento da data de trabalho listando os seguintes campos: *idEventos*, campo da tabela para leitura de números inteiros, sendo chave composta de identificação. *idTrabalhos*, campo da tabela para leitura de números inteiros, sendo chave composta

de identificação. *Data*, campo no formato data, armazena o dia, mês e ano em que o trabalho foi apresentado.

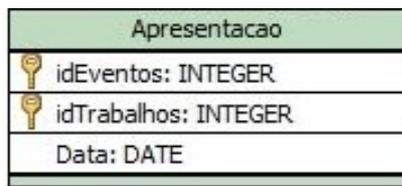


Figura 8. Entidade Apresentacao do modelo relacional

De acordo com a Figura 9, a entidade *Areas* recebe no banco as áreas de conhecimento que relacionam os trabalhos com seguintes atributos: *idAreas*, campo da tabela para leitura de números inteiros, sendo chave primária de identificação com incremento automático na geração de valores. Chave primária de identificação. *Areas*, campo no formato texto, responsável pelo armazenamento de informações das áreas específicas relacionadas ao trabalho.

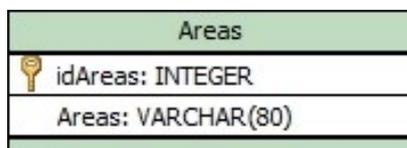


Figura 9. Entidade Areas do modelo relacional

A entidade *Temas*, Figura 10, fica com a responsabilidade de listar o título do trabalho no banco, esta entidade lista os seguintes campos: *idTemas*, campo da tabela para leitura de números inteiros, sendo chave primária de identificação com incremento automático na geração de valores. *Tema*, campo no formato texto, armazena o tema do trabalho postado no sistema.



Figura 10. Entidade Temas do modelo relacional

Já na entidade *Palavras_Chaves* será armazenado no banco palavras para conciliar e facilitar a busca do artigo no sistema, conforme Figura 11: *idPalavrasChaves*, campo da tabela para leitura de números inteiros, sendo chave primária de identificação com incremento automático na geração de valores. *Palavras*, campo no formato texto, onde lista as palavras chaves para facilitar a localização do arquivo.

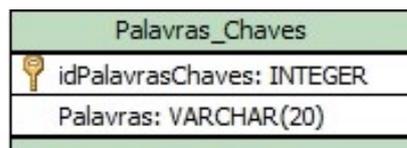


Figura 11. Entidade Palavras_Chaves do modelo relacional

Por fim a entidade associativa *Trabalhos_PalavrasChave* tem a função de relacionar as entidades *Trabalhos* e *PalavrasChaves* devido ao relacionamento de n para n conforme

representada pela Figura 12 contém os seguintes atributos: *idPalavrasChaves*, campo da tabela para leitura de números inteiros, sendo chave composta de identificação. *idTrabalhos*, campo da tabela para leitura de números inteiros, sendo também chave composta de identificação.

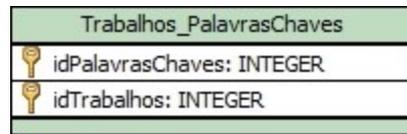


Figura 12. Entidade *Trabalhos_PalavrasChaves* do modelo relacional

O modelo de entidade relacionamento (MER) foi construído para atender os requisitos de uma biblioteca acadêmica, conforme é possível observar na Figura 13 o modelo foi preparado para receber os trabalhos em formato .PDF e convertê-los em objetos binários, dessa forma eliminando a necessidade de um repositório dedicado para armazenamento de arquivos. Veremos a seguir o relacionamento e a cardinalidade de cada entidade conforme a proposta deste trabalho.

