

UNIVERSIDADE REGIONAL DE BLUMENAU
CENTRO DE CIÊNCIAS EXATAS E NATURAIS
CURSO DE SISTEMAS DE INFORMAÇÃO - BACHARELADO

**REGRAS DE ASSOCIAÇÃO NA DISSEMINAÇÃO DA
INFORMAÇÃO DA BIBLIOTECA UNIVERSITÁRIA DA
FURB: ENFOQUE NAS NOVAS AQUISIÇÕES DO ACERVO**

MACIEL FELIPE DA SILVA

BLUMENAU
2010

2010/1-17

MACIEL FELIPE DA SILVA

**REGRAS DE ASSOCIAÇÃO NA DISSEMINAÇÃO DA
INFORMAÇÃO DA BIBLIOTECA UNIVERSITÁRIA DA
FURB: ENFOQUE NAS NOVAS AQUISIÇÕES DO ACERVO**

Trabalho de Conclusão de Curso submetido à
Universidade Regional de Blumenau para a
obtenção dos créditos na disciplina Trabalho
de Conclusão de Curso II do curso de Sistemas
de Informação - Bacharelado.

Prof. Cláudio Ratke, Mestre – Orientador

**BLUMENAU
2010**

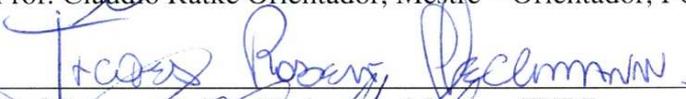
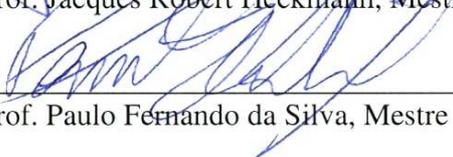
2010/1-17

**REGRAS DE ASSOCIAÇÃO NA DISSEMINAÇÃO DA
INFORMAÇÃO DA BIBLIOTECA UNIVERSITÁRIA DA
FURB: ENFOQUE NAS NOVAS AQUISIÇÕES DO ACERVO**

Por

MACIEL FELIPE DA SILVA

Trabalho aprovado para obtenção dos créditos na disciplina de Trabalho de Conclusão de Curso II, pela banca examinadora formada por:

Presidente:	 _____ Prof. Cláudio Ratke Orientador, Mestre – Orientador, FURB
Membro:	 _____ Prof. Jacques Robert Heckmann, Mestre – FURB
Membro:	 _____ Prof. Paulo Fernando da Silva, Mestre – FURB

Blumenau, 05 de julho de 2010.

Dedico este trabalho a meus pais e a todos que me ajudaram diretamente na realização deste.

AGRADECIMENTOS

A Deus primeiramente, por iluminar o meu caminho.

A meus pais, Sandra e José, que sempre me apoiaram nos estudos e estiveram presentes.

A meu irmão Vitor, pelas dicas de web design e de ferramentas.

À minha namorada Liliane, pela paciência, compreensão e todo o incentivo para realização do trabalho, e também pela ajuda prestada como bibliotecária quanto à revisão e normalização deste trabalho.

A bibliotecária Evanilde M. Moser, chefe da Seção de Automação e Centro de Memória Universitária pela autorização para o desenvolvimento inicial do sistema na Biblioteca, e ao analista de sistemas Marcos R. Cardoso, pela troca de informações e apoio necessário para realização deste trabalho.

Ao professor Alexander Roberto Valdameri, orientador no início da proposta do trabalho, pelo incentivo e contribuição da idéia inicial.

Ao meu orientador, Cláudio Ratke, por ter acrescentado um diferencial ao trabalho e acreditado na conclusão deste.

Aos professores avaliadores e membros da banca, pelas contribuições gerais no trabalho desde a proposta inicial.

A todos que de alguma forma contribuíram para a conclusão deste trabalho.

A mente que se abre a uma nova idéia jamais
voltará ao seu tamanho original.

Albert Einstein

RESUMO

Relato do trabalho desenvolvido utilizando técnica de regras de associação na disseminação de informação sobre as novas aquisições do acervo bibliográfico da Biblioteca Universitária da Universidade Regional de Blumenau. Aborda aspectos da gestão e disseminação da informação em bibliotecas universitárias e da mineração de dados, especificamente da técnica de regras de associação. Especifica o desenvolvimento e a implementação do sistema. O sistema foi desenvolvido para enviar, por e-mail, informações sobre as novas aquisições de acordo com as áreas selecionadas pelo usuário em seu perfil. A técnica de regras de associação foi empregada como um diferencial, objetivando enviar também informações de novas aquisições de outras áreas que podem interessar ao usuário, a partir das correlações detectadas pelo emprego desta. O sistema pode auxiliar nos serviços oferecidos pelas bibliotecas universitárias principalmente em relação às necessidades informacionais de seus usuários.

Palavras-chave: Mineração de dados. Regras de associação. Disseminação seletiva da informação. Bibliotecas universitárias.

ABSTRACT

Work report developed by using the technique of association rules to information dissemination concerning new acquisitions of the Blumenau Regional University's Academic Library bibliographic collection. It approaches aspects about information management and dissemination in academic libraries and data mining, specifically the technique of association rules. It specifies the system's development and implementation. The system was developed to send, by e-mail, new acquisitions information in accordance with the areas selected by the user in his profile. The technique of association rules was used as a differential, with objective to send also new acquisitions information from other areas that can be of interest to the user, from detected correlations using association rules. The system can help in the services offered by academic libraries mainly with relation to informational needs of users.

Key-words: Data mining. Association rules. Selective dissemination of information. Academic libraries.

LISTA DE FIGURAS

Figura 1 - Etapas do Processo da Descoberta de Conhecimento em Banco de Dados.....	20
Figura 2 - Exemplo de conjuntos de itens freqüentes conforme princípio Apriori	29
Figura 3 - Exemplo da redução de conjuntos de itens baseado na poda	29
Figura 4 - Diagrama de caso de uso da interface <i>web</i> para o usuário do sistema proposto.....	39
Figura 5 - Diagrama de caso de uso da ferramenta de envio.....	40
Figura 6 - Diagrama de caso de uso da ferramenta gerencial.....	40
Figura 7 - Diagrama de atividades do <i>website</i> para o usuário configurar seu perfil	41
Figura 8 - Diagrama de atividades da ferramenta de disseminação das informações	42
Figura 9 - Diagrama de entidade e relacionamento	44
Figura 10 - Tela de <i>login</i> do usuário na interface do módulo <i>web</i>	50
Figura 11 - Tela da página principal para preferências do perfil do usuário.....	50
Figura 12 - Tela da página principal para autorização e salvar preferências	51
Figura 13 - Tela do aplicativo para disseminação das informações	52
Figura 14 - Tela principal do aplicativo gerencial do módulo interno	53
Figura 15 - Tela de parametrizações	53
Figura 16 - Tela de consulta das áreas cadastradas para perfil do usuário.....	54
Figura 17 - Tela de consulta por áreas correlatas apontadas pela regra de associação	54
Figura 18 - Tela de relatório de acompanhamento da utilização do sistema.....	55

LISTA DE QUADROS

Quadro 1 - Exemplo de transações realizadas	26
Quadro 2 - Exemplo de representação binária de transações realizadas	26
Quadro 3 - Exemplo cálculo para encontrar suporte e confiança de uma regra.....	27
Quadro 4 - Pseudocódigo do algoritmo Apriori para encontrar conjuntos de itens frequentes	30
Quadro 5 - Exemplo de uma base de dados com transações	31
Quadro 6 - Geração dos conjuntos de itens frequentes	31
Quadro 7 - Exemplo da extração de regras de associação com base no algoritmo Apriori	32
Quadro 8 - Requisitos funcionais	38
Quadro 9 - Requisitos não funcionais	38
Quadro 10 - Função para retornar informações da obra.....	46
Quadro 11 - Exemplo da base de dados de transações da abordagem adotada.....	47
Quadro 12 - Trecho de código para gerar as regras de associação.....	48
Quadro 13 - Trecho de código para envio de <i>e-mails</i>	49
Quadro 14 - Exemplo de regras geradas pelo sistema.....	56
Quadro 15 - Descrição do caso de uso Efetuar <i>Login</i>	64
Quadro 16 - Descrição do caso de uso Selecionar as áreas de preferência	64
Quadro 17 - Descrição do caso de uso Alterar seleção das áreas de preferência	65
Quadro 18 - Descrição do caso de uso Autorizar ou não o envio de informações por <i>e-mail</i> ..	65
Quadro 19 - Descrição do caso de uso Enviar relatórios personalizados das informações.....	66
Quadro 20 - Descrição do caso de uso Consultar as áreas selecionadas no perfil do usuário..	67
Quadro 21 - Descrição do caso de uso Buscar novas aquisições por área	67
Quadro 22 - Descrição do caso de uso Encontrar áreas correlatas através de regras de associação	68
Quadro 23- Descrição do caso de uso Gerar relatório das novas aquisições encontradas	68
Quadro 24 - Dicionário de dados.....	71

LISTA DE SIGLAS

ACCR2 - Código de Catalogação Anglo-americano
BU - Biblioteca Universitária
CDD - Classificação Decimal de Dewey
CSS - *Cascading Style Sheet*
DSI - Disseminação Seletiva da Informação
FURB - Fundação Universidade Regional de Blumenau
GUI - *Graphical User Interface*
HTML - Linguagem de Marcação de Hipertexto
KDD - *Knowledge Discovery in Databases*
MARC - *Machine Readable Cataloging*
MER - Modelagem Entidade e Relacionamento
MFN - *Master File Number*
PHP - *Hypertext Preprocessor*
PL/SQL - *Procedural Language/Structured Query Language*
TID - Identificador único de transação
UML - *Unified Modeling Language*

LISTA DE SÍMBOLOS

\Rightarrow - então
% - por cento
 \subset - está contido
 \cup - união
 \cap - interseção
 σ - contador
 \emptyset - vazio
 \neq - diferente
 \in - pertence

SUMÁRIO

1 INTRODUÇÃO.....	12
1.1 OBJETIVOS DO TRABALHO	13
1.2 ESTRUTURA DO TRABALHO	14
2 FUNDAMENTAÇÃO TEÓRICA	15
2.1 GESTÃO E DISSEMINAÇÃO DA INFORMAÇÃO	15
2.2 MINERAÇÃO DE DADOS.....	18
2.2.1 Tarefas de mineração de dados.....	21
2.2.2 Técnicas de mineração de dados	22
2.2.3 Aplicações da mineração de dados	24
2.2.4 Regras de associação	25
2.3 TRABALHOS CORRELATOS.....	32
3 DESENVOLVIMENTO DO SISTEMA.....	34
3.1 LEVANTAMENTO DE INFORMAÇÕES.....	34
3.2 ESPECIFICAÇÃO	39
3.2.1 Diagrama de caso de uso	39
3.2.2 Diagrama de atividades	41
3.2.3 Modelagem entidade relacionamento.....	43
3.3 IMPLEMENTAÇÃO	45
3.3.1 Técnicas e ferramentas utilizadas	45
3.3.2 Operacionalidade da implementação.....	49
3.4 RESULTADOS E DISCUSSÃO	55
4 CONCLUSÕES.....	58
4.1 EXTENSÕES	59
REFERÊNCIAS	60
APÊNDICE A – Detalhamento dos casos de uso.....	64
APÊNDICE B – Dicionário de dados das tabelas.....	69
ANEXO A - Áreas da CDD utilizadas pela BU para divulgação de informações de novas aquisições	72
ANEXO B – Formato MARC.....	76
ANEXO C – Exemplo de registro bibliográfico no formato MARC	77

1 INTRODUÇÃO

Na sociedade contemporânea a informação, como meio de criação de conhecimento, exerce um papel fundamental na produção de riqueza e na contribuição para o bem-estar e qualidade de vida dos cidadãos. Desta forma, a informação assume um valor estratégico e torna-se o cerne das organizações, que passam a investir na gestão da informação, utilizando as tecnologias, ferramentas e técnicas mais avançadas para obter, selecionar, mapear, organizar e disseminar a informação. Diante das necessidades informacionais das organizações, os sistemas de informação e de banco de dados tornam-se indispensáveis no seu cotidiano, atrelados aos rápidos avanços tecnológicos e as exigências competitivas do mercado.

As bibliotecas são organizações que tem a informação como foco de suas atividades e possuem um importante papel dentro da universidade, pois contribuem efetivamente para a geração do conhecimento ao disponibilizar uma diversidade de materiais bibliográficos e informações aos seus usuários.

Entre os serviços de informação oferecidos pelas bibliotecas está a disseminação da informação, que é um importante veículo de comunicação entre a biblioteca e seus usuários, que têm necessidades informacionais distintas de acordo com o perfil e área de atuação. É papel da biblioteca fornecer ao seu usuário o produto informacional que supra a suas necessidades específicas, quer seja para estudo, para pesquisa ou para a tomada de decisão (TARAPANOFF; ARAÚJO JÚNIOR; CORMIER, 2000).

Na biblioteca universitária a disseminação da informação contribui para a produção do conhecimento no meio acadêmico, incentivando o uso adequado das informações e dos conhecimentos e minimizando os esforços dos usuários (DIAS, 2005).

As tecnologias de informação são aliadas na disponibilização dos serviços de informação, pois permitem melhorar os serviços oferecidos, na divulgação, agilidade ou eficiência. Como os usuários necessitam obter informações mais selecionadas em menos tempo, as bibliotecas precisam utilizar as tecnologias disponíveis, seja para divulgar os serviços oferecidos ou para levar as informações aos usuários de acordo com o seu interesse.

A Biblioteca Universitária (BU) da Universidade Regional de Blumenau (FURB) possui um fluxo constante de aquisições de materiais bibliográficos, que são inseridos ao acervo. Atualmente, para obter informações das novas aquisições da BU o usuário deve realizar uma consulta no *website* da biblioteca, que tem se demonstrado pouco atrativa. Diante

disso, buscou-se incorporar novas tecnologias como uma alternativa para que as informações sejam enviadas ao usuário de forma automatizada e filtrada, possibilitando a escolha das suas áreas de preferência.

Ainda com o intuito de agregar valor na disseminação das informações sobre as novas aquisições da BU, foram utilizadas tecnologias como a mineração de dados, em especial a técnica de regras de associação. Tal técnica é empregada para detectar informações de correlações ou tendências entre dados de transações em bases de dados, de modo que possam ser úteis na disponibilização de serviços de informação.

Neste trabalho foi utilizada técnica de regras de associação para encontrar a correlação entre as áreas das obras emprestadas pelos usuários, objetivando desvendar que se um usuário empresta obras de uma área, em geral, também empresta obras de outra determinada área.

O sistema desenvolvido enviará para os usuários uma lista de obras de acordo com as áreas selecionadas por este e também enviará como sugestão uma lista de obras que podem interessar, esta segunda lista é gerada a partir das áreas detectadas pelo uso das regras de associação.

1.1 OBJETIVOS DO TRABALHO

O objetivo deste trabalho é gerenciar a informação na Biblioteca Universitária (BU) da Universidade Regional de Blumenau (FURB), focando na disseminação da informação sobre novas aquisições do acervo bibliográfico para os usuários da BU.

Os objetivos específicos do trabalho são:

- a) disponibilizar uma interface *web* na qual o usuário possa selecionar as áreas do conhecimento disponíveis para receber informações;
- b) verificar as informações de novas aquisições disponíveis no acervo, separando-as por áreas predefinidas e gerar um relatório personalizado de acordo com o perfil de cada usuário;
- c) promover a disseminação das informações referentes às novas aquisições bibliográficas de acordo com as áreas de preferência selecionadas previamente pelo usuário;
- d) promover a disseminação das informações referentes às novas aquisições bibliográficas sob a forma de sugestões de acordo com o perfil do usuário

detectado através de regras de associação.

1.2 ESTRUTURA DO TRABALHO

Este trabalho está organizado em quatro capítulos, sendo que, no primeiro capítulo foi apresentada a introdução, os objetivos e como o trabalho está estruturado.

No segundo capítulo é apresentada a fundamentação teórica bem como os assuntos que serviram de base para o desenvolvimento do trabalho e a apresentação de alguns trabalhos correlatos.

No terceiro capítulo está descrito o desenvolvimento do sistema proposto, as técnicas e ferramentas utilizadas bem como a elaboração de alguns diagramas para auxiliar na compreensão do sistema.

E por fim, no quarto capítulo as conclusões, limitações e sugestões para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são abordados os principais assuntos e conceitos que auxiliam o melhor entendimento do trabalho e que serviram como base para o desenvolvimento do sistema.

2.1 GESTÃO E DISSEMINAÇÃO DA INFORMAÇÃO

A informação é vista por Davenport e Prusak (1998) como dados que fazem diferença. Enquanto os dados são um conjunto de fatos, a “informação é uma mensagem, geralmente na forma de um documento ou uma comunicação audível ou visível.” (DAVENPORT; PRUSAK, 1998, p. 4).

A informação tem por finalidade mudar o modo como o destinatário vê algo, exercer algum impacto sobre seu julgamento e comportamento. Ela deve informar. O significado original da palavra ‘informar’ é ‘dar forma a’, sendo que a informação visa a modelar a pessoa que a recebe no sentido de fazer alguma diferença em sua perspectiva. (DAVENPORT; PRUSAK, 1998, p. 4).

A informação está inserida em todos os ambientes e está presente em todas as atividades humanas, sociais, científicas, tecnológicas, culturais, políticas e econômicas, assumindo um novo status e importância na sociedade atual. Tais informações são variadas e produzidas de forma contínua, necessitando ser recuperadas, classificadas, organizadas, processadas, analisadas e difundidas pelas organizações no menor tempo possível (STAREC, 2005).

Por isso gerenciar a informação é primordial, e pode se tornar uma vantagem competitiva. Silva e Tomaél (2007, p. 1) afirmam que “para ser utilizada estrategicamente, é fundamental que a informação seja gerida em favor da sobrevivência e competitividade organizacional.” A gestão da informação possibilita às organizações gerir tanto os recursos internos quanto os externos à organização.

A gestão da informação engloba a prospecção, seleção e obtenção da informação; o mapeamento e reconhecimento dos fluxos formais de informação; o tratamento, análise e armazenamento da informação utilizando tecnologias de informação; a disseminação e mediação da informação ao público interessado, e a criação e disponibilização de produtos e serviços de informação (VALENTIM, 2002). O gerenciamento da informação permite um

melhor conhecimento da missão da organização e uma visão do ambiente, no qual os gestores estão inseridos (CARVALHO, 2006).

As organizações têm fluxos de informação específicos e cabe à gestão da informação mapeá-los, identificando pessoas, fontes de informação, tecnologia utilizada, produtos e serviços. Com isso compõem um conjunto estruturado de atividades referentes à maneira como a informação e o conhecimento são obtidos, utilizados e distribuídos (SILVA; TOMAÉL, 2007).

Carvalho (2001, p. 1) acrescenta que “para garantir um adequado fluxo de informação, a construção de sistemas de informação reforça o acesso ao conhecimento para a tomada de decisões nesses espaços, uma vez que a informação está ligada ao conhecimento do ambiente interno e também do externo.”

A informação analisada possui valor estratégico e a evolução das tecnologias da informação e telecomunicação contribui significadamente para esta ação. Carvalho (2001) ressalta que a facilidade de acessar as redes e aos bancos de dados disponibiliza uma quantidade de informação cuja absorção é inviável.

No contexto acadêmico as bibliotecas são as principais responsáveis pela gestão da informação. Uma biblioteca possui um papel importante dentro da universidade, pois disponibiliza diversos tipos de materiais bibliográficos que proporcionam informações aos seus usuários. As bibliotecas universitárias “têm como missão oferecer aos seus usuários informações relevantes para a realização de pesquisas e para o ensino, procurando tornar o acesso, a recuperação e a localização das informações compatíveis com as suas necessidades.” (LUCAS; SOUZA, 2007, p. 2).

