

UNIVERSIDADE REGIONAL DE BLUMENAU  
CURSO DE SISTEMAS DE INFORMAÇÃO



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

**MACIEL FELIPE DA SILVA**

PROF. CLÁUDIO RATKE, Orientador

# ROTEIRO DA APRESENTAÇÃO

- Introdução
- Objetivos
- Fundamentação teórica
- Desenvolvimento do sistema
- Especificações do sistema
- Implementação do sistema
- Resultados e discussões
- Conclusões
- Extensões



# INTRODUÇÃO

- A informação como um importante recurso para as organizações assume valor estratégico;
- O papel das bibliotecas dentro da universidade;
- A disseminação da informação, um importante veículo de comunicação entre a biblioteca e seus usuários (DIAS, 2005).
- Por que desenvolver um sistema para disseminação de informação de novas aquisições na Biblioteca Universitária da Universidade Regional de Blumenau?

## OBJETIVO GERAL

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.



# OBJETIVOS ESPECÍFICOS

- Disponibilizar uma interface *web* na qual o usuário possa selecionar as áreas do conhecimento disponíveis para receber informações;
- 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;
- Promover a disseminação das informações referentes às novas aquisições bibliográficas de acordo com as áreas de preferência selecionadas previamente pelo usuário;
- 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.

# GESTÃO E DISSEMINAÇÃO DA INFORMAÇÃO

- A presença da informação nos mais variados ambientes e atividades humanas (STAREC, 2005);
- No contexto acadêmico as bibliotecas têm a missão de oferecer aos seus usuários informações relevantes para a realização de pesquisas e ensino (LUCAS; SOUZA, 2007);
- O serviço de disseminação seletiva da informação.



# MINERAÇÃO DE DADOS

- Refere-se “à mineração ou descoberta de novas informações em função de padrões ou regras em grandes quantidades de dados.” (ELMASRI; NAVATHE, 2005, p. 624).
- “É uma parte integral da *Knowledge Discovery in Databases* (KDD), que é o processo geral de conversão de dados brutos em informações úteis.” (TAN; STEINBACH; KUMAR, 2009, p. 4);

# MINERAÇÃO DE DADOS

- Aplicações da mineração de dados:
  - Ex.: marketing, comércio, medicina, etc.
- Tarefas da mineração de dados:
  - Ex.: análise de associação.
- Técnicas da mineração de dados:
  - Ex.: regras de associação.



# REGRAS DE ASSOCIAÇÃO

- Objetiva encontrar correlações entre os itens de grandes conjuntos de dados relacionais ou transacionais (HAN; KAMBER, 2006);
- Exemplo: cesta de compras;
- Expressa por  $X \Rightarrow Y$  ( $X$  então  $Y$ ), onde  $X$  é o antecedente e  $Y$  o conseqüente da regra, sendo  $X$  e  $Y$  conjuntos distintos;
- Têm duas medidas utilizadas para análise:
  - o suporte;
  - a confiança.

# REGRAS DE ASSOCIAÇÃO

- Apriori é uma abordagem para minerar regras de associação, proposto por Agrawal e Srikant em 1994;
- Funcionamento do algoritmo Apriori.

# TRABALHOS CORRELATOS

- *Data Mining* aplicado à identificação do perfil dos usuários de uma biblioteca para a personalização de sistemas *Web* de recuperação e disseminação de informações (JESUS, 2004);
- Gestão do conhecimento: aplicação em *data mining* utilizando a teoria dos conjuntos aproximativos para geração do capital intelectual (SCHMITT, 2007);
- Disseminação seletiva da informação em Bibliotecas Universitárias sob o prisma do *Customer Relationship Managment* (LUCAS; SOUZA, 2007).

# LEVANTAMENTO DE INFORMAÇÕES

- Foco em novas aquisições da coleção geral, para usuários que utilizam o serviço de empréstimo;
- Sistemas da BU;
- Padrões internacionais adotados pela BU:
  - Master File Number (MFN);
  - Formato Machine Readable Cataloging (MARC);
  - Classificação Decimal de Dewey (CDD).



## SISTEMA DESENVOLVIDO

- Módulo web: página de Internet para que o usuário possa configurar seu perfil;
- Módulo interno: ferramenta de envio e ferramenta gerencial;
- Envio de e-mail com informações ao usuário conforme seu perfil;
- Dados enviados: título, autor e *link* para a consulta no *website* da BU.