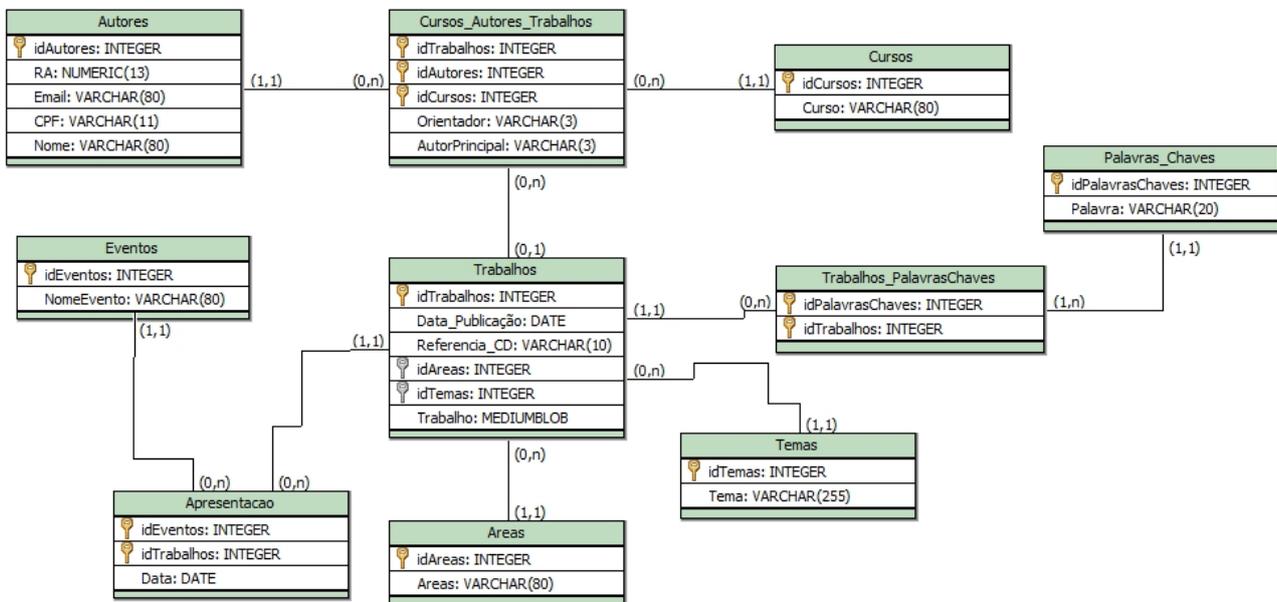


Figura 13. Modelo Entidade Relacionamento

Após identificado todos os atributos das entidades, foram definidos como elas se relacionam: A entidade *Trabalhos* se relaciona com a entidade *Temas* recebendo como chave estrangeira o campo *idTemas* de cardinalidade de n para 1. Da mesma forma é possível visualizar que entidade *Areas* é relacionada com *Trabalhos* utilizando o campo *idAreas* como chave estrangeira de cardinalidade n para 1.

Nota-se que entidade *Trabalhos* se relaciona com a entidade *Eventos* através da entidade associativa *Apresentacao*, que possui como atributos as chaves primárias *idEventos* e *idTrabalhos* se tornando uma chave primária composta. A necessidade de se criar uma entidade associativa deve-se ao fato do relacionamento ser de n para n. Para as entidades *Trabalhos* e a entidade *Palavras_Chaves* foi necessária a criação de mais uma entidade associativa com o nome de *Trabalhos_PalavrasChaves*, que tem

como atributos as chaves primárias *idPalavrasChaves* e *idTrabalhos* se tornando uma chave primária composta, novamente a necessidade de se criar uma entidade associativa deve-se ao fato do relacionamento de ambas serem de n para n. A entidade *Trabalhos* relaciona-se com a entidade *Cursos* e *Autores* através uma entidade associativa *Autores_Trabalhos* no qual as chaves compostas para este relacionamento são: *idTrabalhos*, *idAutores* e *IdCursos*, ambas as entidades com cardinalidade definida de n para n.