De acordo com Dias (2005), uma das formas na produção de conhecimento a partir do meio acadêmico é a implantação de um serviço de disseminação da informação na biblioteca universitária, pois promove o uso intensivo da informação e dos conhecimentos relevantes de forma adequada, poupando tempo e esforço por parte de seus usuários.

Disseminar significa difundir, propagar, semear ou espalhar por muitas partes (MICHAELIS, 1998). A disseminação da informação é uma importante etapa da gestão da informação, completando o processo de seleção, tratamento e análise da informação.

As organizações que necessitam disseminar a informação devem utilizar as tecnologias de informação a seu favor, para melhorar cada vez mais os serviços oferecidos, seja no âmbito de divulgação, agilidade ou eficiência. É primordial “que as bibliotecas universitárias ofereçam serviços de informação que satisfaçam as expectativas e necessidades dos usuários.” (CORTES; LOPES, 2008, p.117).

Entretanto, com um volume cada vez maior de informações disponíveis é fundamental que a informação enviada a um usuário seja o mais próximo daquilo que este necessita. Diante disto, “disseminar ou apenas fornecer informação no mundo atual tornou-se uma missão até certo ponto árdua.” (EIRÃO, 2009, p. 21). Com a variedade de ferramentas tecnológicas disponíveis se torna mais difícil fornecer algo desejado e relevante para uma pessoa. Por isso, que

os centros de informação, representados principalmente por bibliotecas, desenvolvem seus serviços visando à satisfação de seu público. É cada vez mais comum o surgimento de serviços/produtos focados na particularidade de cada usuário, desenvolvidos a partir do conceito de segmentação de interesses. (EIRÃO, 2009, p. 21).

Souto e Portela (2003/2004, p. 125) acrescentam que “a Disseminação Seletiva da Informação (DSI) é um serviço que fornece ao usuário uma relação periódica de fontes informacionais, relacionadas com sua área, mediante o preenchimento de um perfil de interesse.”

No caso específico das bibliotecas universitárias os usuários têm necessidades informacionais distintas de acordo com o perfil de cada um e de sua área de atuação. Esses usuários também necessitam estar sintonizados com as publicações de seu interesse e, principalmente, com as novas aquisições efetuadas pela biblioteca. De acordo com Souto e Portela (2003/2004, p. 129), “a aplicação do DSI voltado para a prestação de serviços aos alunos de faculdades e universidades auxilia no suporte de atividades de ensino e pesquisa.”

No cenário atual, em que as pessoas têm cada vez menos tempo, os serviços *online* são os aliados para atender essas demandas, pois com o uso da Internet não existem barreiras de tempo e espaço, e o serviço é amplamente acessível. Além disso, os usuários “para atender às exigências da sociedade atual, estão cada vez mais à procura de praticidade e rapidez, vantagens estas, que as inovações tecnológicas proporcionam.” (CORTES; LOPES, 2008, p.117). E a “[...] Internet trouxe um aumento considerável quanto à oferta de informações e às possibilidades de sua disseminação.” (DIAS, 2005, p.86).

A automatização do envio de informações de acordo com o interesse do usuário proporciona a rapidez e a praticidade almejadas pelos usuários, pois a personalização do serviço e o recebimento automático das informações diminuem o tempo gasto pelo usuário para procurar e selecionar manualmente as informações de que precisa. “O serviço eletrônico de DSI torna os pesquisadores das bibliotecas mais próximos delas, uma vez que passam efetivamente a utilizar seus serviços de maneira personalizada.” (LUCAS; SOUZA, 2007, p. 3).

Os mesmos aparatos tecnológicos da tecnologia da informação, que possibilitaram as mudanças no processo de comunicação científica, permitiram o avanço das técnicas de tratamento das informações e a partir deles é possível desenvolver serviços voltados para atender satisfatoriamente às necessidades dos usuários. (SOUTO; PORTELA, 2003/2004, p. 125).

Segundo Lucas e Souza (2007, p.3), é importante que as bibliotecas ofereçam este tipo de serviço aos seus usuários, pois “com isso pode-se perceber maior agregação de valor aos serviços, mantendo-se a qualidade, a confiabilidade e a competitividade, que hoje são fatores indispensáveis no processo de tomada de decisão de uma biblioteca.”

Além disso, com a necessidade de obter informações cada vez mais rápidas surgem novas metodologias utilizadas para a análise qualitativa e quantitativa, como a mineração de dados, que é uma tecnologia muito útil no gerenciamento de informação. Feraud (2004, p. 39) acrescenta que “um conceito em moda na gestão da informação é o de mineração de dados, ou a extração de dados específicos de enormes campos de informações.”

2.2 MINERAÇÃO DE DADOS

A mineração de dados, conhecida também por *Data Mining*, “como indica o próprio nome, se refere à mineração ou descoberta de novas informações em função de padrões ou regras em grandes quantidades de dados.” (ELMASRI; NAVATHE, 2005, p. 624).

Conforme Tan, Steinbach e Kumar (2009, p. 3), “a mineração de dados é o processo de descoberta automática de informações úteis em grandes depósitos de dados.” Ainda para estes autores,

As técnicas de mineração de dados são organizadas para agir sobre grandes bancos de dados com o intuito de descobrir padrões úteis e recentes que poderiam, de outra forma, permanecer ignorados. Elas também fornecem capacidade de previsão do resultado de uma observação futura, como a previsão de se um cliente recém chegado gastará mais de 100 dólares em uma loja de departamento. (TAN; STEINBACH; KUMAR, 2009, p. 3).

Como os “dados são apenas um conjunto de fatos distintos e objetivos relativos a eventos.” (DAVENPORT; PRUSAK, 1998, p. 2) a mineração de dados torna-se importante para as organizações como uma ferramenta de gerenciamento de informação, na medida em que possibilita a descoberta de informações vantajosas, pois,

Dados descrevem apenas parte daquilo que aconteceu, não fornecem julgamento nem interpretação e nem qualquer base sustentável para a tomada de ação. Embora a

matéria prima do processo decisório possa incluir dados, eles não podem dizer o que fazer. Dados nada dizem sobre a própria importância ou irrelevância. Porém, os dados são importantes para as organizações, porque são matéria prima essencial para a criação da informação. (DAVENPORT; PRUSAK, 1998, p. 3).

Atualmente com a grande disponibilidade de dados na sociedade e indústria de informação a mineração de dados tem atraído a atenção devido à necessidade de transformar dados em informações e em conhecimentos úteis (HAN; KAMBER, 2006). “A mineração de dados é uma tecnologia que combina métodos tradicionais de análise de dados com algoritmos sofisticados para processar grandes volumes de dados.” (TAN; STEINBACH; KUMAR, 2009, p. 1).

No entanto, Tan, Steinbach e Kumar (2009) afirmam que,

A extração de informação útil, entretanto, tem provado ser extremamente desafiadora. Muitas vezes, ferramentas e técnicas tradicionais de análise de dados não podem ser usadas devido ao tamanho do conjunto dos dados ser muito grande. Às vezes a natureza não trivial dos dados significa que as abordagens tradicionais não podem ser aplicadas, mesmo se o conjunto de dados for relativamente pequeno. Em outras situações, questões que precisam ser respondidas não podem ser abordadas usando-se as técnicas existentes para análise de dados e, assim, novos métodos precisam ser desenvolvidos. (TAN; STEINBACH; KUMAR, 2009, p. 1).

O termo mineração de dados vem sendo amplamente discutido para descoberta de conhecimento a partir de dados, tanto que alguns autores optam pelo seu uso como um sinônimo para outro termo conhecido por Descoberta de Conhecimento em Base de Dados ou *Knowledge Discovery in Databases* (KDD). Porém, outros autores, abordam a mineração de dados como uma etapa ou processo essencial do KDD (HAN; KAMBER, 2006).

Desta forma, os autores Han e Kamber (2006) concordam que a mineração de dados é uma etapa do KDD, entretanto devido ao termo estar se popularizando, eles optam pelo conceito de mineração de dados como sendo,

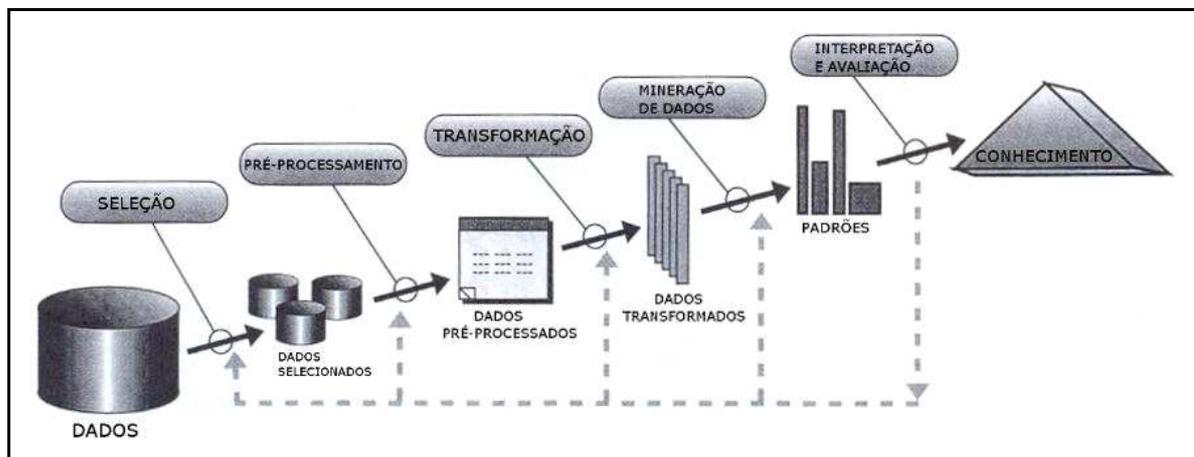
[...] uma visão ampla da funcionalidade de mineração de dados: *data mining* é o processo de descoberta de conhecimento interessante a partir de grandes quantidades de dados armazenados em bases de dados, data warehouses ou repositórios de outras informações. (HAN; KAMBER, 2006, p. 7, tradução nossa)

Já para Tan, Steinbach e Kumar (2009, p. 4), “a mineração de dados é uma parte integral da descoberta de conhecimento em banco de dados (KDD), que é o processo geral de conversão de dados brutos em informações úteis.”

Além disso, Fayyad, Piatetsky-Shapiro e Smyth (1996), acrescentam que o crescimento de empresas, governo e bancos de dados científicos excedeu a capacidade de interpretar e analisar os dados gerados, surgindo assim a necessidade de novas ferramentas e

técnicas para analisar bancos de dados automatizados e inteligentes. Tais ferramentas e técnicas são o foco do KDD, visto como “um processo não trivial de identificação de padrões válidos, novos (desconhecidos até então), potencialmente úteis, e compreensíveis.” (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996, p. 6, tradução nossa).

Assim, o processo de descoberta de conhecimento envolve várias etapas não lineares, onde o processo envolve interação e a iteração significativa, que pode conter *loops* entre duas etapas ou mais, sendo a mineração de dados uma das etapas do KDD (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996). A Figura 1 apresenta as etapas do KDD.



Fonte: adaptado de Fayyad, Piatetsky-Shapiro e Smyth (1996, p.10).

Figura 1 - Etapas do Processo da Descoberta de Conhecimento em Banco de Dados

As etapas do KDD são definidas em:

- dados: desenvolver uma compreensão do domínio da aplicação, o conhecimento prévio relevante, e os objetivos do usuário final;
- seleção: selecionar um conjunto de dados para análise em questão, ou focar em um subconjunto de variáveis ou amostras de dados, em que a descoberta deve ser realizada;
- pré-processamento e limpeza dos dados: operações básicas tais como remover ruídos se necessário, coletar informações necessárias para modelar ou estimar ruídos, escolher estratégias para lidar com campos de dados faltantes, nulos ou repetidos, correção de prováveis erros relevantes para a mineração dos dados;
- transformação: localização de características úteis para representar os dados dependendo do objetivo da tarefa. Utilização de redução de dimensionalidade ou métodos de transformação para reduzir o número de variáveis em consideração;
- mineração de dados: decidir o método e a técnica a ser utilizada para a descoberta de conhecimento, definir o algoritmo que irá realizar a extração dos padrões de

- interesse, ajustando-o se necessário;
- f) interpretação e avaliação (pós-processamento): retorno possível para qualquer um dos passos anteriores para iteração, identificar os padrões de interesse conforme critérios do usuário;
 - g) conhecimento: incorporar o conhecimento obtido à performance do sistema, ou simplesmente documentar e reportar os resultados às partes interessadas, incluindo também a resolução de conflitos em potencial resultantes da mineração de dados (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996).

2.2.1 Tarefas de mineração de dados

Para Han e Kamber (2006), em geral, as tarefas de mineração de dados podem ser classificadas em preditivas e descritivas. Preditivas por realizarem a dedução sobre os dados atuais, a fim de fazer previsões, e descritivas por caracterizarem as propriedades gerais dos dados nas bases de dados.

A classificação é considerada um tipo de tarefa preditiva utilizada em variáveis alvo discretas, cuja tarefa é aprender uma função na qual mapeia cada conjunto de dado em rótulos de classes predefinidas (TAN; STEINBACH; KUMAR, 2009).

Han e Kamber (2006) definem a classificação como sendo o processo para encontrar um modelo ou função que descreva e distinga classes diferentes de dados e conceitos, com a finalidade de ser capaz de usar o modelo para prever a classe de objetos cujo rótulo da classe seja desconhecido.

A partir de um conjunto de treinamento contendo registros cujos rótulos sejam conhecidos é aplicado um algoritmo de aprendizagem que irá construir um modelo de classificação, e posteriormente é aplicado a um conjunto teste, na qual seja possível prever os rótulos dos registros cujas classes sejam desconhecidas. Técnicas como, por exemplo, redes neurais artificiais, algoritmos genéticos e árvores de decisão, podem ser empregados para tarefas de classificação (TAN; STEINBACH; KUMAR, 2009).

A regressão também é considerada uma tarefa preditiva, e se diferencia da classificação, pois é utilizada para variáveis alvo contínuas, prevendo valores reais futuros, como por exemplo, prever o preço futuro de uma ação, considerando que o preço seja um atributo de valor contínuo (TAN; STEINBACH; KUMAR, 2009). Métodos estatísticos e

Redes Neurais podem ser utilizados nestes tipos de tarefas.

Já as tarefas descritivas, segundo Tan, Steinbach e Kumar (2009, p. 8), “objetivam derivar padrões, correlações, tendências, grupos, trajetórias e anomalias, que sintetizam os relacionamentos subjacentes nos dados”. Dentre as tarefas descritivas estão:

- a) análise de agrupamento (*clustering*): busca localizar grupos de observações intimamente relacionados, cujas semelhanças entre si sejam maiores como, por exemplo, agrupar conjuntos de clientes relacionados por um fator em comum;
- b) detecção de anomalias: consiste na tarefa de identificar observações na qual as características sejam expressivamente distintas dos demais dados, aplicações desta tarefa incluem a detecção de fraudes, a detecção de intrusão em redes, padrões incomuns de doenças e perturbações no meio ambiente;
- c) análise de associação: utilizada para descobrir padrões que descrevem características altamente associadas dentro de dados, com o intuito de encontrar os padrões mais interessantes de maneira eficiente como, por exemplo, definir quais os produtos mais frequentemente comprados juntos pelos clientes. Regras de associação é uma técnica que pode ser adotada para este tipo de tarefa (TAN; STEINBACH; KUMAR, 2009, p. 11-13).

Amo (2004, p. 3) afirma que “é importante distinguir o que é uma tarefa e o que é uma técnica de mineração.” Segundo esta autora, a tarefa incide na especificação do que se quer buscar nos dados, que tipo de regularidades ou categoria de padrões interessa encontrar, ou que tipo de padrões poderiam nos surpreender (por exemplo, um gasto exagerado de um cliente de cartão de crédito, fora dos padrões usuais de seus gastos). A técnica de mineração aborda sobre a especificação de métodos que nos garantem como descobrir os padrões que nos interessam.

2.2.2 Técnicas de mineração de dados

Existem várias técnicas para mineração de dados, e a utilização delas depende da tarefa que se pretendem realizar, sendo que cada uma tem sua finalidade específica para um determinado tipo de problema em questão, das quais podemos citar redes neurais artificiais, algoritmos genéticos, árvores de decisão e regras de associação.

O trabalho com redes neurais artificiais tem como motivação o reconhecimento de que

o cérebro humano processa informações de uma forma inteiramente diferente do computador digital convencional, pois,

O cérebro é um ‘computador’ (sistema de processamento de informação) altamente complexo, não-linear e paralelo. Ele tem a capacidade de organizar seus constituintes estruturais, conhecidos por neurônios, de forma a realizar certos processamentos como, por exemplo, reconhecer padrões, percepção e controle remoto, muito mais rapidamente que o mais rápido computador digital hoje existente. (HAYKIN, 2001, p. 27-28).

Uma rede neural artificial é vista como “uma máquina que é projetada para modelar a maneira como o cérebro realiza uma tarefa particular ou função de interesse, sendo a rede normalmente implementada utilizando-se componentes eletrônicos ou é simulada por programação em um computador digital.” (HAYKIN, 2001, p. 27-28).

Conforme Elmasri e Navathe (2005, p. 640), “as redes neurais são auto-adaptativas, isto é, elas aprendem de informações sobre um problema específico, entretanto as saídas são altamente quantitativas e não são fáceis de entender, elas geralmente têm também problemas para modelar dados de séries temporais.”

Os algoritmos genéticos tentam encontrar a melhor solução para problemas de otimização e busca com base nas idéias da seleção natural (HAN; KAMBER, 2006). É uma classe de procedimentos para busca randômica, que realiza buscas adaptativas e robustas sobre uma vasta faixa de topologias de espaço de busca, e são também utilizados para solução e agrupamentos de problemas. Possuem uma habilidade para resolver problemas em paralelo e fornece uma poderosa ferramenta para mineração de dados. Entretanto, uma desvantagem apresentada é a alta demanda de processamento computacional imprescindível para localizar alguma coisa com significado (ELMASRI; NAVATHE, 2005).

Árvore de decisão “é um fluxograma como a estrutura de uma árvore, onde cada nó representa um teste em um valor de atributo, e cada ramo representa um resultado do teste, e as folhas da árvore representam classes ou distribuições de classe.” (HAN; KAMBER, 2006, p. 24, tradução nossa).

As regras de associação permitem “encontrar padrões, associações ou correlações em conjuntos de itens (objetos) de uma base de dados transacional, relacional ou de outros tipos de repositórios de informação.” (NEVES, 2002, p. 6). Por ser o foco deste trabalho, esta técnica será mais detalhada na seção 2.2.4.

2.2.3 Aplicações da mineração de dados

A mineração de dados utiliza-se de um conjunto de disciplinas, incluindo sistemas de banco de dados, estatística, aprendizagem de máquina, visualização e ciência da informação (HAN; KAMBER, 2006). Dependendo do tipo de dados a ser extraído ou do aplicativo de mineração de dados, o sistema pode incluir técnicas “de análise de dados espaciais, de recuperação de informação, reconhecimento de padrões, análise de imagens, processamento de sinal, computação gráfica, tecnologia *web*, economia, negócios, a bioinformática, ou psicologia.” (HAN; KAMBER, 2006, p. 29, tradução nossa).

As tecnologias de mineração de dados podem ser aplicadas em grande variedade de contextos de tomada de decisões empresariais, dentre os quais Elmasri e Navathe (2005) destacam:

- a) marketing, em aplicações como análise de comportamento do consumidor baseados em padrões de consumo, definição de estratégias de marketing incluindo propaganda, localização de lojas e mala direta direcionada, segmentação de clientes, lojas e produtos;
- b) finanças, em aplicações como análise de crédito de clientes, segmentação de contas a receber, análise de desempenho de investimentos financeiros como ações, avaliação de opções de financiamentos e detecção de fraudes;
- c) produção, em aplicações de otimização de recursos como máquinas, força de trabalho e materiais, projeto ótimo de fabricação, layouts em chão de fábrica e projeto de produtos baseados nos requisitos dos clientes;
- d) saúde, em aplicações de descoberta de padrões em imagens radiológicas, análise de dados experimentais para relação com doenças, análise de efeitos colaterais de remédios.