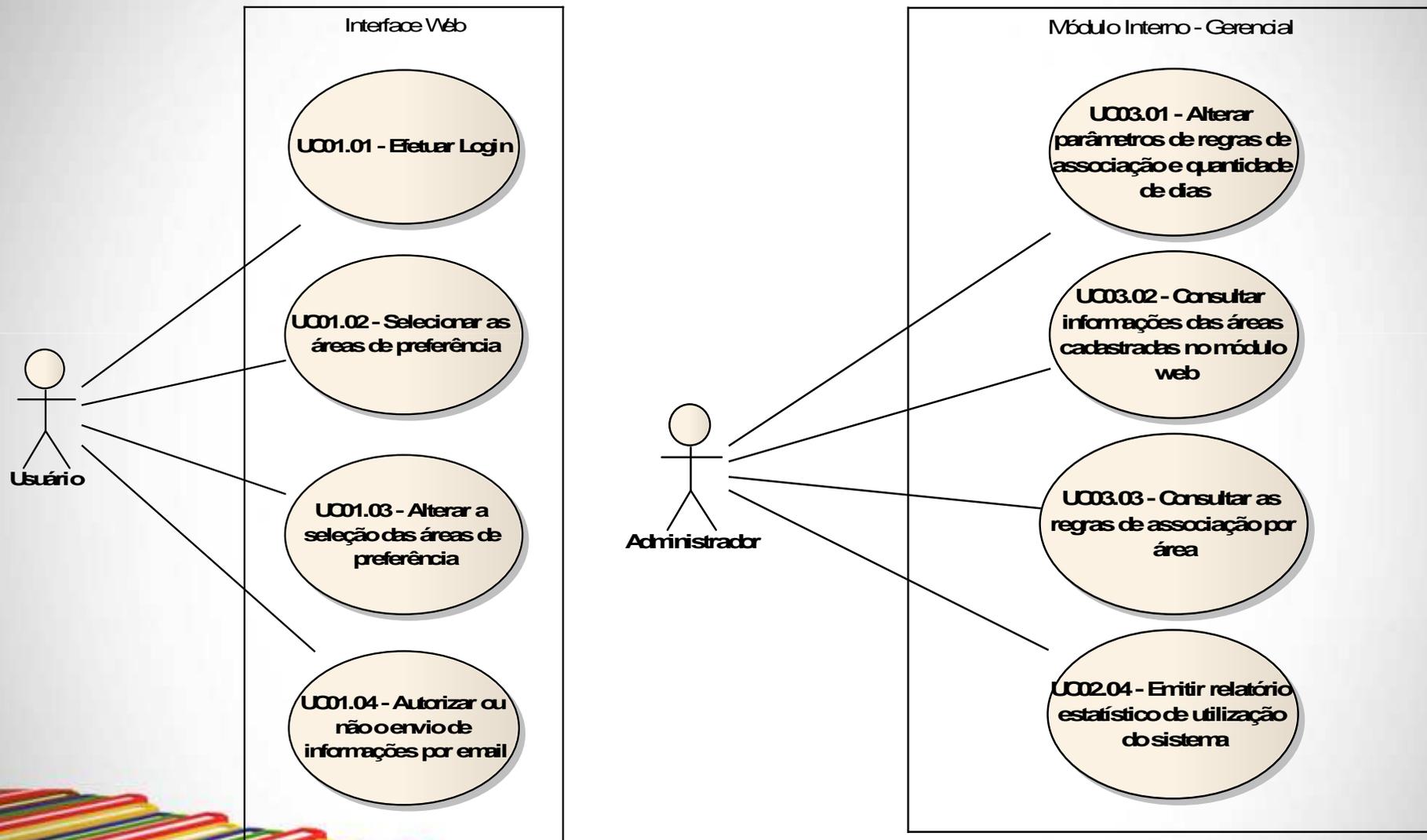
# ALGUNS REQUISITOS FUNCIONAIS

- RF02: O sistema deverá permitir ao usuário selecionar as áreas de preferência em seu perfil;
- RF04: O sistema deverá permitir ao usuário autorizar ou não o envio de informações para seu *e-mail*;
- RF05: O sistema deverá permitir enviar relatórios personalizados das informações apuradas;
- RF07: O sistema deverá permitir efetuar a busca por área de novas aquisições incorporadas ao acervo em um determinado período;
- RF08: O sistema deverá permitir gerar sugestões através de regras de associação, verificando quais são as áreas correlatas em relação ao perfil do usuário;
- 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;
- 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.