4 Resultados

Diante da ideia proposta, foi aplicada regras no banco para que o armazenamento de arquivos receba documentos do tamanho de até 16MB com a utilização do MEDIUMLOB. A problemática encontrada a princípio foi na configuração do servidor Apache do arquivo *PHP.ini* onde limitava o *upload* de arquivos para 2MB, sendo necessária sua alteração para conforme mostrado no comando abaixo, que é responsável por delimitar o tamanho máximo de arquivos enviados: *upload_max_filesize=16M*

Conforme Figura14, foi criada uma classe no PHP para automatizar o processo de conversão dos arquivos no banco de dados. A variável *\$_FILES* é responsável por selecionar os arquivos em formato PDF com nome temporário e o seu tamanho. A função *fopen()* é utilizada para abrir o arquivo, por se tratar de um arquivo binário, é necessário utilizar opção “rb” onde a letra “r” solicita a leitura do arquivo e a letra “b” força a abertura do arquivo em modo binário. Após esta etapa, o comando *fread()* é responsável pela leitura dos dados do arquivo aberto e repassa a informação de volta para o PHP. Este comando é passado como *fread* (variável que recebe a informação do *fopen()*, tamanho). O arquivo é então aberto na variável *\$conteudo* onde seu tamanho é medido em bytes e encapsulado.

```
$arquivo = $_FILES['pdf']['tmp_name'];
$tamanhoArquivo = $_FILES['pdf']['size'];
$fp = fopen($arquivo, "rb");
$conteudo = fread($fp, $tamanho);
$conteudo = addslashes($conteudo);
fclose($fp);
```

Figura14. Código em linguagem PHP de conversão do arquivo PDF para binário.

O script de busca padrão do sistema conforme ilustrado pela Figura15 relaciona todos os trabalhos conforme a pesquisa feita pelo usuário. As informações aparecerão listadas na parte inferior da página na seção de Trabalhos buscados. Conforme ilustrado pela Figura16 o usuário digitou a palavra *Oracle* com retorno de uma linha no sistema. Na construção do script a entidade Tema é acessada pela variável do PHP *\$tema* digitada pelo usuário e ordenada em ordem crescente.

```
SELECT tema, idTemas FROM temas
WHERE tema LIKE '%" . $tema . "%' ORDER BY tema ASC;
```

Figura15. Query para selecionar os trabalhos filtrando de acordo com o tema e classificando em ordem alfabética.



Figura16. Página de Busca por Temas de trabalhos.

A construção do script para leitura dos dados relaciona os itens da entidade *Temas* associados ao curso desejado como mostrado na Figura17. O sistema ainda possibilita a busca de trabalhos através por Cursos e Áreas relacionadas, A Figura18 demonstra a tela onde o usuário realizou a busca de todos os trabalhos de Banco de Dados disponíveis no sistema através da categoria *Cursos*.

```
SELECT t.tema, t.idtemas, c.idcursos FROM cursos_autores_trabalhos cat
INNER JOIN cursos c ON cat.idcursos= c.idcursos
INNER JOIN trabalhos tr ON cat.idtrabalhos= tr.idtrabalhos
INNER JOIN temas t ON t.idtemas = tr.idtemas
WHERE c.curso ='"$_GET["curso"]"'
ORDER BY tema ASC;
```

Figura17. Query para selecionar os trabalhos filtrando de acordo com o Curso desejado e classificando em ordem alfabética.



Figura18. Página de Busca por Cursos – Banco de dados

5 Conclusão

Foi possível provar com esta proposta que é viável a implementação de um banco de dados sem a necessidade de um repositório de arquivos para gestão de seus documentos, através de parâmetros como o BLOB para armazenar os arquivos de forma binária diretamente no banco de dados. Além de possibilitar mais segurança e integridade a todo conteúdo nele mantido já que todos os dados, inclusive os arquivos são administrados pelo SGBD. Tendo como ponto positivo a redução de custos na implementação da ferramenta, e a vantagem de disponibilizar todo material produzido na rede em um ambiente de alta disponibilidade com acessos simultâneos. Outro item importante a ser destacado é que o tempo de converter e desconverter o objeto binário não foi significativo, não comprometendo a exibição ou recuperação dos arquivos.