Han e Kamber (2006) afirmam que a mineração de dados possa ser utilizada para controle de qualidade, marketing, comércio, análise de dados financeiros, indústria das telecomunicações, medicina, aplicações e pesquisas científicas diversas, detecção de fraudes e intrusão.

São muitos os contextos em que a mineração de dados pode ser empregada. Podemos destacar também outras utilizações, a exemplo de Machado e Becker (2002) que propõem um estudo voltado para o uso da mineração de dados na *web* aplicado a um ambiente de ensino a distância, onde a partir da análise comportamental do usuário possam ser extraídos padrões de

navegação, e possibilitar melhorias e adequações em *websites* de ensino.

Jesus (2004) propõe a utilização de mineração de dados para a identificação de perfis de usuários de bibliotecas, visando à personalização dos serviços de recuperação e de disseminação da informação. Neste sistema, o histórico de empréstimos, consultas e reservas dos usuários da biblioteca foram armazenados e as técnicas de mineração de dados foram aplicadas para identificar o perfil dos usuários e sugerir obras de forma personalizada a partir das transações realizadas.

2.2.4 Regras de associação

Com o aumento significativo da quantidade de registros armazenados em bancos de dados extrair informações úteis a partir dessas bases, para que sirva de apoio na tomada de decisões para organizações, vem se tornando uma tarefa mais complexa, e para tal, surgem técnicas para auxiliar nessa exploração, como por exemplo, regras de associação, que visa detectar na base itens ditos frequentes e que apresentem associações conjuntas.

Elmasri e Navathe (2005, p. 627) defendem que “uma das principais tecnologias em mineração de dados envolve a descoberta das regras de associação ou conjunto de itens frequentes. Um banco de dados é considerado uma coleção de transações, cada um envolvendo um conjunto de itens.”

A “mineração de itens frequentes leva à descoberta de associações e correlações entre os itens de grandes conjuntos de dados relacionais ou transacionais.” (HAN; KAMBER, 2006, p. 227, tradução nossa).

Um exemplo comum da utilização para mineração de regras de associação na descoberta de padrões através das compras realizadas, transações feitas, por clientes em supermercados ou lojas, na qual uma mercadoria está relacionada com outra, “[...] conhecida também como transações de cestas de compras.” (TAN; STEINBACH; KUMAR, 2009, p. 389). A partir desses padrões encontradas, o estabelecimento pode realizar promoções de vendas em determinados produtos ou também conhecer o perfil de seus consumidores, por exemplo.

No Quadro 1, podemos observar um exemplo, onde cada linha representa uma transação, seu identificador e o conjunto de itens comprados por um determinado consumidor.

TID	Itens
1	Pão, Café
2	Suco, Fraldas, Cerveja, Ovos
3	Fraldas, Cerveja, Refrigerante
4	Pão, Leite, Fraldas, Cerveja
5	Café, Leite, Fraldas, Refrigerante

Fonte: adaptado de Tan, Steinbach e Kumar (2009, p. 390).

Quadro 1 - Exemplo de transações realizadas

No exemplo citado acima, observa-se que existe uma forte relação entre fraldas e cerveja, pois do total de cinco transações, quatro delas apresentam o item fraldas, e dessas quatro transações, três possuem o item cerveja. Com base nessas informações, poderíamos afirmar que, dos consumidores que compram fraldas, 75% deles também tendem a comprar cerveja.

A representação binária, zero ou um, é outra forma de apresentar os dados, onde cada linha é equivalente a uma transação e cada coluna a um item, e os itens por sua vez assumem valores como verdadeiro (um) se estiver presente na transação ou falso (zero) se não estiver (TAN; STEINBACH; KUMAR, 2009). Observa-se um exemplo da representação binária no Quadro 1, com base no exemplo citado no Quadro 2.

TID	Pão	Café	Suco	Fraldas	Cerveja	Ovos	Leite	Refrigerante
1	1	1	0	0	0	0	0	0
2	0	0	1	1	1	1	0	0
3	0	0	0	1	1	0	0	1
4	1	0	0	1	1	0	1	0
5	0	1	0	1	0	0	1	1

Fonte: adaptado de Tan, Steinbach e Kumar (2009, p. 391).

Quadro 2 - Exemplo de representação binária de transações realizadas

“Uma propriedade importante de um conjunto de itens é o contador de suporte, que se refere ao número de transações que contém um determinado conjunto de itens.” (TAN; STEINBACH; KUMAR, 2009, p. 392) como, no exemplo apresentado no Quadro 2, {fraldas} tem contador do suporte igual a quatro, pois está presente em quatro das cinco transações.

Formalizando o problema para mineração de regras de associação conforme Agrawal et al (1996), seja $I = \{i_1, i_2, \dots, i_m\}$ um conjunto de itens e D um conjunto de transações, onde cada transação T é um conjunto de itens (*itemset*) tal que $T \subset I$. Associado a cada transação está um único identificador, chamado identificador único de transação (TID). Assim, uma regra de associação é uma implicação na forma expressa por $X \Rightarrow Y$ (X então Y), sendo X

denominado de antecedente e Y conseqüente da regra, onde X e Y são *itemsets* distintos, representados na expressão $X \subset I, Y \subset I$ e $X \cap Y = \emptyset$.

Regras de associação basicamente têm duas medidas utilizadas para análise, o suporte e a confiança. “O suporte determina a freqüência na qual uma regra é aplicável a um determinado conjunto de dados, e a confiança determina a freqüência na qual os itens em Y aparecem em transações que contenham X .” (TAN; STEINBACH; KUMAR, 2009, p. 392).

Dessa maneira, diz-se que, “a regra $X \Rightarrow Y$, tem suporte s no conjunto de transações D , se $s\%$ das transações em D contém $X \cup Y$. A regra $X \Rightarrow Y$ pertence ao conjunto de transações D , com confiança c se $c\%$ das transações em D que contém X também contém Y .” (AGRAWAL et al., 1996, p. 308, tradução nossa).

O suporte da regra $X \Rightarrow Y$ é definido na expressão como $sup(X \Rightarrow Y) = \sigma(X \cup Y) / D$, ou seja, é a razão entre o contador de suporte contendo X e Y e o total de transações. E a confiança da regra $X \Rightarrow Y$ é definida na expressão como $conf(X \Rightarrow Y) = \sigma(X \cup Y) / \sigma(X)$, é a razão entre o contador do suporte contendo X e Y e o contador do suporte contendo X (TAN; STEINBACH; KUMAR, 2009).

Baseando-se no Quadro 2, e retomando como exemplo a regra $\{fraldas\} \Rightarrow \{cerveja\}$ teríamos a representação em números para encontrar o suporte e a confiança desta regra, conforme o Quadro 3.

<p>Regra: $\{fraldas\} \Rightarrow \{cerveja\}$</p> <p>Suporte da regra: $\{fraldas\} = 4$ Total de transações = 5 $4 / 5 = 0,8$ (ou 80% de suporte em relação ao banco de dados)</p> <p>Confiança da regra: $\{cerveja\} = 3$ $\{fraldas\} = 4$ $3/4 = 0,75$ (ou 75% de grau de confiança para esta regra)</p>

Quadro 3 - Exemplo cálculo para encontrar suporte e confiança de uma regra

Uma regra de baixo suporte pode indicar uma simples coincidência e também não ser interessante dependendo do problema em questão, sendo o suporte muitas vezes utilizado para eliminar regras sem interesse. Já a confiança, mede a confiabilidade da inferência feita por uma regra, seja uma regra $X \Rightarrow Y$, quanto maior a confiança, maior a probabilidade de que Y

esteja presente em transações que possuam X , fornecendo também uma avaliação da probabilidade condicional de Y dado X (TAN; STEINBACH; KUMAR, 2009).

Tan, Steinbach e Kumar (2009) afirmam que, minerar regras de associação pode ser um processo computacional muito custoso, visto que, mesmo um pequeno conjunto de dados pode gerar um número muito grande de regras, já que cada item se relaciona com os demais formando vários conjuntos.

Para que uma regra seja interessante, ou dita forte, deve atender os critérios, tenha suporte $\geq \textit{minsup}$ e a confiança $\geq \textit{minconf}$, sendo *minsup* e *minconf* os limites de suporte e confiança respectivamente definidos (TAN; STEINBACH; KUMAR, 2009).

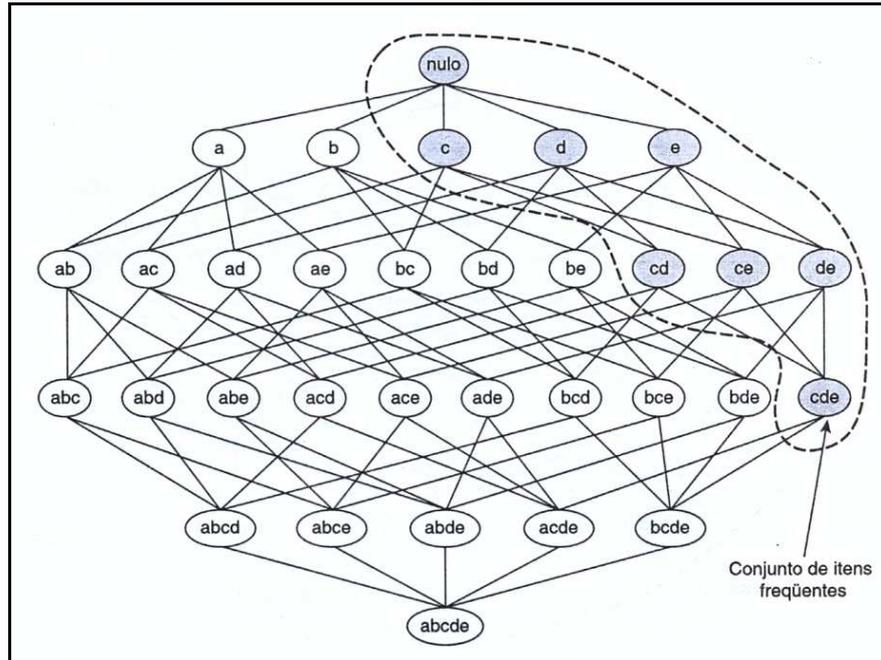
Então, a estratégia utilizada é decompor o problema em duas etapas:

- a) encontrar todos os conjuntos de itens que tenham o seu suporte maior que o suporte mínimo estabelecido, estes conjuntos são chamados de conjuntos de itens freqüentes;
- b) após encontrar os conjuntos de itens freqüentes, gerar as regras que atendam a um mínimo de confiança (AGRAWAL et al., 1996).

Como o número de conjunto de itens distintos é $2^k - 1$, e k é o número de itens, cresce exponencialmente, para reduzir o número do conjunto de itens candidatos utiliza-se o contador de suporte (TAN; STEINBACH; KUMAR, 2009).

Segundo o princípio de Apriori, sendo um conjunto de itens considerado frequentes $\{X, Y, Z\}$, por exemplo, então todos os seus subconjuntos $\{X\}$, $\{Y\}$, $\{Z\}$, $\{X, Y\}$, $\{X, Z\}$ e $\{Y, Z\}$ também serão frequentes (TAN; STEINBACH; KUMAR, 2009).

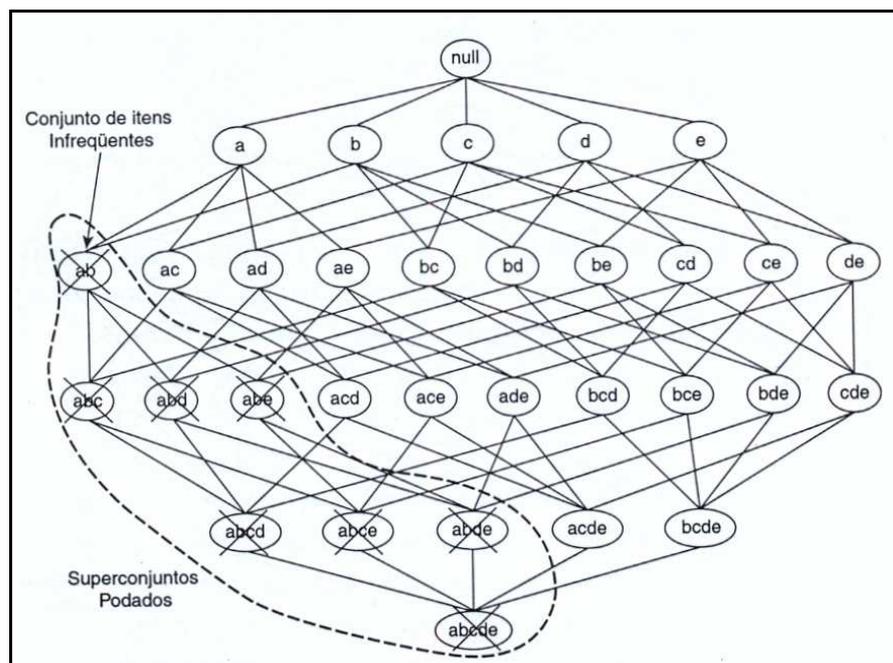
A Figura 2 apresenta um exemplo do princípio de Apriori para conjuntos de itens frequentes.



Fonte: Tan, Steinbach e Kumar (2009, p. 397).

Figura 2 - Exemplo de conjuntos de itens freqüentes conforme princípio Apriori

A recíproca também é verdadeira, se um conjunto não for freqüente, seus subconjuntos também não serão. E através de uma técnica baseada no suporte para efetuar a poda de conjunto de itens que não interessam, diminuindo a busca exponencial (TAN; STEINBACH; KUMAR, 2009), conforme o exemplo na Figura 3.



Fonte: Tan, Steinbach e Kumar (2009, p. 398).

Figura 3 - Exemplo da redução de conjuntos de itens baseado na poda

Existem diversos algoritmos desenvolvidos, dentre eles, o algoritmo conhecido por Apriori, considerado “o primeiro algoritmo de mineração de regras de associação que foi precursor do uso de poda baseado em suporte para controlar de forma sistemática o crescimento exponencial dos conjuntos de itens candidatos.” (TAN; STEINBACH; KUMAR, 2009, p. 399).

O algoritmo é dividido em duas etapas e funciona iterativamente de modo a encontrar conjuntos de itens frequentes em uma determinada base de dados para que posteriormente sejam extraídas as regras de associação.

Na primeira etapa, inicia com todos os itens da base como sendo um conjunto de um item, é feito a contagem do suporte para cada conjunto, são eliminados os conjuntos que não atendem ao suporte mínimo definido, restando somente os conjuntos de um item ditos frequentes. A partir desse conjunto, são gerados novos conjuntos candidatos, contendo agora dois itens cada conjunto, verifica-se novamente o suporte e em seguida são eliminados novamente aqueles, cujo suporte é inferior ao mínimo definido. Assim, o processo é repetido n vezes até não haver mais conjuntos de itens candidatos, resultando nos conjuntos contendo os itens mais frequentes da base de dados (TAN; STEINBACH; KUMAR, 2009).

O Quadro 4 apresenta um pseudocódigo do algoritmo Apriori para encontrar conjuntos de itens frequentes, adotando-se como notação para descrevê-lo, seja *itemset* um conjunto de itens, *k-itemset* um conjunto contendo k itens, L_k conjunto de *k-itemsets* frequentes e C_k conjunto de *k-itemsets* candidatos (AGRAWAL et al., 1996).

```

 $L_1 = \{1\text{-itemsets frequentes}\};$ 
for (  $k = 2; L_{k-1} \neq \emptyset; k++$  ) do begin
     $C_k = \text{apriori-gen}(L_{k-1});$  //Novos conjuntos candidatos
    forall (transações  $t \in D$ ) do begin
         $C_t = \text{subset}(C_k, t);$  //Conjutos candidatos contidos em  $t$ 
        forall (candidatos  $c \in C_t$ ) do
             $c.\text{count}++;$ 
        end
         $L_k = \{c \in C_k / c.\text{count} \geq \text{minsup}\}$ 
    end
 $\text{Resposta} = \bigcup_k L_k;$ 

```

Fonte: adaptado de Agrawal et al (1996, p. 312).

Quadro 4 - Pseudocódigo do algoritmo Apriori para encontrar conjuntos de itens frequentes

Encontrados então os conjuntos de itens frequentes, na segunda etapa são geradas as regras de associação candidatas e extraídas as regras fortes que atendam a um suporte mínimo de confiança utilizado como parâmetro.

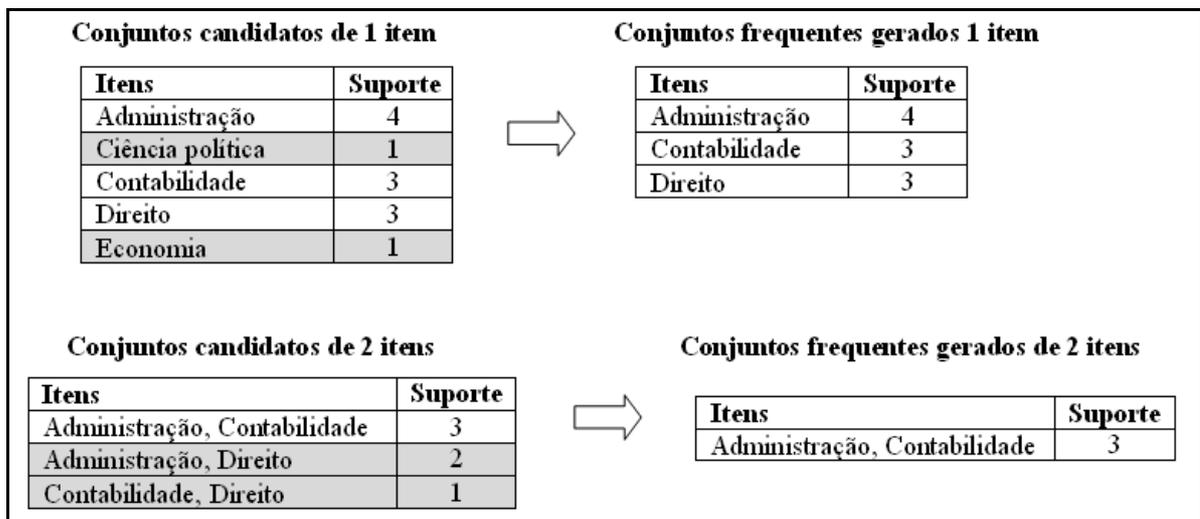
O algoritmo Apriori usa uma abordagem de níveis para gerar regras de associação, onde cada nível corresponde ao número de itens que pertencem ao conseqüente da regra. Inicialmente, todas as regras de confiança alta que tenham apenas um item no conseqüente da regra são extraídas. Estas regras são então usadas para gerar novas regras candidatas. (TAN; STEINBACH; KUMAR, 2009, p. 417).

A seguir é feita uma demonstração básica de como seria a extração de regras de nível 1 em um conjunto de transações, conforme os conceitos vistos anteriormente. Seja uma base de dados de transações qualquer, onde os itens são algumas áreas (Quadro 5), deseja-se encontrar itens frequentes para extrair regras, cujo valor do suporte mínimo deve ser de 50% ou mais, (contador do suporte maior ou igual a 3), e a confiança mínima da regra 70% ou superior.

TID	Itens
1	Administração, Contabilidade
2	Direito, Administração, Contabilidade
3	Administração, Economia, Direito
4	Administração, Contabilidade
5	Direito, Ciência política

Quadro 5 - Exemplo de uma base de dados com transações

São gerados os conjuntos candidatos e encontrados os conjuntos de itens frequentes dessa base, conforme Quadro 6.



Quadro 6 - Geração dos conjuntos de itens frequentes

Com base nos conjuntos de itens frequentes gerados, são extraídas então as regras de associação que possuem relação com outros itens e que satisfaçam a confiança mínima definida para o problema em questão, de acordo com o Quadro 7.