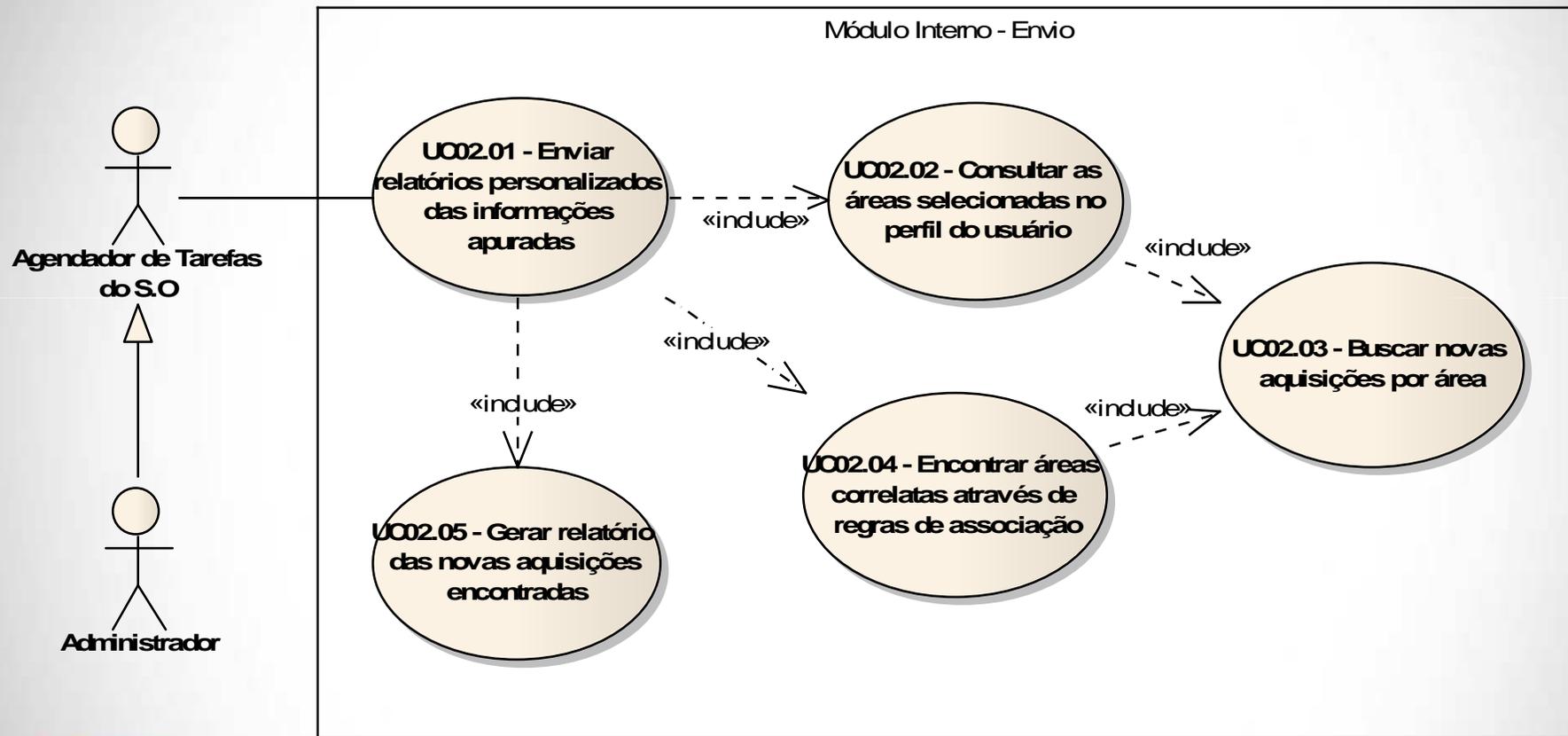
# REQUISITOS NÃO FUNCIONAIS

- RNF01: O módulo web do sistema deverá ser desenvolvido na linguagem PHP e HTML basicamente;
- RNF02: O módulo web do sistema deverá ter compatibilidade com os navegadores Internet Explorer 7.0 ou superior e Mozilla Firefox 3.5.x ou superior;
- RNF03: O módulo interno do sistema deverá ser desenvolvido na linguagem Delphi;
- RNF04: O sistema deverá suportar o banco de dados Oracle 10g;
- RNF05: O sistema deverá possuir interface web para acesso do usuário;
- RNF06: A interface web deverá ser hospedada em um servidor Apache;
- RNF07: O módulo interno deverá ter como sistema operacional base plataforma Windows;
- 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.