Como trabalhos futuros destacamos a realização de mais pesquisas relacionadas ao armazenamento de arquivos de forma binária, tanto em BDs relacionais, quanto os não relacionais.

Referências

- Blattmann, U. (2001) “Modelo de Gestão da informação Digital online em bibliotecas acadêmicas na educação à distância: biblioteca virtual”. Tese de doutorado em Engenharia de Produção, Universidade Federal de Santa Catarina, SC.
- Blattmann, U., Fachin, G. R. B., Rados, G. J.V. (1999) “Recuperar a informação eletrônica pela Internet”. Em: Revista da ACB: Biblioteconomia em Santa Catarina, Florianópolis, v.4, n.1, <https://revista.acb.org.br/racb/article/view/376/453>, Outubro.
- Cayres, P.H. (2015) “Modelagem de Banco de Dados”. Rede Nacional de Ensino e Pesquisa.
- Dionísio, E.J. “Armazenando Imagens no MySQL”
<http://www.devmedia.com.br/armazenando-imagens-no-mysql/32104>, Maio.
- Guimarães, M. C. S. et al. “Métricas em saúde coletiva: bases quantitativas e qualitativas para a criação de um índice de citação da literatura nacional em Saúde Coletiva”. Relatório de pesquisa para o Projeto CNPq - Processo 403522/2008-0. 2011. (Não publicado).
- Levacov, M. (1997) “Bibliotecas virtuais: (r) evolução?”, Ciência da Informação. Inf. v. 26 n. 2.
- Marchiori, P. Z. (1997) “Ciberteca ou biblioteca virtual: uma perspectiva de gerenciamento de recursos de informação”. Ciência da Informação, v.26, n.2, p. 115-124.
- Martins, W. (2001) “A palavra escrita: história do livro, da imprensa e da biblioteca”. Ática. 3ª edição.
- Meneghini, R. “O projeto Scielo (Scientific Electronic Library on Line): e a visibilidade da literatura científica Periférica”, <http://dx.doi.org/10.1590/S0100-40422003000200001>, Junho.
- Morales, A. B. T, Rossetti A. G. (2007) “O papel da tecnologia da informação na gestão do conhecimento”, Ciência da Informação. Inf. v. 36 n. 1, p. 124-135.

- Muccioli, C. et al. (2007) “A Produção Científica no Brasil,
http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0004-27492007000400001,
Novembro.
- Nascimento, L. L., Pinto V. B., Vale H. C. P. (2013) “O livro, a biblioteca e leitura:
conhecer o passado para entender a (r) evolução tecnológica”. XXV Congresso
Brasileiro de Biblioteconomia, Documento e Ciência da Informação, Florianópolis,
SC.
- Negroponte, N. (1995) “A vida digital”. Tradução de Sérgio Tellaroli. São Paulo, SP,
Companhia das Letras. 2ª edição.
- Santos Neto, J. A e Santos, R. P. (2013) “A Gestão de documentos nos arquivos
acadêmicos e a portaria MEC nº. 1.224/2013”, <http://www.uel.br/revistas/infoprof>,
Novembro.
- Oracle. (2016) “O banco de dados de código aberto mais conhecido do mundo”,
<https://www.oracle.com/br/mysql/index.html>, Novembro.
- Rodrigues, E. (1995) “Bibliotecas virtuais e cibertecários: o futuro já começou”,
<http://repositorium.sdum.uminho.pt/handle/1822/423>, Outubro.
- Silva, N. R. (2014) “O Bibliotecário mediante as novas tecnologias da informação e
comunicação” Universidade Federal do Ceará, CE.
- Teixeira, M. V. (2011) “Você sabe o que são as bases de dados científicas?”,
<https://bibliotecaucs.wordpress.com/2011/10/27/voce-sabe-o-que-sao-as-bases-de-dados-cientificas>, Novembro.