Conjuntos de itens frequentes gerados	
Conjuntos	Suporte
Administração	4 (80%)
Contabilidade	3 (60%)
Direito	3 (60%)
Administração, Contabilidade	3 (60%)

Regras de associação geradas	
Regras geradas	Confiança
Administração => Contabilidade	$3 / 4 = 0,75$ (75%)
Contabilidade => Administração	$3 / 3 = 1$ (100%)

Quadro 7 - Exemplo da extração de regras de associação com base no algoritmo Apriori

Das duas regras de associação geradas, podemos verificar que com relação à base de dados mencionada, a primeira regra mostra que 75% das transações onde há Administração, encontramos Contabilidade. Já para a outra regra, o nível da confiança é ainda maior, pois 100% das transações contendo Contabilidade também há Administração.

Assim podemos perceber a correlação entre os itens de uma determinada base de dados contendo dados de transações, e a incidência em que um item pode apresentar em relação a outro, sugerindo então informações de tendências que podem ser aproveitadas nas organizações de diversas formas.

2.3 TRABALHOS CORRELATOS

Jesus (2004), em sua dissertação para o mestrado em Ciências da Computação da Universidade Federal de Santa Catarina propõe o desenvolvimento de um *data warehouse* para armazenar o histórico de empréstimos, consultas e reservas dos usuários da biblioteca, e a aplicação de técnicas de mineração de dados para identificar o perfil dos usuários e possibilitar a personalização de um sistema *web* para recuperação e disseminação de

informações ao usuário, sendo que o sistema irá sugerir obras de forma personalizada a cada usuário de acordo com seu perfil gerado a partir das transações realizadas. Os usuários da aplicação foram alunos e professores da pós-graduação da BU/FURB, e os dados utilizados para mineração em relação aos usuários eram dados parciais do ano de 2003. Para o desenvolvimento do sistema, foi utilizada linguagem de programação *Hypertext Preprocessor* (PHP) e Linguagem de Marcação de Hipertexto (HTML), com hospedagem em um servidor *Apache* com plataforma *Windows 2000 Server* e banco de dados *Oracle 8i*.

Silva (2004), em sua dissertação para o mestrado em Ciências da Computação da Universidade Federal do Rio Grande do Sul, cujo objetivo foi estudar os fundamentos da mineração de dados com foco na técnica de regras de associação e desenvolver um protótipo baseado no algoritmo Apriori e seus conceitos, e a partir do protótipo aplicado na Secretaria Municipal de Saúde de Londrina no Paraná, analisar uma base de dados de saúde relacionando características sócio-econômicas com dados de procedimentos realizados em internações hospitalares. O protótipo foi desenvolvido utilizando linguagem *Delphi*, e banco de dados *Microsoft Access*.

O artigo publicado por Lucas e Souza (2007), apresenta uma proposta de disseminação seletiva da informação em bibliotecas universitárias em um estudo baseado em *Customer Relationship Management* (CRM), visando à melhoria de processos nas BUs e a contribuição com a satisfação de seus usuários. Como resultado de um estudo desenvolvido por meio de pesquisas bibliográficas, análise documental e estudos de caso sob o prisma do CRM, sendo assim, o artigo aborda conceitualmente o assunto disseminação seletiva da informação, citando algumas sugestões para implementação de sistemas, como por exemplo, a utilização de *websites* com conteúdo personalizado para cada tipo de perfil de usuário.

3 DESENVOLVIMENTO DO SISTEMA

Este capítulo abrange a metodologia adotada para o desenvolvimento do sistema. Inicialmente será abordado o levantamento de informações e os requisitos funcionais e não funcionais que devem ser atendidos no sistema. Logo após, será apresentada a especificação do sistema sob a forma de diagramas da *Unified Modeling Language* (UML). Posteriormente, a implementação com as técnicas e ferramentas utilizadas e a operacionalidade de implementação, e por fim os resultados obtidos.

3.1 LEVANTAMENTO DE INFORMAÇÕES

O principal objetivo do sistema desenvolvido neste trabalho é promover a disseminação de informações referentes às novas aquisições da coleção geral, coleção constituída por livros, para os usuários da BU/FURB que utilizam o serviço de empréstimos.

A BU/FURB tem como missão,

[...] desenvolver e colocar à disposição da comunidade universitária um acervo bibliográfico e serviços que atendam às necessidades de informação para as atividades de ensino, pesquisa e extensão, adotando modernas tecnologias para o tratamento, recuperação e transferência da informação (UNIVERSIDADE REGIONAL DE BLUMENAU, 2006b).

O Sistema de Automação de Bibliotecas (SABIBLIO), que é o sistema para gerenciamento da BU, atende a todas as unidades e é composto por vários módulos, dentre eles:

- a) módulo de controle de aquisição: controla e executa as rotinas de cada etapa da sugestão e aquisição de obras;
- b) módulo de recuperação da informação: visa a recuperar informações, desde livros, teses, normas técnicas, artigos de periódicos, títulos de periódicos, materiais especiais, bem como informações dos usuários do sistema;
- c) módulo de controle de empréstimos: executa as rotinas de empréstimo, devolução e reserva de material bibliográfico;
- d) módulo de processamento técnico: executa as rotinas de cadastramento das informações bibliográficas, de autoridades (nomes) e de assuntos;

- e) módulo de parametrização: permite a manutenção dos parâmetros do sistema (UNIVERSIDADE REGIONAL DE BLUMENAU, 2006c).

Toda obra que é adquirida pela BU/FURB é recebida na Seção de Seleção e Aquisição e em seguida é encaminhada para a Seção de Processamento Técnico, onde é catalogada no módulo de processamento técnico do SABIBLIO. Cada obra recebe um número único de registro no sistema que o identifica, chamado *Master File Number* (MFN). Os diversos exemplares de uma obra também recebem um registro para diferenciá-los como exemplar e possibilitar o empréstimo destes.

O processo técnico envolve a catalogação (descrição bibliográfica do material), a classificação e a indexação da obra. Para isso, são adotados padrões internacionais de catalogação, como o Código de Catalogação Anglo-americano (ACCR2) e o formato *Machine Readable Cataloging* (MARC).

O formato MARC define uma estrutura para registros bibliográficos que permite o seu armazenamento em meio magnético e tratamento adequado pelos sistemas de computador, contendo recursos que permitem ao computador identificar e interpretar a informação contida em um registro catalográfico. (FUNDAÇÃO GETULIO VARGAS, 2010).

O formato MARC (Anexos B e C) permite ainda o intercâmbio de registros bibliográficos e catalográficos entre bibliotecas que utilizam esse padrão, já que os dados estão no mesmo formato, os registros podem ser importados ou exportados no sistema, poupando tempo para os profissionais que fazem a catalogação.

A classificação da obra é feita de acordo com seu assunto, e para tal a BU adota a Classificação Decimal de Dewey (CDD), que é um sistema de classificação documentária que organiza todo o conhecimento em dez classes principais e atribui números decimais para suas categorias. Desta forma cada número representa um assunto e este aumenta decenalmente à medida que o assunto se torna mais específico (DEWEY, 1979). A cada obra é atribuído somente um número que representa o seu assunto, assim cada obra terá apenas um número de classificação.

Ainda, há o módulo *web*, neste caso o *website* da BU, e que permite o acesso às informações sobre os recursos e o fornecimento de serviços na Internet, como consulta ao catálogo do acervo, renovação de empréstimos, reserva de material, sugestão de aquisição, alteração de senha e orientações diversas (UNIVERSIDADE REGIONAL DE BLUMENAU, 2006c).

Atualmente, um dos serviços *online* disponibilizados no *website* da biblioteca é uma

lista de novas aquisições que pode ser consultada através de uma busca na qual o usuário informa um período, que compreende uma data inicial e uma data final, e o tipo de material, permitindo que seja realizada uma consulta.

Diante do exposto, a aplicação desenvolvida neste trabalho agrega valor aos serviços oferecidos, pois ao permitir a disseminação seletiva da informação dos livros adquiridos recentemente pela BU/FURB irá minimizar o tempo e os esforços dos usuários que receberão automaticamente as informações que desejam, mantendo-os informados frequentemente das aquisições incorporadas recentemente ao acervo da BU.

O sistema desenvolvido divide-se basicamente em dois módulos, sendo um módulo uma interface via *web* para interação dos usuários da biblioteca e o outro um módulo interno com interface gráfica, ou *Graphical User Interface* (GUI), para gerenciamento da aplicação por parte do administrador e disseminação das informações apuradas de acordo com os perfis cadastrados.

Quanto ao módulo *web*, foi desenvolvida uma página de Internet para que o usuário possa configurar seu perfil, e que futuramente esta possa ser acoplada ao atual *website* da BU/FURB.

No seu perfil o usuário poderá selecionar quais as áreas do conhecimento lhe interessam de acordo com uma lista de áreas predefinidas pela biblioteca conforme a classificação da obra (Anexo A). O usuário terá também a opção de autorizar ou não autorizar o envio destas informações para o seu *e-mail*.

O módulo interno divide-se na ferramenta de envio e na ferramenta gerencial. A ferramenta de envio é responsável por processar as informações, verificando as áreas selecionadas no perfil de cada usuário, e buscando as obras ou livros que foram incorporadas recentemente no acervo bibliográfico, conforme cada área selecionada.

Ainda, incrementando a proposta inicial do trabalho, também são enviadas sugestões de novas aquisições incorporadas ao acervo, através da utilização de regras de associação, onde é feito uma busca por novas aquisições que tenham correlação com as áreas selecionadas pelo usuário em seu perfil. Essas regras de associação são extraídas com base no histórico de empréstimos feitos pelos usuários dentro de um período de um ano.

Por fim, o sistema envia um relatório personalizado sobre as novas aquisições para o *e-mail* do usuário, com as informações filtradas de acordo com as áreas escolhidas previamente por ele e também sugestões, encontradas pelo sistema a partir das regras de associações geradas, que possam também interessar ao usuário.

As informações que os usuários receberão em seus *e-mails* serão referentes às obras,

tais como título e autor, e ainda o *link* que remete para a consulta *online* no atual *website* da biblioteca, facilitando a busca do material para o usuário, que poderá visualizar informações mais detalhadas de cada obra, como disponibilidade, localização e biblioteca a qual pertence o material.

A ferramenta de envio é executada através de um evento do agendador de tarefas do sistema operacional em um determinado horário agendado, pelo menos uma vez por semana, de modo que, sejam executadas todas as rotinas necessárias para a disseminação das informações encontradas, conforme mencionado anteriormente. Deste modo, promove o envio de relatórios personalizados para o *e-mail* do usuário referentes às informações apuradas das aquisições incorporadas recentemente ao acervo bibliográfico, de acordo com cada área selecionada em seu perfil cadastrado.

Já a ferramenta gerencial, tem como finalidade auxiliar o administrador do sistema, quanto à parametrização de confiança para geração das regras de associação e a quantidade de dias retroativos que será feito a busca por novas aquisições, consultar informações das áreas cadastradas no módulo *web* e consultar as regras de associação por área, também permitirá visualizar um relatório estatístico de utilização do sistema, como por exemplo, a quantidade de usuários que está utilizando o sistema e o ranking das áreas mais selecionadas no módulo *web* pelos usuários.

Para o desenvolvimento da interface *web* foi adotada linguagem de programação *Hypertext Preprocessor* (PHP) com Linguagem de Marcação de Hipertexto (HTML), com hospedagem em um servidor *Apache*. O módulo interno foi desenvolvido em linguagem de programação *Delphi*.

Para o armazenamento dos dados foi utilizado *Oracle 10g*, com base na estrutura de tabelas atualmente utilizadas na biblioteca, com algumas alterações, para que o sistema desenvolvido tivesse portabilidade em uma integração futura.

O Quadro 8 apresenta os requisitos funcionais previstos para o sistema e sua rastreabilidade, ou seja, vinculação com os casos de uso associados.

REQUISITOS FUNCIONAIS	CASO DE USO
RF01: O sistema deverá permitir que o usuário faça <i>login</i> .	UC01.01
RF02: O sistema deverá permitir ao usuário selecionar as áreas de preferência em seu perfil.	UC01.02
RF03: O sistema deverá permitir ao usuário alterar a seleção das áreas em seu perfil.	UC01.03
RF04: O sistema deverá permitir ao usuário autorizar ou não o envio de informações para seu <i>e-mail</i> .	UC01.04
RF05: O sistema deverá permitir enviar relatórios personalizados das informações apuradas.	UC02.01
RF06: O sistema deverá permitir verificar as áreas selecionadas no perfil do usuário para poder encontrar novas aquisições.	UC02.02
RF07: O sistema deverá permitir efetuar a busca por área de novas aquisições incorporadas ao acervo em um determinado período.	UC02.05
RF08: O sistema deverá permitir gerar sugestões através de regras de associação, verificando quais são as áreas correlatas em relação ao perfil do usuário.	UC02.03
RF09: O sistema deverá permitir gerar relatório com as novas aquisições encontradas conforme as áreas do perfil e das sugestões baseadas em regras de associação.	UC02.04
RF10: O sistema deverá permitir ao administrador alterar os parâmetros de confiança das regras de associação e a quantidade de dias retroativos que será feito a busca por novas aquisições.	UC03.01
RF11: O sistema deverá permitir ao administrador consultar informações das áreas que estão cadastradas no módulo <i>web</i> .	UC03.02
RF12: O sistema deverá permitir ao administrador consultar quais são as regras de associação por área, ou seja, a relação de uma área com outra.	UC03.03
RF13: O sistema deverá permitir ao administrador visualizar um relatório estatístico de utilização do sistema, com informações sobre de quantidade de usuários cadastrados e usuários que utilizam o sistema e o <i>ranking</i> das áreas mais selecionadas no módulo <i>web</i> pelos usuários.	UC03.04

Quadro 8 - Requisitos funcionais

O Quadro 9 apresenta a lista dos requisitos não funcionais previstos para o sistema.

REQUISITOS NÃO FUNCIONAIS
RNF01: O módulo <i>web</i> do sistema deverá ser desenvolvido na linguagem <i>PHP</i> e <i>HTML</i> basicamente.
RNF02: O módulo <i>web</i> do sistema deverá ter compatibilidade com os navegadores <i>Internet Explorer 7.0</i> ou superior e <i>Mozilla Firefox 3.5.x</i> ou superior.
RNF03: O módulo interno do sistema deverá ser desenvolvido na linguagem <i>Delphi</i> .
RNF04: O sistema deverá suportar o banco de dados <i>Oracle 10g</i> .
RNF05: O sistema deverá possuir interface <i>web</i> para acesso do usuário.
RNF06: O módulo <i>web</i> deverá ser hospedado em um servidor <i>Apache</i> .
RNF07: O módulo interno deverá ter como sistema operacional base a plataforma <i>Windows 32 bits</i> .
RNF08: O agendador de tarefas do sistema operacional deverá disparar um evento para executar a ferramenta de envio de informações no horário agendado em um determinado dia da semana.

Quadro 9 - Requisitos não funcionais

3.2 ESPECIFICAÇÃO

Nesta seção estão descritos os diagramas e modelos elaborados para desenvolvimento do sistema, dentre eles: os principais diagramas de caso de uso e seu detalhamento e o diagrama de atividades, ambos com base na *Unified Modeling Language* (UML), utilizando a ferramenta *Enterprise Architect 7*, e a Modelagem Entidade e Relacionamento (MER), desenvolvido no *DB Designer Fork*.

3.2.1 Diagrama de caso de uso

Os principais diagramas de caso de uso do sistema estão dispostos a seguir, sendo que os detalhes dos casos de uso mais relevantes estão descritos no Apêndice A.

Na Figura 4, tem-se o diagrama de caso de uso do módulo *web* para o usuário do sistema.

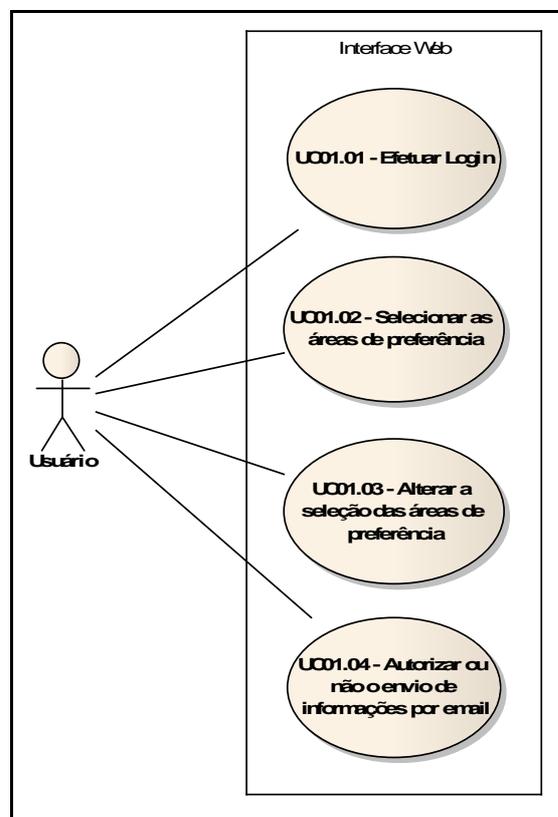


Figura 4 - Diagrama de caso de uso da interface *web* para o usuário do sistema proposto

Na Figura 5, tem-se o diagrama de caso de uso do módulo interno da ferramenta de envio.

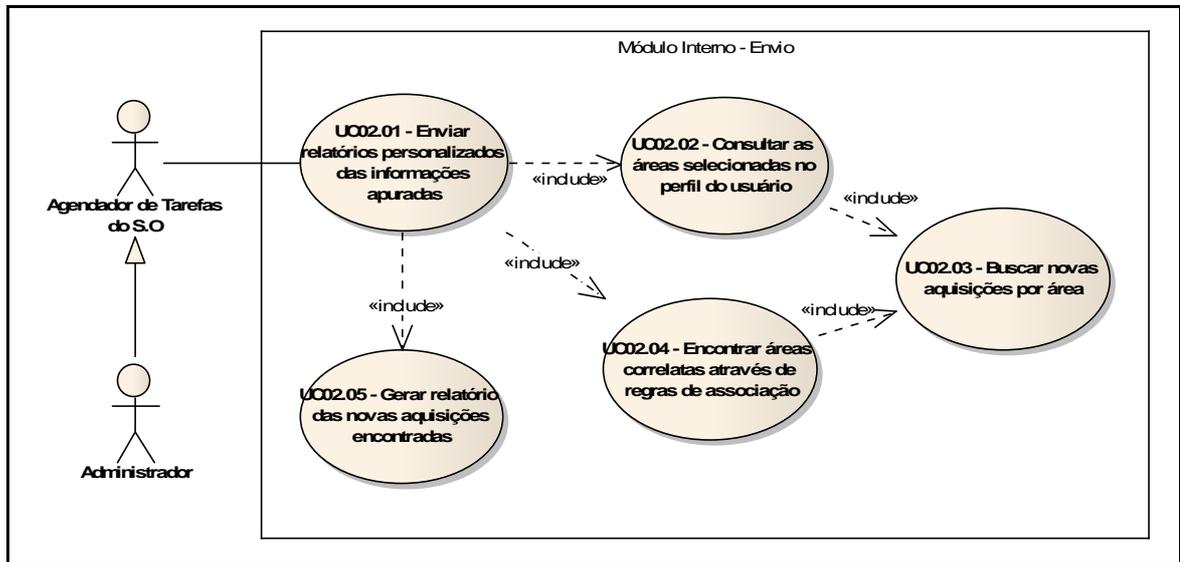


Figura 5 - Diagrama de caso de uso da ferramenta de envio

Na Figura 6, tem-se o diagrama de caso de uso do módulo interno da ferramenta gerencial.