# DIAGRAMAS DE CASO DE USO



# DIAGRAMAS DE CASO DE USO



# BANCO DE DADOS

- Simulado estrutura atual da BU quanto ao banco de dados;
- Entidades e dados exportados;
- Novas entidades para o desenvolvimento do sistema;
- Criação de uma *view* materializada com base nas transações de empréstimos, para extrair as regras de associação.

# TÉCNICAS E FERRAMENTAS UTILIZADAS

- Módulo Web:
  - *Dreamweaver* da Adobe;
  - linguagem de programação para internet PHP;
  - linguagem de marcação HTML;
  - linguagem de estilos *Cascading Style Sheets* (CSS);
  - linguagem *JavaScript*.
- Módulo interno:
  - ambiente de programação *Delphi*.

# TÉCNICAS E FERRAMENTAS UTILIZADAS

- Armazenamento dos dados:
  - gerenciador de banco de dados *Oracle 10g* versão *Express Edition*;
  - funções em linguagem *Procedural Language/Structured Query Language* (PL/SQL);
- Gerar as sugestões que são enviadas aos usuários:
  - adaptação de Apriori para o caso específico da biblioteca.

# PÁGINA PRINCIPAL DO PERFIL

Novas Aquisições Biblioteca Universitária - FURB - Windows Internet Explorer

http://localhost/sugestor/home.php?p=options

Arquivo Editar Exibir Favoritos Ferramentas Ajuda

★ Favoritos Novas Aquisições Biblioteca Universitária - FURB

**Bem-Vindo! Maciel Felipe da Silva**  
As informações serão encaminhadas para: maciel.f@gmail.com

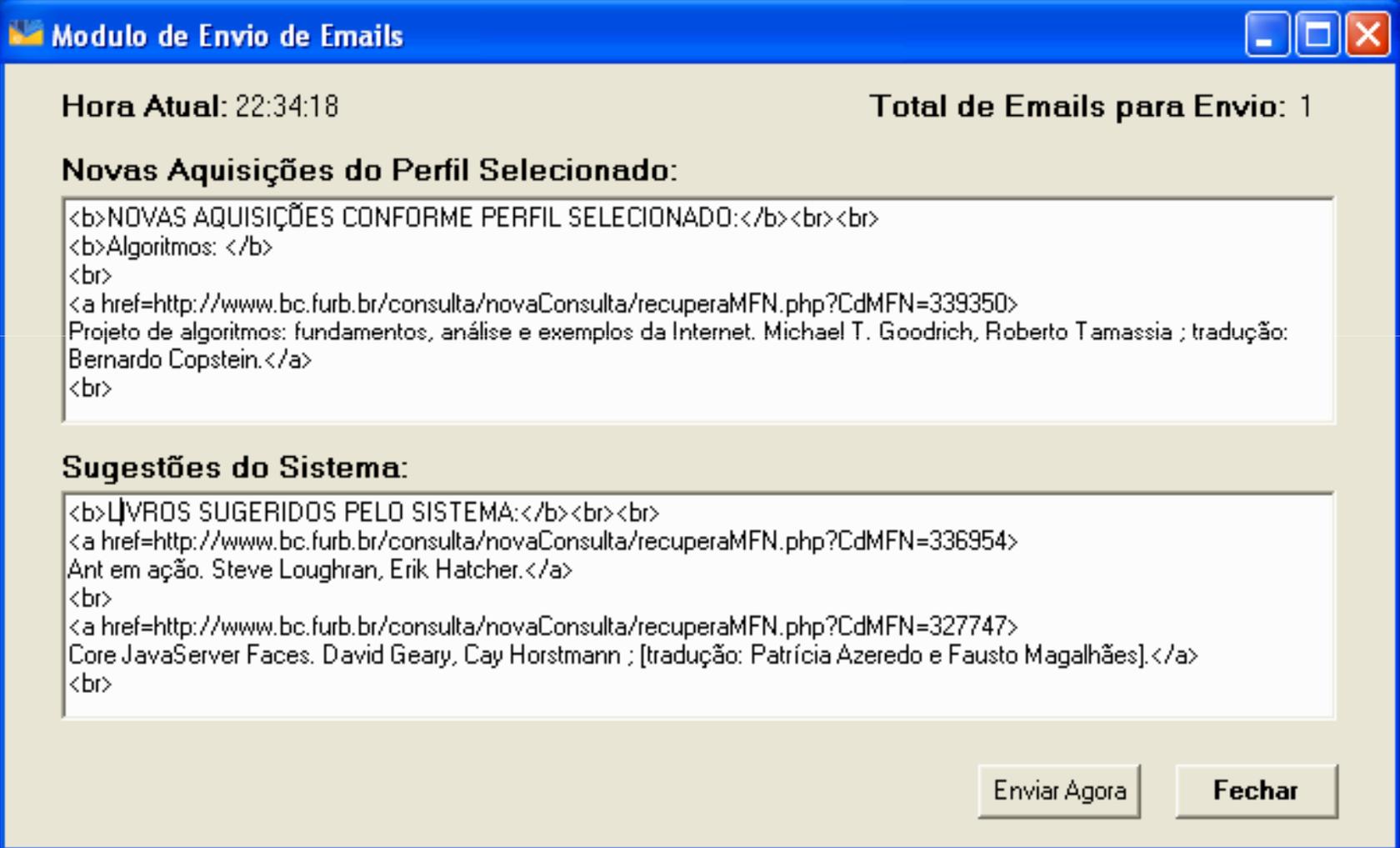
 **NOVAS AQUISIÇÕES**  
via email

Alterar Email Como Funciona? Sair

**Selecione abaixo as áreas do seu interesse:**

<input type="checkbox"/> Administração Pública	<input type="checkbox"/> Administração e Serviços Auxiliares	<input type="checkbox"/> Artes
<input type="checkbox"/> Agricultura	<input type="checkbox"/> Administração	<input type="checkbox"/> Escultura
<input type="checkbox"/> Arquitetura e Urbanismo	<input type="checkbox"/> Contabilidade	<input type="checkbox"/> Fotografia