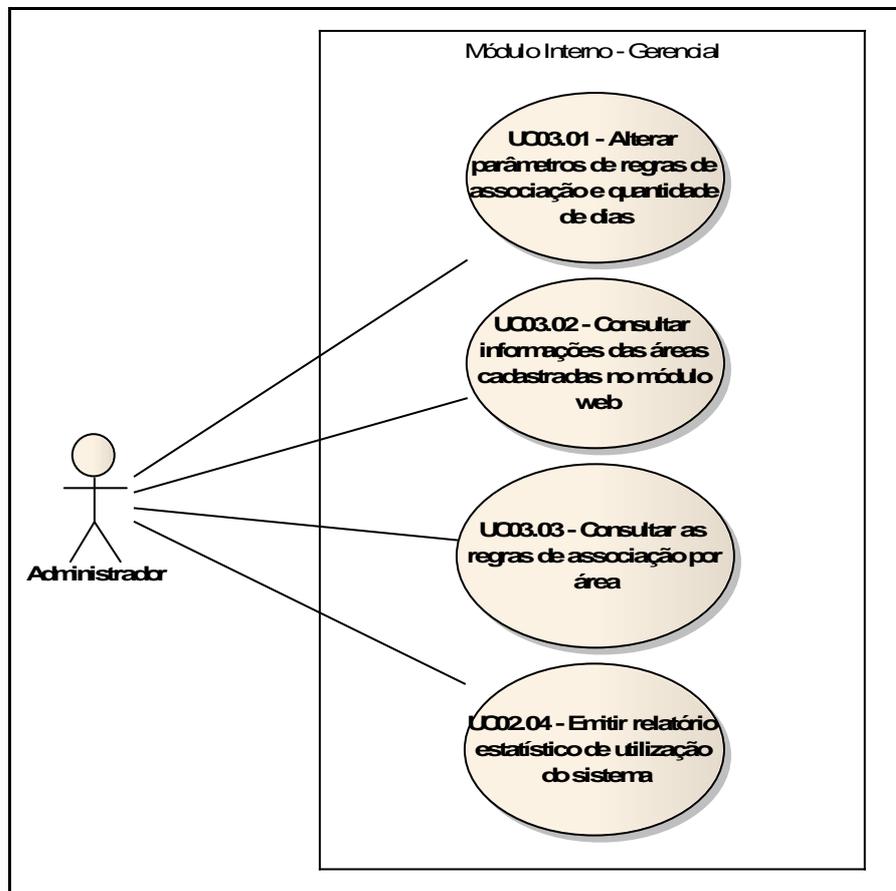


Figura 6 - Diagrama de caso de uso da ferramenta gerencial

3.2.2 Diagrama de atividades

O diagrama de atividades descreve o fluxo dos processos no sistema. Na figura 7, é apresentado o diagrama de atividades do módulo *web*, que representa o fluxo dos processos realizados pelo usuário em seu perfil através de um *website*.

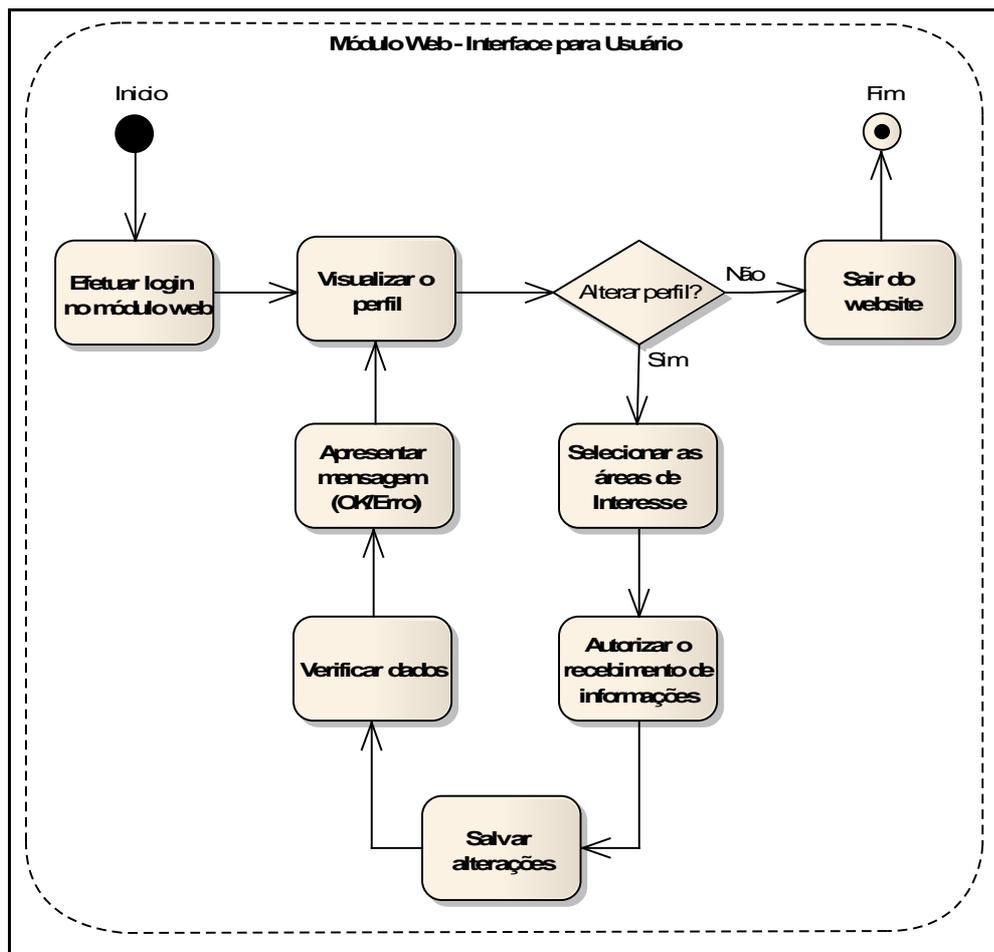


Figura 7 - Diagrama de atividades do *website* para o usuário configurar seu perfil

O usuário deverá efetuar o *login* no *website* e ao se autenticar é apresentada a página principal na tela, com todas as opções disponíveis de áreas. O usuário por sua vez configura seu perfil, fazendo a seleção das áreas de acordo com sua necessidade, depois habilita ou não a autorização de envio de informações para seu *e-mail* e salva suas preferências. O sistema irá verificar as preferências selecionadas, se a operação foi realizada corretamente o sistema emite uma mensagem de que a “operação foi realizada com sucesso”, caso contrário emite uma mensagem de “erro na operação”. Após a mensagem, o usuário é redirecionado

novamente para a página principal do seu perfil e pode optar por sair do *website*.

Já no módulo interno a ferramenta de envio tem o objetivo de disseminar as informações de novas aquisições para o *e-mail* dos usuários. Na Figura 8, é apresentado o diagrama de atividades da ferramenta de envio do módulo interno.

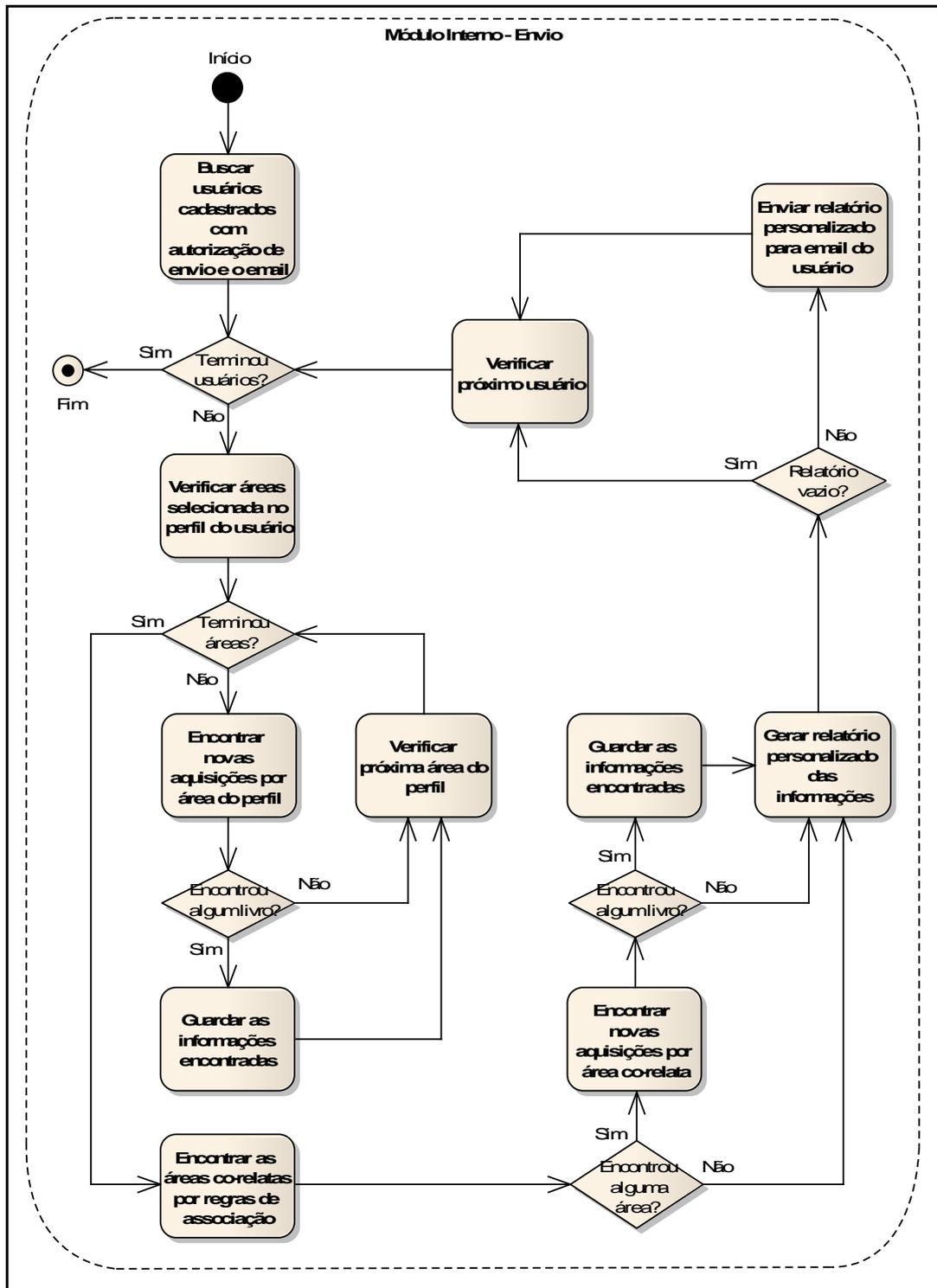


Figura 8 - Diagrama de atividades da ferramenta de disseminação das informações

A ferramenta de envio é automaticamente acionada em um determinado horário por um escalonador associado ao sistema operacional ou manualmente pelo administrador quando necessário. O fluxo dos processos inicia com a busca por usuários que tenham seus perfis cadastrados e a autorização para envio de informações habilitada. Para cada usuário, o sistema verifica as áreas selecionadas e faz uma busca por novas aquisições conforme cada área do perfil, armazenando as informações das obras encontradas. Finalizada essa primeira etapa, inicia-se o processo para gerar as sugestões, ou seja, encontrar áreas que tenham correlação com as áreas selecionadas no perfil, no qual para cada área correlata encontrada também é feita a busca por novas aquisições, armazenando-se as informações encontradas. Ao término das buscas, é gerado um relatório personalizado das informações que foram armazenadas. Caso este relatório não esteja vazio é então enviado por *e-mail* para o usuário, contendo informações das obras conforme as áreas selecionadas previamente em seu perfil e outras obras como sugestões apontadas na correlação. Os processos são repetidos enquanto houver algum usuário para enviar relatório das informações.

3.2.3 Modelagem entidade relacionamento

O sistema foi desenvolvido utilizando os padrões de algumas tabelas existentes atualmente no banco de dados da BU, e que são utilizadas em outros sistemas, como por exemplo, a consulta *online* para a lista de novas aquisições.

Para que fosse possível simular um ambiente parecido com o que a BU possui algumas tabelas e seus respectivos dados foram exportados. Isso permitiu o desenvolvimento do protótipo do sistema e projetando para que futuramente possa ser integrado aos serviços da BU.

As tabelas *acervo_dados_mfn* e *acervo_dicionario_mfn* fazem parte das tabelas importadas da estrutura atual, sendo que a tabela *acervo_dados_mfn* em específico contém informações das obras no padrão do formato MARC.

Assim como a tabela *area_cdd* (Anexo A), que mantém registros sobre a classificação da CDD, a tabela *dw_novas_aquisicoes* é utilizada para armazenar o número de registro de todos os exemplares das obras incorporadas ao acervo da BU, bem como a data de entrada. Já a tabela *circulação_historico_ano_base* mantém todo o histórico de empréstimos realizados em um período de um ano. Por fim, a tabela *pessoa* mantém informações dos alunos e

funcionários da FURB, e *pessoa_servicos_preferencias*, tabela utilizada para controle de preferências dos usuários ou de serviços *online* da BU.

Além das tabelas importadas da estrutura atual, foram criadas mais três tabelas novas, dentre elas, uma tabela para controle das áreas do perfil do usuário (*perfil_selecionado_web*), outra para as áreas que são disponibilizadas no módulo web (*area_web*) e a terceira para controle de parâmetros do sistema (*parametro_regra_assoc*). Foi acrescentado também mais um campo na tabela *pessoas_servicos_referencias* (*fl_autorizacao_novas_aquisicoes*) para o controle da autorização de envio. Na Figura 9, tem-se o diagrama de entidade relacionamento que representam as entidades que são utilizadas no banco de dados pelo sistema.

Para agilizar e permitir extrair as regras de associação, criou-se uma tabela do tipo visão materializada denominada *tab_padrao*, com base nos registros de históricos de empréstimos de livros, contendo alguns campos da tabela *circulacao_historico_ano_base*, e incluído ainda outros campos como, código da área, descrição da área e o código de MFN a qual o exemplar pertence. Esta visão materializada será atualizada pelo próprio gerenciador do banco de dados, neste caso o *Oracle*, e foi definido que as atualizações devem acontecer em um período semanal. Assim, as regras de associação geradas estarão sempre em conformidade com a demanda atual dos empréstimos realizados.

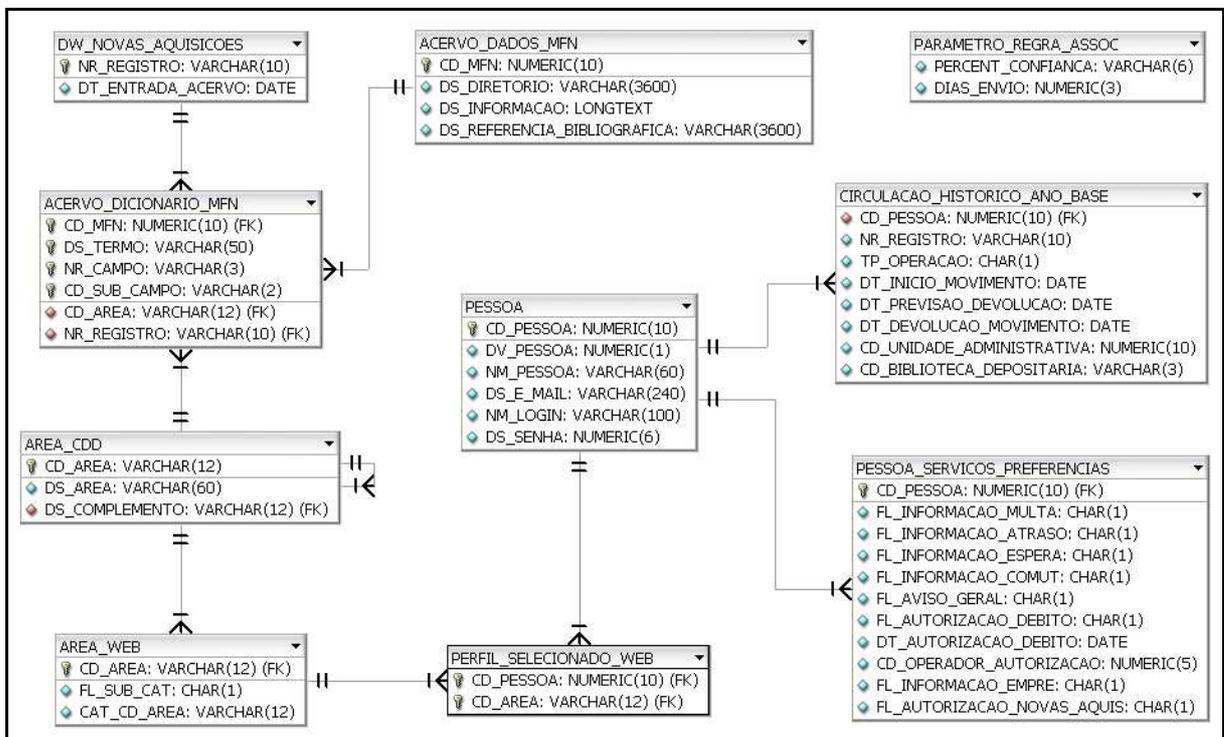


Figura 9 - Diagrama de entidade e relacionamento

A função de cada entidade do banco de dados está descrita a seguir após o nome da mesma:

- a) *acervo_dados_mfn*: contêm todas as informações cadastradas em relação à obra;
- b) *acervo_dicionario_mfn*: contêm termos indexados por palavras referente às obras;
- c) *area_cdd*: contêm informações como código e descrição de algumas áreas da CDD definidas pela biblioteca como principais ;
- d) *area_web*: contêm informações das áreas disponíveis para seleção apresentadas no perfil do usuário;
- e) *circulação_historico_ano_base*: contêm os registros das operações de empréstimos feitos pelos usuários no período de um ano até a data atual;
- f) *dw_novas_aquisicoes*: contêm os registros dos exemplares das obras que entram no acervo, bem como a data de entrada;
- g) *parametro_regra_assoc*: contêm informações de parâmetros para confiança das regras de associação e dias para envio;
- h) *perfil_selecionado_web*: contêm as áreas selecionadas pelo usuário;
- i) *pessoa*: entidade que contêm informações do usuário;
- j) *pessoa_servicos_preferencias*: contêm informações e *flags* referentes às preferências de serviços *online* do usuário.

3.3 IMPLEMENTAÇÃO

A seguir são mostradas as técnicas e ferramentas utilizadas e a operacionalidade da implementação do sistema.

3.3.1 Técnicas e ferramentas utilizadas

O projeto foi desenvolvido com ferramentas similares às utilizadas na BU/FURB, justamente para que futuramente possa ser integrado ao ambiente e sistemas da biblioteca. Quanto ao módulo *web*, foi adotada a ferramenta de desenvolvimento *Dreamweaver* da Adobe, e a linguagem de programação para internet PHP, sendo que para o visual da página

do perfil do usuário foram adotadas ainda, a linguagem de marcação HTML e também a linguagem de estilos *Cascading Style Sheets* (CSS). Fez-se necessário ainda, em alguns trechos de código, a utilização da linguagem *JavaScript* para tratamentos de alguns componentes da página.

No módulo interno, tanto para a ferramenta de disseminação das informações como para a ferramenta de gerenciamento do sistema, foi adotado o ambiente de programação *Delphi*.

Para o armazenamento dos dados, foi utilizado o gerenciador de banco de dados *Oracle 10g* versão *Express Edition*, disponibilizado gratuitamente no *website* da empresa Oracle. Houve a necessidade de criação duas funções no banco de dados, sendo utilizada a linguagem *Procedural Language/Structured Query Language* (PL/SQL), próprio do gerenciador de banco de dados. Uma destas funções, conforme o Quadro 10 é utilizada para ler os dados da tabela *acervo_dados_mfn*. Passando o número do campo do MARC a função irá retornar as informações correspondentes. Neste trabalho, é utilizada para retornar informação do campo 082, que corresponde à área de classificação da obra, e também o campo 245, que corresponde ao título e autor da obra. Já a outra função, é utilizada para encontrar a descrição da área a qual a obra pertence.

```

CREATE OR REPLACE FUNCTION OBRA(v_CD_MFN ACERVO_DADOS_MFN.CD_MFN%TYPE, v_NR_CAMPO NUMBER)
RETURN VARCHAR2
IS
v_DS_DIRETORIO VARCHAR2(3600);
v_DS_INFORMACAO LONG;
v_INICIO NUMBER :=1;
v_CAMPO VARCHAR2(3600) :='';
v_CAMPO_VER VARCHAR2(3600) :='';
BEGIN
SELECT DS_DIRETORIO, DS_INFORMACAO
INTO v_DS_DIRETORIO, v_DS_INFORMACAO
FROM ACERVO_DADOS_MFN
WHERE CD_MFN = v_CD_MFN;

WHILE (v_INICIO < LENGTH(v_DS_DIRETORIO)) LOOP
IF (SUBSTR(v_DS_DIRETORIO,v_INICIO,3)= v_NR_CAMPO) THEN
v_CAMPO:=SUBSTR(v_DS_INFORMACAO,SUBSTR(v_DS_DIRETORIO,v_INICIO+7,5)+2,SUBSTR(v_DS_DIRETORIO,v_INICIO+3,4)-1);
v_CAMPO_VER:=v_CAMPO;

IF (INSTR(v_CAMPO_VER,'^a')>0) THEN
v_CAMPO:=SUBSTR(v_CAMPO_VER,3,LENGTH(v_CAMPO_VER));
v_CAMPO_VER:=v_CAMPO;

IF (INSTR(v_CAMPO_VER,'^|0')>0) THEN
v_CAMPO:=SUBSTR(v_CAMPO_VER,1,LENGTH(v_CAMPO_VER)-4);
v_CAMPO_VER:=v_CAMPO;

ELSIF (INSTR(v_CAMPO_VER,'^|1')>0) THEN
v_CAMPO:=SUBSTR(v_CAMPO_VER,1,LENGTH(v_CAMPO_VER)-4);
v_CAMPO_VER:=v_CAMPO;

ELSE
v_CAMPO:=v_CAMPO_VER;
END IF;
END IF;
v_INICIO:= v_INICIO+12;
END LOOP;

RETURN v_CAMPO;
END OBRA;

```

Quadro 10 - Função para retornar informações da obra

Para gerar as sugestões que são enviadas aos usuários foi feita uma adaptação de Apriori para o caso específico da biblioteca. Apriori encontra primeiramente os conjuntos de itens frequentes da base de transações e posteriormente gera as regras a partir dos itens frequentes. Neste trabalho, era necessário encontrar as áreas correlatas a partir de cada área disponibilizada no módulo *web* para o usuário. Assim, as áreas do perfil são tratadas como itens frequentes, já que são consideradas mais relevantes, logo foram também consideradas antecedentes da regra. Então, depois de calculado o suporte da área que antecede a regra, são encontradas as áreas correlatas, ou seja, os consequentes da regra. Essas por sua vez, devem satisfazer um suporte de confiança mínimo definido previamente pelo administrador do sistema, e com base nas áreas correlatas resultantes é feita então a busca por novas aquisições que serão encaminhadas aos usuários como sugestão.

O modelo de regra de associação adotada para este trabalho diz-se do primeiro nível, sendo apenas uma área para o antecedente e uma área para o consequente da regra gerada, por exemplo, a área X do perfil possui a área Y como consequente.

O identificador único de transação (TID) utilizado para a extração das regras, a partir das transações de livros emprestados, é o código do usuário. E é importante ressaltar que o foco deste trabalho está na área que o livro pertence, conforme o número de classificação da CDD. Assim, independente da quantidade de livros ou títulos emprestados pelo usuário, cada área é contabilizada uma única vez por usuário. Um exemplo dessa abordagem está representado no Quadro 11.

TID (cód. usuário)	Áreas (cód. área)					
	005.1	342.1	630	658	370	813
001	1	0	0	1	0	1
002	0	1	0	0	1	0
003	0	1	1	0	0	1
004	1	1	0	1	0	1
005	0	1	1	1	0	1
Contador	2	4	2	3	1	4

Quadro 11 - Exemplo da base de dados de transações da abordagem adotada

Para encontrar essas áreas correlatas e gerar as regras de associação, utiliza-se como base de dados a visão materializada (*tab_padrao*) que contém os registros das transações de histórico de empréstimos realizados. A partir de uma determinada área selecionada no perfil do usuário é verificado na base quais são os usuários que fizeram empréstimos de livros daquela determinada área, e depois verifica quais são as outras áreas dos outros livros emprestados por esses mesmos usuários, por exemplo, usuários que emprestam livros da área

X, também emprestam livros da área Y, digamos que esta seja uma regra gerada $X \Rightarrow Y$, dessa maneira, um usuário que selecionou a área X em seu perfil, estaria recebendo informações da área X e também informações livros da área Y sob a forma de sugestões.

No Quadro 12 é apresentado o trecho de código que gera as regras de associação ou áreas correlatas em relação às áreas selecionadas pelo usuário em seu perfil.

```

qry_area_regra_assoc.Close:
qry_area_regra_assoc.SQL.Clear:
qry_area_regra_assoc.SQL.Add('SELECT RS.CD_AREA2');
qry_area_regra_assoc.SQL.Add('FROM (SELECT A.CD_AREA AS CD_AREA1, A.DS_AREA AS DS_AREA1, B.CD_AREA AS CD_AREA2, ');
qry_area_regra_assoc.SQL.Add('      B.DS_AREA AS DS_AREA2, COUNT(*) AS QTD_AREA2');
qry_area_regra_assoc.SQL.Add('      FROM (SELECT DISTINCT CD_AREA, DS_AREA, CD_PESSOA FROM TAB_PADRAO) B,');
qry_area_regra_assoc.SQL.Add('      (SELECT DISTINCT CD_AREA, DS_AREA, CD_PESSOA FROM TAB_PADRAO WHERE CD_AREA=' + cd_area + ') A');
qry_area_regra_assoc.SQL.Add('      WHERE A.CD_PESSOA = B.CD_PESSOA');
qry_area_regra_assoc.SQL.Add('      AND A.CD_AREA <> B.CD_AREA');
qry_area_regra_assoc.SQL.Add('      GROUP BY A.CD_AREA, A.DS_AREA, B.CD_AREA, B.DS_AREA');
qry_area_regra_assoc.SQL.Add('      ORDER BY QTD_AREA2 DESC, B.DS_AREA) RS,');
qry_area_regra_assoc.SQL.Add('      (SELECT COUNT(*) AS TOTAL_AREA1');
qry_area_regra_assoc.SQL.Add('      FROM (SELECT DISTINCT CD_AREA, DS_AREA, CD_PESSOA);
qry_area_regra_assoc.SQL.Add('      FROM TAB_PADRAO WHERE CD_AREA = ' + cd_area + ') QTD, PARAMETRO_REGRA_ASSOC PR');
qry_area_regra_assoc.SQL.Add('WHERE (RS.QTD_AREA2 / QTD.TOTAL_AREA1) * 100 >= PR.PERCENT_SUPORTE_CONF');
qry_area_regra_assoc.Open:

if not qry_area_regra_assoc.IsEmpty then
begin
    qry_area_regra_assoc.First:
    while not qry_area_regra_assoc.Eof do
    begin
        if cd_area_regra_assoc = '' then
        begin
            cd_area_regra_assoc := ('' + qry_area_regra_assoc.FieldByName('CD_AREA2').AsString + '');
        end
        else
        begin
            cd_area_regra_assoc := (cd_area_regra_assoc + ',' + '' + qry_area_regra_assoc.FieldByName('CD_AREA2').AsString + '');
        end;
        //end if cd_area_regra_assoc = ''
        qry_area_regra_assoc.Next:
    end;
    //end while qry_area_regra_assoc
end;
//end if qry_area_regra_assoc
qry_area_regra_assoc.Close:

```

Quadro 12 - Trecho de código para gerar as regras de associação

Depois de geradas as áreas correlatas, é feita uma busca por novas aquisições para cada área. As informações encontradas irão compor um relatório que posteriormente será encaminhado ao usuário através de *e-mail*. No Quadro 13 é apresentado o trecho de código que faz o envio de *e-mails*.

```

if encontrou = true then
begin

    IdMessage.Recipients.EmailAddresses := ds_e_mail;
    IdMessage.Subject := 'Novas Aquisições';
    IdMessage.ContentType := 'text/html';
    IdMessage.Body.Clear;
    IdMessage.Body.Add('<html><body>');
    IdMessage.Body.AddStrings(Memo1.Lines); //Novas Aquis areas do perfil
    IdMessage.Body.AddStrings(Memo2.Lines); //Novas Aquis sugestões
    IdMessage.Body.Add('</html></body>');

    IdSMTP.Connect;
    try
        IdSMTP.Send(IdMessage);
    finally
        IdSMTP.Disconnect;
    end; //end try
end; //end if encontrou

```

Quadro 13 - Trecho de código para envio de *e-mails*

3.3.2 Operacionalidade da implementação

Esta seção apresenta as funcionalidades do sistema desenvolvido, bem como telas ou interfaces para o usuário e ainda alguns trechos mais relevantes do código fonte do sistema.

Para um melhor entendimento, o funcionamento do sistema será descrito sequencialmente, iniciando com a configuração do perfil por parte do usuário, posteriormente a disseminação das informações pelo sistema e por fim o gerenciamento pelo administrador.

Na Figura 10, é demonstrada a interface *web* na qual o usuário irá autenticar-se com seu nome de usuário e senha para poder acessar seu perfil através de um *browser* de internet.



Figura 10 - Tela de *login* do usuário na interface do módulo *web*

Ao se autenticar é apresentado ao usuário o seu perfil, no qual ele irá selecionar as suas preferências. No topo da página é mostrado ao usuário seu *e-mail* para o qual serão enviadas informações. Caso ele queira alterar o *e-mail* cadastrado, o usuário deve clicar em “Alterar E-mail” e será redirecionado para uma página da BU/FURB onde poderá efetuar a alteração. Ainda no topo da página do perfil, existe o menu “Como Funciona” para ajuda ao usuário e o menu “Sair”. A Figura 11 apresenta a página principal do perfil do usuário.



Figura 11 - Tela da página principal para preferências do perfil do usuário

Após o usuário selecionar as áreas de sua preferência, mais ao final da página o usuário irá optar por autorizar ou não o recebimento de informações em seu *e-mail*. Por fim, irá clicar no botão “Salvar” para gravar as suas informações, conforme Figura 12.

The screenshot shows a web interface for selecting preferences. It features three columns of checkboxes for different areas: Serviço Social, Sociologia, Tecnologia de Alimentos, Turismo, Zoologia, Nutrição (with sub-options Culinária and Dietética), and Veterinária (with sub-options Animais Domésticos and Animais Selvagens). Below these is a section for email preferences with radio buttons for 'SIM' (selected) and 'NÃO'. A prominent blue 'Salvar' button with a checkmark is centered below. At the bottom, the FURB logo and the text 'Biblioteca Universitária da FURB Sistema para disseminação seletiva da informação' are visible.

Figura 12 - Tela da página principal para autorização e salvar preferências

Já a disseminação das informações é feita pela ferramenta de envio do módulo interno. A ferramenta será executada semanalmente de forma automatizada, sendo que, uma vez iniciada a execução do aplicativo, serão processadas todas as rotinas necessárias, desde a apuração das informações e geração dos relatórios personalizados para cada usuário, até o seu envio. Basicamente o aplicativo possui uma tela com alguns componentes visuais, como por exemplo, caixas de texto do tipo memorando, para o administrador somente acompanhar o que está sendo processado no momento da execução do aplicativo, conforme Figura 13. O administrador ainda poderá estar efetuando a execução manual do aplicativo, apenas abrindo a ferramenta.



Figura 13 - Tela do aplicativo para disseminação das informações

A ferramenta irá primeiramente efetuar a busca por novas aquisições, correspondente a cada área que o usuário selecionou em seu perfil desejando receber informações pertinentes.

Em seguida, são extraídas as áreas correlatas do perfil do usuário e posteriormente é feito uma nova busca, mas agora por novas aquisições conforme cada área correlata encontrada, e as informações retornadas da busca servem como sugestões que possam ser de interesse ao usuário.

Ao final de todas as rotinas executadas, se houver alguma informação de nova aquisição no relatório gerado, tanto informações conforme as áreas do perfil como sugestões, são então enviadas para o *e-mail* do usuário.

Já no aplicativo gerencial do módulo interno, Figura 14, o administrador poderá acessar as funcionalidades do aplicativo através do menu disposto da seguinte forma:

- a) menu parametrização: permite parametrizar percentual para regra de associação e quantidade de dias para busca de novas aquisições;
- b) menu consulta: permite consultar informações das áreas cadastradas para o perfil dos usuários e também consultar áreas correlatas;
- c) menu relatório: permite visualizar um relatório estatístico de utilização do sistema;
- d) sobre: descrição do nome do sistema, versão e autor;
- e) sair: encerra a execução do aplicativo.



Figura 14 - Tela principal do aplicativo gerencial do módulo interno

A partir do menu de parametrização do sistema, o administrador tem acesso à tela de parametrizações, Figura 15, na qual irá definir o percentual desejado para a confiança da regra que uma área tem em relação à outra e também a quantidade de dias retroativos em que será efetuada a busca por novas aquisições. Pretende-se que o envio das informações seja semanal, então por padrão adota-se sete dias, de forma configurável, possibilitando ao administrador da aplicação alterar caso necessário.

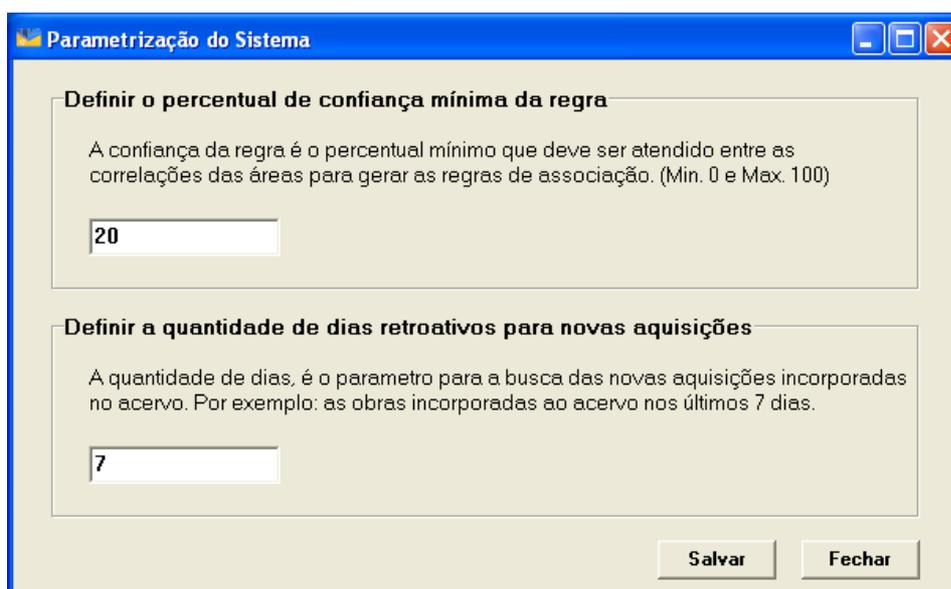


Figura 15 - Tela de parametrizações

Já no menu consulta, o administrador poderá consultar quais são as áreas que estão

cadastradas e que aparecerão no perfil do usuário, Figura 16. Ao abrir esta tela de consulta, todas as áreas aparecerão listadas em ordem alfabética pela descrição, e ainda pode ser feita uma pesquisa por uma determinada área.

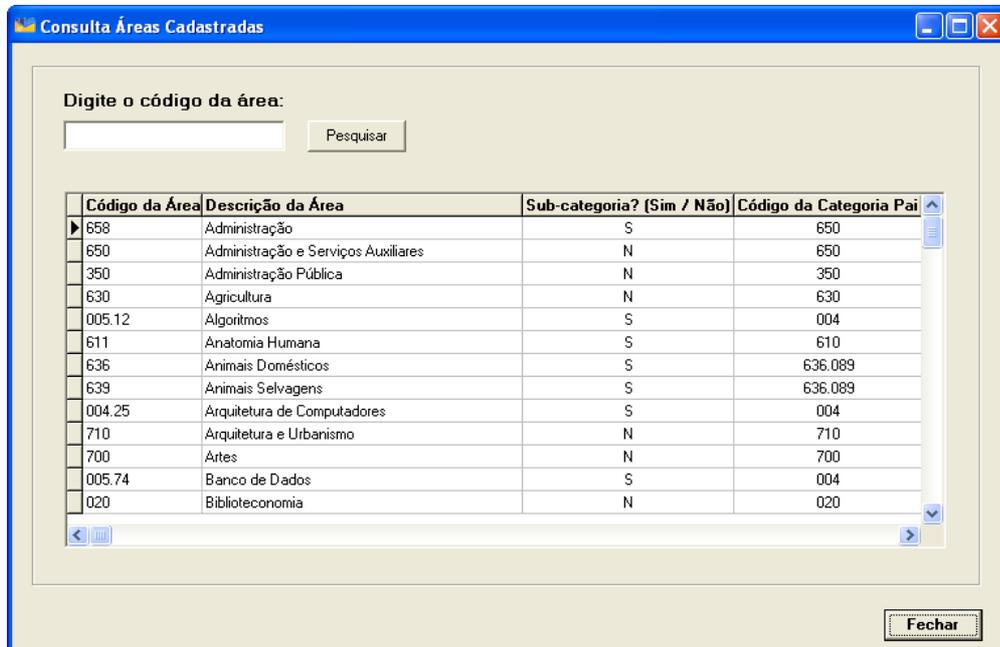


Figura 16 - Tela de consulta das áreas cadastradas para perfil do usuário

No mesmo menu ainda, existe a consulta de regras de associação, onde o administrador poderá pesquisar quais são as áreas correlatas de uma determinada área. Conforme Figura 17.



Figura 17 - Tela de consulta por áreas correlatas apontadas pela regra de associação

E por fim, no menu relatório, é apresentado um relatório estatístico de utilização do sistema, Figura 18, com informações como: total de perfis de usuário cadastrados no sistema, total de usuários que estão utilizando atualmente o sistema, ranking das áreas selecionadas nos perfis e total de áreas cadastradas.



Estatísticas de Utilização do Sistema

Usuários Cadastrados no Sistema: 2 **Data:** 06/05/2010

Usuários Utilizando o Sistema Atualmente: 2

Ranking das áreas selecionadas:

Código da área:	Descrição:	% Assinaturas:
658	Administração	100
657	Contabilidade	100
658.8	Marketing e Propaganda	100
650	Administração e Serviços Auxiliares	50
700	Artes	50
341.3	Direito Administrativo	50
342.1	Direito Civil	50
730	Escultura	50
770	Fotografia	50
780	Música	50
658.401	Planejamento Estratégico	50

Figura 18 - Tela de relatório de acompanhamento da utilização do sistema

3.4 RESULTADOS E DISCUSSÃO

O trabalho desenvolvido permitiu que informações filtradas sobre novas aquisições chegassem ao usuário através de seu *e-mail*, poupando esforços. Dessa maneira, ele receberá periodicamente informações sobre as obras mais recentes incorporadas ao acervo da biblioteca, ao invés de ter que buscar por essas informações de seu interesse no *website*.

A utilização da técnica de regras de associação para minerar as áreas correlatas, a partir do perfil do usuário, objetivou disseminar as informações sobre novas aquisições e ao encontrar a correlação das áreas permitiu enviar sugestões de obras ao usuário, agregando um

diferencial ao trabalho desenvolvido.

Embora ele não tivesse por objetivo analisar os motivos que geram as correlações entre as áreas, observou-se que os resultados obtidos através da técnica empregada foram satisfatórios e interessantes em relação às regras geradas das áreas. Entre os casos que podem ser mencionados para ilustrar o uso das regras de associação, por exemplo, com o percentual de confiança mínimo definido em 30%, e as áreas do perfil Artes e Literatura Inglesa, teríamos as seguintes regras geradas pelo sistema, conforme Quadro 14.

Artes (700) => História da Moda (391) = 42%
Artes (700) => Arte e sociedade (701) = 37%
Artes (700) => Desenho industrial (745.2) = 37%
Literatura Inglesa (823) => Literatura Americana (813) = 58%
Literatura Inglesa (823) => Literatura Brasileira (869.93) = 32%

Quadro 14 - Exemplo de regras geradas pelo sistema

Diante dessas regras, os usuários que selecionaram Artes e Literatura Inglesa em seu perfil, estariam recebendo como sugestões livros incorporados recentemente ao acervo e que foram classificados, conforme a CDD, em História da Moda, Arte e Sociedade, Desenho industrial, Literatura Americana e Literatura Brasileira.

Além da abordagem no uso da técnica de regras de associação, o que difere também dos demais trabalhos correlatos citados é o objetivo geral, que buscou focar somente na disseminação seletiva das informações sobre as obras recentemente incorporadas ao acervo, promovendo a divulgação destas aos usuários, ao mesmo tempo que fomenta o seu uso. Ressalta-se ainda, a disponibilização de um *website* onde o próprio usuário define seu perfil escolhendo suas áreas de preferência.

O trabalho desenvolvido por Jesus (2004) tinha como objetivo um sistema para recuperação e disseminação seletiva da informação, através de uma interface *web* personalizada de forma dinâmica para cada usuário. Já Silva (2004), desenvolveu uma ferramenta para identificação de padrões em base de dados de pacientes da Secretaria Municipal de Saúde de Londrina no Paraná, utilizando técnica de regras de associação. Por último Lucas e Souza (2007) propõem melhorar os processos da disseminação seletiva da informação em bibliotecas universitárias a partir do uso do CRM.

Por fim, com o protótipo do sistema desenvolvido para a BU/FURB, projetado com base na estrutura de banco de dados pré-existente, mostra que a proposta do trabalho pode ser integrada à estrutura atual sem alterações significativas e ainda agregar valor aos serviços oferecidos, de maneira mais atrativa ao usuário, principalmente no que se refere às novas aquisições.

4 CONCLUSÕES

Gerenciar a informação na Biblioteca Universitária (BU) da Universidade Regional de Blumenau (FURB), especificamente a disseminação da informação sobre novas aquisições do acervo bibliográfico para os usuários da BU mostrou-se importante para a criação de um canal de comunicação eficaz com os usuários, possibilitando a disseminação de informações sobre o acervo diretamente para o *e-mail* do usuário, minimizando tempo gasto com buscas e maximizando o uso dos recursos tecnológicos.

Foi proposto um sistema para a disseminação de informações das novas aquisições da BU/FURB e após o desenvolvimento do protótipo observa-se que os objetivos específicos destacados neste trabalho de conclusão de curso foram alcançados, tais como: o desenvolvimento de uma interface *web* na qual o usuário pode selecionar as áreas do conhecimento disponíveis para receber informações; a verificação das informações de novas aquisições disponíveis no acervo, separando-as por áreas predefinidas e gerando um relatório personalizado de acordo com o perfil de cada usuário. Pretende-se com a implantação deste protótipo que sejam disseminadas as informações referentes às novas aquisições bibliográficas de acordo com as áreas de preferência selecionadas previamente pelo usuário, juntamente com as informações detectadas através de regras de associação.

A utilização da tecnologia de mineração de dados para extração de informações em base de dados, especificamente da técnica de regra de associação, foi útil não só na ferramenta de envio, mas também na ferramenta gerencial, pois permitiu encontrar correlações entre algumas áreas que ainda eram desconhecidas e em outros casos possibilitou constatar associações que se acreditava existir.

Além disso, o emprego das regras de associação, permitiu criar um sistema mais elaborado que além de enviar as informações de interesse do usuário, conforme a seleção feita previamente por este, também envia informações de obras que podem interessar a estes usuários.

Entre as principais contribuições da utilização das regras de associação para o desenvolvimento do sistema está, a previsão de possíveis necessidades informacionais, detectadas pelo emprego das regras de associação ao verificar que quem empresta obras de uma área também empresta de outra determinada área. Assim, o diferencial do sistema desenvolvido é o envio das informações de acordo com o interesse do usuário por *e-mail*, associado ao envio de informações que também podem interessar ao usuário, detectadas pelo

uso das regras de associação.

Entretanto, foram encontradas algumas limitações durante o desenvolvimento do sistema, como a disponibilização para o usuário na interface *web* de uma quantidade limitada de áreas, áreas estas que já são predefinidas pela Biblioteca para a divulgação de novas aquisições. As áreas disponíveis atualmente foram definidas de acordo com as áreas mais gerais da CDD relacionadas aos cursos da Universidade, desta forma, não considera áreas mais específicas.

Por fim, o sistema proposto irá proporcionar que a BU/FURB disponibilize um serviço de disseminação da informação mais condizente com as necessidades dos usuários. A disponibilização de serviços de disseminação da informação em bibliotecas universitárias auxilia as atividades de ensino e pesquisa. O usuário ao se beneficiar do serviço o terá como um instrumento de atualização na sua área de atuação.

4.1 EXTENSÕES

É importante destacar que o sistema desenvolvido se restringiu a atender somente os requisitos propostos no trabalho, pois foi limitado pelo tempo disponível para sua elaboração, como um trabalho de conclusão de curso, e por depender da estrutura e dos dados dos sistemas já existentes na BU/FURB.

Por isso, sugere-se que sejam desenvolvidos trabalhos futuros ou o aperfeiçoamento e implementação de novas funcionalidades como:

- a) incluir novas áreas ao módulo *web* para que o usuário tenha mais opções;
- b) trabalhar com as áreas mais específicas ao buscar informações das novas aquisições referentes ao perfil selecionado;
- c) gerar as regras de associação para encontrar áreas correlatas das obras emprestadas por usuários de acordo com cursos de graduação, pós-graduação e departamentos;
- d) implementar o uso da técnica de regras de associação para encontrar outras correlações entre os empréstimos de obras realizados na BU/FURB;
- e) permitir que o usuário escolha se deseja receber as sugestões que são geradas pelas regras de associação.

REFERÊNCIAS

AGRAWAL, Rakesh et al. Fast discovery of association rules. In: FAYYAD, Usama M. et al. **Advances in knowledge discovery and data mining**. Menlo Park: AAAI: MIT, 1996. p. 307-328.

AMO, Sandra de. Técnicas de mineração de dados. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO. JORNADA DE ATUALIZAÇÃO EM INFORMATICA, 24. Salvador, 2004. **Anais...** Salvador: UFBA, 2004.

CARVALHO, Elizabeth Leão de. Importância da gestão da informação para o processo decisório nas organizações. In: VALENTIM, Marta Lígia Pomim (org.). **Informação, conhecimento e inteligência organizacional**. Marília: FUNDEPE ED, 2006. p. 81-98.

CARVALHO, Kátia de. Disseminação da informação e informação de inteligência organizacional. **DataGramZero**: Revista de Ciência da Informação, Rio de Janeiro, v.2, n.3, p 1-8, jun. 2001. Disponível em: <http://www.dgz.org.br/jun01/Art_04.htm>. Acesso em: 01 jun. 2010.

CORTES, Márcia Della Flora; LOPES, Marilisa Leite. As bibliotecas universitárias federais brasileiras e a acessibilidade das informações em seus websites. **Revista ACB: Biblioteconomia em Santa Catarina, Florianópolis**, v. 13, n.1, p.117-129, 2008. Disponível em: <<http://revista.acbsc.org.br/index.php/racb/article/viewArticle/552>>. Acesso em: 01 jun. 2010.

DAVENPORT, Thomas H; PRUSAK, Laurence. **Conhecimento empresarial: como as organizações gerenciam o seu capital intelectual**. Rio de Janeiro: Campus, 1998. 237p.

DIAS, Simone Lopes. **A disseminação da informação mediada por novas tecnologias e a educação do usuário na Biblioteca Universitária**. 2005. 138 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2005. Disponível em: <http://polo1.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/dias_sl_me_mar.pdf>. Acesso em 01 jun. 2010.

DEWEY, Melvil. **Dewey decimal classification and relative index**. 19 ed. Albany, N.Y: Forest Press, 1979. 3 v.

EIRÃO, Thiago Gomes. Disseminação seletiva da informação: uma abordagem. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v.7, n. 1, p. 20-29,

jul./dez. 2009. Disponível em: <<http://www.sbu.unicamp.br/seer/ojs/viewarticle.php?id=183>>. Acesso em: 12 mar. 2010.

ELMASRI, Ramez; NAVATHE, Shamkant B. **Sistemas de banco de dados**. 4. ed. São Paulo: Pearson Education do Brasil: Addison Wesley, 2005.

FAYYAD, Usama M.; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. From data mining to knowledge discovery: an overview. In: FAYYAD, Usama M. et al. **Advances in knowledge discovery and data mining**. Menlo Park: AAAI: MIT, 1996. p. 1-34.

FERAUD, Genevière. Um século de gestão da informação. In: DAVENPORT, Thomas H.; MARCHAND, Donald A.; DICKSON, TIM. **Dominando a gestão da informação**. Porto Alegre: Bookman, 2004. p. 39-44.

FUNDAÇÃO GETULIO VARGAS. Rede Bibliodata. **O que é formato MARC?** 2010. Disponível em: <<http://www8.fgv.br/bibliodata/geral/modelos/faqs.htm#2>>. Acesso em: 05 jun. 2010.

HAN, Jiawei; KAMBER, Micheline. **Data Mining: concepts and techniques**. 2nd ed. Amsterdam: Elsevier, 2006.

HAYKIN, Simon S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001.

JESUS, Alberto Pereira de. **Data Mining aplicado à identificação do perfil dos usuários de uma biblioteca para a personalização de sistemas Web de recuperação e disseminação de informações**. 2004. 120 f. Dissertação (Mestrado em Ciência da Computação) - Programa de Pós-Graduação em Ciência da Computação, Universidade Federal de Santa Catarina, Florianópolis, 2004. Disponível em: <http://www.bc.furb.br/docs/TE/2004/273389_1_1.pdf>. Acesso em: 10 abr. 2010.

LUCAS, Elaine R. de Oliveira; SOUZA, Nicole Amboni de. Disseminação seletiva da informação em Bibliotecas Universitárias sob o prisma do Customer Relationship Management. **Informação & Informação**, Londrina, v. 12, n. 1, p. 1- 17, 2007. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/viewArticle/1745>>. Acesso em: 01 jun. 2010.

MACHADO, Leticia dos Santos; BECKER, Karin. O uso da mineração de dados na web aplicado a um ambiente de ensino a distância. In: WORKSHOP DE TESES E DISSERTAÇÕES EM BANCO DE DADOS, 1., 2002, Gramado. **Anais eletrônicos...** Gramado: BDBComp, 2002. Disponível em: <<http://www.lbd.dcc.ufmg.br:8080/colecoes/wtdbd/2002/012.pdf>>. Acesso em: 22 mar. 2010.

MARANHÃO, Ana Maria Neves; MENDONÇA, Maria de Lourdes dos Santos. **Formato MARC 21**. Rio de Janeiro: PUC-Rio, 2010. Disponível em: <<http://www.dbd.puc-rio.br/MARC21/>>. Disponível em: 15 jun. 2010.

MICHAELIS: moderno dicionário da língua portuguesa. 4. ed. São Paulo : Melhoramentos, 1998.

NEVES, João Manuel Poças Marques das. **Ambiente de pós-processamento para regras de associação**. 2002. 118 f. Dissertação (Mestrado em Análise de Dados e Sistemas de Apoio à Decisão) - Faculdade de Economia, Universidade do Porto, Porto, 2002. Disponível em: <http://www.fep.up.pt/cursos/mestrados/madsad/teses/tese_mestrado_joao_pocas_pos_processamento_ra.pdf>. Acesso em: 30 maio 2010.

SILVA, Glauco Carlos. **Mineração de regras de associação aplicada a dados da Secretaria Municipal de Saúde de Londrina - PR**. 2004. 94 f. Dissertação (Mestrado em Ciência da Computação) - Programa de Pós-Graduação em Ciência da Computação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2004. Disponível em: <<http://www.lume.ufrgs.br/bitstream/handle/10183/8696/000586835.pdf?sequence=1>>. Acesso em: 01 jun. 2010.

SILVA, Terezinha Elisabeth da; TOMAÉL, Maria Inês. A gestão da informação nas organizações. **Informação e Informação**, Londrina, v. 12, n. 2, p. 1-2, jul./dez. 2007. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/viewFile/1806/1540>>. Acesso em: 01 jun. 2010.

SOUTO, Leonardo Fernandes; PORTELA, Patrícia de Oliveira. O SDI como instrumento de educação continuada: a responsabilidade das universidades no treinamento dos usuários. **Revista da ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v. 8/9, p. 123-133, 2003/2004. Disponível em: <<http://revista.acbsc.org.br/index.php/racb/article/view/411>>. Acesso em 05 mar. 2010.

STAREC, Claudio. A dinâmica da informação: a gestão estratégica da informação para a tomada de decisão nas organizações. In: STAREC, Claudio; GOMES, Elisabeth; BEZERRA, Jorge. **Gestão estratégica da informação e inteligência competitiva**. São Paulo: Saraíva, 2005. p. 47-64.

TAN, Pang-Ning; STEINBACH, Michael; KUMAR, Vipin. **Introdução ao data mining: mineração de dados**. Rio de Janeiro: Ciência Moderna, 2009.

TARAPANOFF, Kira; ARAÚJO JÚNIOR, Rogério Henrique de; CORMIER, Patricia Marie Jeanne. Sociedade da informação e inteligência em unidades de informação. **Ciência da Informação**, Brasília, v. 29, n.3, p. 91-100, set./dez. 2000. Disponível em: <http://www.scielo.br/scielo.php?pid=S0100-19652000000300009&script=sci_arttext&tlng=es>. Acesso em: 30 maio 2010.

UNIVERSIDADE REGIONAL DE BLUMENAU. Biblioteca Universitária. **Biblioteca online**. 2006a. Disponível em: <<http://www.bc.furb.br/consulta/novaConsulta/recuperaMfn.php?CdMFN=261406>>. Acesso em: 10 jun. 2010.

UNIVERSIDADE REGIONAL DE BLUMENAU. Biblioteca Universitária. 2006b. **Missão**. Disponível em: <<http://www.bc.furb.br/mambo/index.php?option=content&task=view&id=15&Itemid=33>>. Acesso em: 28 maio 2010.

UNIVERSIDADE REGIONAL DE BLUMENAU. Biblioteca Universitária. **Recursos de informática**. 2006c. Disponível em: <<http://www.bc.furb.br/mambo/index.php?option=content&task=view&id=133&Itemid=218>>. Acesso em: 28 maio 2010c.

VALENTIM, Marta Ligia Pomim. Inteligência Competitiva em Organizações: dado, informação e conhecimento. **DataGramZero**: Revista de Ciência da Informação, Rio de Janeiro, - v.3, n.4, p.1-13, ago. 2002. Disponível em: <http://www.dgz.org.br/ago02/Art_02.htm>. Acesso em: 16 maio 2010.

APÊNDICE A – Detalhamento dos casos de uso

No Quadro 15 apresenta-se o caso de uso "UC01.01 - Efetuar *Login*".

Nome do Caso de Uso	Efetuar <i>Login</i>
Descrição	Este caso de uso tem por objetivo permitir ao usuário acessar o módulo <i>web</i> via navegador de internet e informa seu <i>login</i> e senha para autenticação.
Ator	Usuário
Pré-condição	Sistema deve estar hospedado no servidor <i>web</i> ; Usuário deve estar cadastrado no banco de dados.
Fluxo principal	<ol style="list-style-type: none"> 1. Usuário preenche seu <i>login</i> e sua senha; 2. Sistema valida os dados de <i>login</i> e senha do usuário; 3. Sistema direciona o Usuário para a página Situação do Usuário.
Fluxo exceção	<ul style="list-style-type: none"> • nome de usuário e/ou senha inválido(s).
Pós-condição	Usuário conectado ao sistema.

Quadro 15 - Descrição do caso de uso Efetuar *Login*

No Quadro 16 apresenta-se o caso de uso "UC01.02 - Selecionar as áreas de preferência".

Nome do Caso de Uso	Selecionar as áreas de preferência
Descrição	Este caso de uso tem por objetivo apresentar ao usuário o seu perfil, na qual ele irá selecionar as suas áreas de preferências na qual deseja receber informações.
Ator	Usuário
Pré-condição	Usuário deve fazer <i>login</i> no sistema.
Fluxo principal	<ol style="list-style-type: none"> 1. Usuário acessa seu perfil; 2. Sistema apresenta na tela as áreas pré-cadastradas disponíveis; 3. Usuário seleciona as áreas de seu interesse; 4. Usuário autoriza o envio de informações para seu <i>e-mail</i>; 5. Usuário salva as informações.
Cenário – Visualização	Sistema mostra os registros das informações cadastradas pelo Usuário.
Fluxo exceção	<ul style="list-style-type: none"> • Nenhuma área selecionada pelo usuário então apresenta mensagem de erro no passo 5.
Pós-condição	Perfil do usuário cadastrado conforme suas preferências.

Quadro 16 - Descrição do caso de uso Selecionar as áreas de preferência

No Quadro 17 apresenta-se o caso de uso "UC01.03 - Alterar seleção das áreas de preferência".

Nome do Caso de Uso	Alterar seleção das áreas de preferência
Descrição	Este caso de uso tem por objetivo permitir ao usuário alterar a seleção das áreas de seu interesse na qual deseja receber informações, selecionando outras áreas ou removendo a seleção de áreas já selecionadas.
Ator	Usuário
Pré-condição	Usuário deve fazer <i>login</i> no sistema; Usuário já deve possuir o perfil cadastrado.
Fluxo principal	No passo 3 do fluxo principal do caso de uso "Selecionar as áreas de preferência", se o usuário desejar alterar a seleção de uma área ou selecionar outras áreas: 3.1 Usuário seleciona suas áreas marcando ou desmarcando a seleção; 3.2 Usuário salva as alterações realizadas; 3.3 Sistema apresenta mensagem de que as informações foram salvas.
Cenário – Visualização	Sistema apresenta o novo perfil gerado com as alterações feitas pelo usuário.
Fluxo exceção	<ul style="list-style-type: none"> Nenhuma área selecionada pelo usuário então apresenta mensagem de erro no passo 3.2.
Pós-condição	Perfil do usuário alterado conforme suas preferências.

Quadro 17 - Descrição do caso de uso Alterar seleção das áreas de preferência

No Quadro 18 apresenta-se o caso de uso "UC01.04 - Autorizar ou não o envio de informações por *e-mail*".

Nome do Caso de Uso	Autorizar ou não o envio de informações por <i>e-mail</i>
Descrição	Este caso de uso tem por objetivo permitir ao usuário autorizar ou não autorizar o envio de informações para seu <i>e-mail</i> .
Ator	Usuário
Pré-condição	Usuário deve fazer <i>login</i> no sistema; Usuário já deve possuir o perfil cadastrado.
Fluxo principal	No passo 4 do fluxo principal do caso de uso "Selecionar as áreas de preferência", se o usuário desejar alterar a autorização de envio de informações por <i>e-mail</i> : 4.1 Usuário escolhe e seleciona a autorização (sim ou não) de envio de informações para seu <i>e-mail</i> ; 4.2 Usuário salva a alteração realizada; 4.3 Sistema apresenta mensagem de que as informações foram salvas.
Cenário – Visualização	Sistema apresenta o novo perfil gerado com a alteração feita pelo usuário.
Pós-condição	Autorização alterada.

Quadro 18 - Descrição do caso de uso Autorizar ou não o envio de informações por *e-mail*

No Quadro 19 apresenta-se o caso de uso "UC02.01 - Enviar relatórios personalizados das informações apuradas".

Nome do Caso de Uso	Enviar relatórios personalizados das informações apuradas
Descrição	A ferramenta de envio é executada semanalmente, e este caso de uso tem por objetivo permitir o envio de um relatório personalizado para o <i>e-mail</i> do usuário, sobre as obras (livros) que foram incorporadas recentemente no acervo bibliográfico, conforme cada área do perfil, e também como sugestões que possam ser de interesse.
Ator	Agendador de Tarefas do S.O ou Administrador
Pré-condição	Haver pelo menos um perfil cadastrado; Existir obras incorporadas recentemente ao acervo.
Fluxo principal	<ol style="list-style-type: none"> 1. Agendador de Tarefas do S.O ou Administrador executam a aplicação; 2. Sistema faz a busca por perfis habilitados para receber informações; 3. Sistema consulta informações em cada perfil; <ol style="list-style-type: none"> 3.1 [Inclui caso de uso Consultar as áreas selecionadas no perfil do usuário]; 4. Sistema encontra as áreas correlatas; <ol style="list-style-type: none"> 4.1 [Inclui caso de uso Encontrar áreas correlatas através de regras de associação]; 5. Sistema gera relatório personalizado; <ol style="list-style-type: none"> 5.1 [Inclui caso de uso Gerar relatório das novas aquisições encontradas]; 6. Sistema verifica se relatório contém informações; 7. Sistema envia o relatório para <i>e-mail</i> do usuário; 8. Sistema finaliza a execução.
Fluxo exceção	<ul style="list-style-type: none"> • O relatório gerado está vazio, ou seja, não recuperou nenhuma informação sobre novas aquisições naquele período conforme as determinadas áreas do perfil, então não é enviado <i>e-mail</i> ao respectivo usuário.
Inclusão (includes)	UC02.02 - Consultar as áreas selecionadas no perfil do usuário; UC02.04 - Encontrar áreas correlatas através de regras de associação; UC02.05 - Gerar relatório das novas aquisições encontradas.
Pós-condição	Relatório enviado ao usuário.

Quadro 19 - Descrição do caso de uso Enviar relatórios personalizados das informações

No Quadro 20 apresenta-se o caso de uso "UC02.02 - Consultar as áreas selecionadas no perfil do usuário".

Nome do Caso de Uso	Consultar as áreas selecionadas no perfil do usuário
Descrição	Este caso de uso tem por objetivo permitir consultar as áreas selecionadas no perfil de cada usuário, para possibilitar uma busca por novas aquisições para cada área encontrada no perfil.
Ator	Agendador de Tarefas do S.O ou Administrador
Pré-condição	Haver pelo menos um perfil cadastrado.
Fluxo principal	1. Sistema consulta as áreas selecionadas em cada perfil encontrado; 2. Sistema guarda as áreas encontradas no perfil; 3. Sistema faz a busca por novas aquisições; 3.1 [Inclui caso de uso Buscar novas aquisições por área].
Fluxo exceção	• Nenhum perfil autorizado para envio de informações.
Inclusão (includes)	UC02.03 - Buscar novas aquisições por área.
Pós-condição	Áreas do perfil consultadas.

Quadro 20 - Descrição do caso de uso Consultar as áreas selecionadas no perfil do usuário

No Quadro 21 apresenta-se o caso de uso "UC02.03 - Buscar novas aquisições por área".

Nome do Caso de Uso	Buscar novas aquisições por área
Descrição	Este caso de uso tem por objetivo permitir a busca por novas aquisições por área, tanto as áreas selecionadas no perfil do usuário como as áreas correlatas. A busca sobre as novas aquisições é feita a partir da data atual de forma retroativa de dias parametrizados com base na data da entrada do exemplar no acervo.
Ator	Agendador de Tarefas do S.O ou Administrador
Pré-condição	Receber as áreas para busca por novas aquisições.
Fluxo principal	1. Sistema verifica a quantidade de dias retroativos que está parametrizado; 2. Sistema faz uma busca por novas aquisições para cada área; 3. Sistema guarda as informações encontradas sobre as novas aquisições.
Fluxo exceção	• Nenhuma informação encontrada, então não armazena nada.
Pós-condição	Busca realizada.

Quadro 21 - Descrição do caso de uso Buscar novas aquisições por área

No Quadro 22 apresenta-se o caso de uso "UC02.04 - Encontrar áreas correlatas através de regras de associação".

Nome do Caso de Uso	Encontrar áreas correlatas através de regras de associação
Descrição	Este caso de uso tem por objetivo permitir encontrar as áreas correlatas a partir de cada área do perfil do usuário, através de regras de associação, para que posteriormente seja feita a busca por novas aquisições de cada área correlata.
Ator	Agendador de Tarefas do S.O ou Administrador
Pré-condição	Haver pelo menos um perfil cadastrado.
Fluxo principal	<ol style="list-style-type: none"> 1. Sistema encontra as áreas correlatas para cada área selecionada no perfil do usuário utilizando de regras de associação; 2. Sistema guarda as áreas correlatas encontradas; 3. Sistema faz a busca por novas aquisições; <ol style="list-style-type: none"> 3.1 [Include caso de uso Buscar novas aquisições por área].
Fluxo exceção	<ul style="list-style-type: none"> • Nenhuma área correlata gerada pela regra de associação.
Inclusão (includes)	UC02.03 - Buscar novas aquisições por área.
Pós-condição	Áreas correlatas encontradas.

Quadro 22 - Descrição do caso de uso Encontrar áreas correlatas através de regras de associação

No Quadro 23 apresenta-se o caso de uso "UC02.05 - Gerar relatório das novas aquisições encontradas".

Nome do Caso de Uso	Gerar relatório das novas aquisições encontradas
Descrição	Este caso de uso tem por objetivo permitir organizar na forma de um relatório personalizado as informações encontradas sobre novas aquisições, obras referentes às áreas do perfil e as sugestões geradas pelo sistema.
Ator	Agendador de Tarefas do S.O ou Administrador
Pré-condição	Haver informações encontradas sobre novas aquisições.
Fluxo principal	<ol style="list-style-type: none"> 1. Sistema recupera as informações apuradas na busca feita de novas aquisições por área do perfil; 2. Sistema recupera as informações apuradas na busca feita de novas aquisições por áreas correlatas; 3. Sistema organiza as informações na forma de relatório.
Fluxo exceção	<ul style="list-style-type: none"> • Nenhuma informação encontrada, relatório fica vazio.
Pós-condição	Relatório gerado.

Quadro 23- Descrição do caso de uso Gerar relatório das novas aquisições encontradas

APÊNDICE B – Dicionário de dados das tabelas

O dicionário de dados descreve em mais detalhes as entidades da modelagem do banco de dados utilizada no trabalho desenvolvido. Os tipos de dados de cada campo são descritos a seguir:

- a) CHAR: tipo de campo para armazenamento de strings de caracteres de largura fixa, e os valores entre parênteses definem o comprimento máximo em bytes de caracteres;
- b) DATE: tipo de campo para armazenamento de datas;
- c) LONG: tipo de campo para armazenamento de strings de caracteres e seu tamanho é definido em bytes com largura variável;
- d) NUMERIC: tipo de campo para armazenamento de valores numéricos decimais, e os valores entre parênteses definem o comprimento máximo em dígitos;
- e) VARCHAR2: tipo de campo para armazenamento de strings de caracteres e seu tamanho é definido em bytes com largura variável, os valores entre parênteses definem o comprimento máximo em bytes de caracteres.

O dicionário de dados utiliza o seguinte padrão: nome da tabela, nome do campo, tipo do campo, descrição do campo e observação para o caso do campo fazer parte de chave primária e/ou chave estrangeira. No Quadro 24, pode ser visualizado o dicionário de dados do sistema.

Tabela: ACERVO_DADOS_MFN			
Nome da coluna	Tipo	Descrição	Observação
CD_MFN	NUMERIC (10)	Código do registro da obra	Chave Primária
DS_DIRETORIO	VARCHAR2 (3600)	String numérica que define a posição dos dados da obra conforme formato MARC	
DS_INFORMACAO	LONG	Dados da obra formato MARC	
DS_REFERENCIA_BIBLIOGRAFICA	VARCHAR2 (3600)	Referência da obra	
Tabela: ACERVO_DICIONARIO_MFN			
Nome da coluna	Tipo	Descrição	Observação
CD_MFN	NUMERIC (10)	Código do registro da obra	Chave Primária e Estrangeira
DS_TERMO	VARCHAR2 (50)	Descrição do termo da obra	Chave Primária
NR_CAMPO	VARCHAR2 (3)	Número do campo do formato MARC da informação	Chave Primária
CD_SUB_CAMPO	VARCHAR2 (2)	Código do subcampo do formato MARC	Chave Primária
CD AREA	VARCHAR2 (2)	Código de área CDD	Chave Estrangeira
NR_REGISTRO	VARCHAR2 (10)	Número de registro do exemplar da obra	Chave Estrangeira

Tabela: AREA_CDD			
Nome da coluna	Tipo	Descrição	Observação
CD_AREA	VARCHAR2 (12)	Código de área CDD	Chave Primária
DS_AREA	VARCHAR2 (60)	Descrição da área	
DS_COMPLEMENTO	VARCHAR2 (12)	Código complementar da área	Chave Estrangeira
Tabela: AREA_WEB			
Nome da coluna	Tipo	Descrição	Observação
CD_AREA	VARCHAR (12)	Código de área CDD disponibilizada no <i>website</i>	Chave Primária e Estrangeira
FL_SUB_CAT	CHAR (1)	Define se é subcategoria para <i>website</i>	
CAT_CD_AREA	VARCHAR (12)	Código de área CDD pai no <i>website</i>	
Tabela: CIRCULAÇÃO_HISTORICO_ANO_BASE			
Nome da coluna	Tipo	Descrição	Observação
CD_PESSOA	NUMERIC (10)	Código de pessoa	Chave Estrangeira
NR_REGISTRO	VARCHAR (10)	Número de registro do exemplar da obra	
TP_OPERACAO	CHAR (1)	Tipo de operação realizada	
DT_INICIO_MOVIMENTO	DATE	Data do empréstimo	
DT_PREVISAO_DEVOLUCAO	DATE	Data prevista para devolução	
DT_DEVOLUCAO_MOVIMENTO	DATE	Data da devolução do exemplar	
CD_UNIDADE_ADMINISTRATIVA	NUMERIC (10)	Código da unidade administrativa	
CD_BIBLIOTECA_DEPOSITARIA	VARCHAR (3)	Código da unidade setorial	
Tabela: DW_NOVAS_AQUISICOES			
Nome da coluna	Tipo	Descrição	Observação
NR_REGISTRO	VARCHAR2 (10)	Número de registro do exemplar da obra	Chave Primária
DT_ENTRADA_ACERVO	DATE	Data da entrada no acervo	
Tabela: PARAMETRO_REGRA_ASSOC			
Nome da coluna	Tipo	Descrição	Observação
PERCENT_CONFIANCA	VARCHAR (6)	Percentual de confiança mínimo	
DIAS_ENVIO	NUMERIC (3)	Número de dias retroativos para busca	
Tabela: PERFIL_SELECIONADO_WEB			
Nome da coluna	Tipo	Descrição	Observação
CD_AREA	VARCHAR2 (12)	Código de área CDD selecionada no perfil	Chave Primária e Estrangeira
CD_PESSOA	NUMERIC (10)	Código de pessoa	Chave Primária e Estrangeira
Tabela: PESSOA			
Nome da coluna	Tipo	Descrição	Observação
CD_PESSOA	NUMERIC (10)	Código de pessoa	Chave Primária
NM_PESSOA	VARCHAR2 (60)	Nome da pessoa	
DS_E_MAIL	VARCHAR2 (240)	<i>E-mail</i>	
NM_LOGIN	VARCHAR2 (100)	Nome de <i>login</i>	
DS_SENHA	NUMERIC (6)	Senha	

Tabela: PESSOA_SERVICOS_PREFERENCIAS			
Nome da coluna	Tipo	Descrição	Observação
CD_PESSOA	NUMERIC (10)	Código de pessoa	Chave Primária e Estrangeira
FL_INFORMACAO_MULTA	VARCHAR2 (1)	Controle de envio de informações de multa	
FL_INFORMACAO_ATRASO	VARCHAR2 (1)	Controle de envio de informações de atraso	
FL_INFORMACAO_ESPERA	VARCHAR2 (1)	Controle de envio de informações de espera	
FL_INFORMACAO_COMUT	VARCHAR2 (1)	Controle serviços de comutação	
FL_AVISO_GERAL	VARCHAR2 (1)	Controle de avisos gerais	
FL_AUTORIZACAO_DEBITO	VARCHAR2 (1)	Controle de autorização de débito de multa	
DT_AUTORIZACAO_DEBITO	DATE	Data de autorização do débito	
CD_OPERADOR_AUTORIZACAO	NUMERIC (5)	Código do operador	
FL_INFORMACAO_EMPRE	VARCHAR2 (1)	Controle de envio de informações de empréstimo	
FL_AUTORIZACAO_NOVAS_AQUIS	VARCHAR2 (1)	Controle da autorização para envio de informações sobre novas aquisições	

Quadro 24 - Dicionário de dados

ANEXO A - Áreas da CDD utilizadas pela BU para divulgação de informações de novas aquisições

Código	Descrição	Código grande área
000	Generalidades	000
001.4	Metodologia Científica	001.4
001.5	Generalidades	000
004	Informática	004
004.25	Arquitetura de Computadores	004
004.26	Informática	004
004.6	Redes	004
004.7	Informática	004
005.1	Programação	004
005.12	Algoritmos	004
005.13	Linguagem de Programação	004
005.43	Sistemas Operacionais	004
005.74	Banco de Dados	004
006.3	Inteligência Artificial	004
007	Generalidades	000
020	Biblioteconomia	020
028.5	Literatura Infante-Juvenil	028.5
028.6	Biblioteconomia	020
030	Dicionários e Enciclopédias	030
040	Generalidades	000
100	Filosofia	100
150	Psicologia	150
160	Filosofia	100
200	Religião	200
289.92	Testemunha de Jeová	200
300	Sociologia	300
302.23	Comunicação de Massa	658.8
302.24	Sociologia	300
304.2	Educação Ambiental	304.2
304.3	Sociologia	300
310	Estatística Demográfica	310
320	Ciência Política	320
330	Economia	330
333.7	Recursos Naturais	304.2
334	Economia - Cooperativas	330
340	Direito	340
341.2	Direito Constitucional	340
341.3	Direito Administrativo	340
341.39	Direito Tributário	340
341.43	Direito Processual Penal	340
341.46	Direito Processual Civil	340
341.5	Direito Penal	340
341.6	Direito do Trabalho	340
342.1	Direito Civil	340
342.14	Obrigações e Contratos	340

342.2	Direito comercial	340
350	Administração Pública	350
360	Serviço Social	360
363.7	Problemas e Serviços Ambientais	304.2
363.79	Poluição	360
370	Educação	370
378	Ensino Superior	370
379	Educação	370
380	Comércio, Comunicação e Transportes	380
380.145	Turismo	380.145
380.146	Comércio e Serviços	380
382	Comércio exterior	382
390	Costumes	390
391	Moda	391
392	Folclore	390
395	Etiqueta	390
400	Linguagem	400
500	Ciências Puras	500
510	Matemática	510
520	Astronomia	520
530	Física	530
540	Química	540
550	Geologia	550
560	Paleontologia	560
570	Biologia	570
574.192	Bioquímica	574.192
574.5	Ecologia e Meio Ambiente	304.2
574.6	Biologia econômica	570
580	Botânica	580
581.5	Ecologia das Plantas	304.2
581.6	Economia Botânica	580
590	Zoologia	590
591.5	Ecologia dos Animais	304.2
591.6	Zoologia Econômica	590
600	Tecnologia	600
610	Medicina	610
610.73	Enfermagem	610.73
610.74	Medicina	610
611	Anatomia Humana	610
612	Fisiologia Humana	610
612.3	Nutrição	612.3
612.4	Medicina	610
613.2	Dietética	612.3
613.4	Medicina	610
614	Saúde Pública	610
615	Farmácia	615
615.8	Fisioterapia	615.8
615.9	Medicina - Toxicologia	610
616	Medicina Interna	610
616.8	Neurologia	610
616.853	Epilepsia	610
616.89	Psiquiatria	610

617	Cirurgia	610
617.6	Odontologia	617.6
617.7	Medicina - Oftalmologia	610
618.1	Ginecologia	610
618.2	Obstetrícia	610
618.92	Pediatria	610
618.97	Geriatrics	610
619	Medicina Experimental	610
620	Engenharia	620
621.3	Engenharia Elétrica	621.3
621.382	Telecomunicações	621.382
621.4	Engenharia do Calor	620
624	Engenharia Civil	624
625	Engenharia Rodoviária	620
627	Engenharia Hidráulica	304.2
628	Engenharia Sanitária	304.2
629	Engenharia	620
630	Agricultura	630
634.9	Engenharia Florestal	634.9
635	Agricultura - Horticultura	630
636	Animais Domésticos	636.089
636.089	Veterinária	636.089
636.99	Agricultura - Horticultura	630
639	Animais Selvagens	636.089
640	Economia Doméstica	640
641.5	Culinária	612.3
641.878	Economia Doméstica	640
650	Administração e Serviços Auxiliares	650
651	Secretariado	650
657	Contabilidade	657
658	Administração	650
658.401	Planejamento Estratégico	658
658.8	Marketing e Propaganda	658.8
658.9	Administração	650
659	Propaganda	658.8
659.2	Relações Públicas	650
660	Engenharia Química	660
664	Tecnologia de Alimentos	664
665	Engenharia Química - Tecnologia Industrial	660
666	Cerâmica	660
667	Engenharia Têxtil	667
668	Engenharia Produtos Orgânicos	660
677	Engenharia Têxtil	667
678	Engenharia Química - Elastômeros	660
690	Construção Civil	690
700	Artes	700
710	Arquitetura e Urbanismo	710
730	Escultura	700
741	Design	741
745.929	Artes	700
760	Design	741
770	Fotografia	700

780	Música	700
790	Educação Física e Recreação	790
800	Literatura	800
813	Literatura Americana	800
823	Literatura Inglesa	800
869.93	Literatura Brasileira	800
900	História	900
910	Geografia	910
920	Biografia	920
930	História	900

Fonte: Dados do Sistema de Bibliotecas da Universidade Regional de Blumenau.

ANEXO B – Formato MARC

00X - Campos de Controle
02X - 09X - Campos de Números e Códigos
1XX - Entrada Principal – Campos
20X-24X - Título e Título Relacionado
25X-28X - Edição, Impressão, etc.
3XX - Descrição Física, etc.
4XX - Séries
5XX - Notas
6XX - Assuntos
70X-75X - Entrada Secundária
76X-78X - Entradas de Ligação
80X-840 - Entrada Secundária – Série
841-88X - Coleções, Localização, Gráficos alternados, etc.
9XX - Reservado para uso individual das bibliotecas e sistemas

Fonte: Maranhão e Mendonça (2010).

ANEXO C – Exemplo de registro bibliográfico no formato MARC

Dados do exemplo: OLIVEIRA, Adelize Generini de. **Delphi 7: conceitos básicos**. São Paulo: Relativa, 2003. 214 p, il.

```

Parágrafo 001 : UBU8000756151
Parágrafo 002 : n
Parágrafo 008 : 040316s2003 scba 000 0 por d
Parágrafo 020 : ^a8588503204 (broch.)
Parágrafo 082 : ^a005.369 DELPHI
Parágrafo 090 : ^a005.369 DELPHI^bO48dr^dCG
Parágrafo 100 : ^aOliveira, Adelize Generini de
Parágrafo 245 : ^aDelphi 7 : ^bconceitos básicos / ^cAdelize Generini de Oliveira. -
Parágrafo 260 : ^aSão Paulo : ^bRelativa, ^c2003.
Parágrafo 300 : ^a214 p. : ^bil.
Parágrafo 650 : ^aDelphi (Programa de computador)
Parágrafo 908 : am 22 a 4500
Parágrafo 949 : ^aCG224294^bBC^eC^tCG^y20040304
Parágrafo 949 : ^aCG224295^bBC^eC^tCG^y20040304
Parágrafo 999 : 261406

```

Fonte: Universidade Regional de Blumenau (2006a).