<input type="checkbox"/> Biblioteconomia	<input type="checkbox"/> Marketing e Propaganda	<input type="checkbox"/> Música
<input type="checkbox"/> Biografia	<input type="checkbox"/> Planejamento Estratégico	
<input type="checkbox"/> Biologia	<input type="checkbox"/> Secretariado	
<input type="checkbox"/> Bioquímica	<input type="checkbox"/> Direito	<input type="checkbox"/> Educação Ambiental
<input type="checkbox"/> Botânica	<input type="checkbox"/> Direito Administrativo	<input type="checkbox"/> Ecologia das Plantas
<input type="checkbox"/> Ciência Política	<input type="checkbox"/> Direito Civil	<input type="checkbox"/> Ecologia dos Animais
<input type="checkbox"/> Comércio exterior	<input type="checkbox"/> Direito comercial	<input type="checkbox"/> Ecologia e Meio Ambiente
<input type="checkbox"/> Design	<input type="checkbox"/> Direito Constitucional	<input type="checkbox"/> Problemas e Serviços Ambientais
	<input type="checkbox"/> Direito do Trabalho	<input type="checkbox"/> Recursos Naturais

# TELA DA FERRAMENTA DE ENVIO



**Modulo de Envio de Emails**

Hora Atual: 22:34:18 Total de Emails para Envio: 1

**Novas Aquisições do Perfil Selecionado:**

**NOVAS AQUISIÇÕES CONFORME PERFIL SELECIONADO:**  
**Algoritmos:**  
<http://www.bc.furb.br/consulta/novaConsulta/recuperaMFN.php?CdMFN=339350>  
Projeto de algoritmos: fundamentos, análise e exemplos da Internet. Michael T. Goodrich, Roberto Tamassia ; tradução: Bernardo Copstein.

**Sugestões do Sistema:**

**LIVROS SUGERIDOS PELO SISTEMA:**  
<http://www.bc.furb.br/consulta/novaConsulta/recuperaMFN.php?CdMFN=336954>  
Ant em ação. Steve Loughran, Erik Hatcher.  
<http://www.bc.furb.br/consulta/novaConsulta/recuperaMFN.php?CdMFN=327747>  
Core JavaServer Faces. David Geary, Cay Horstmann ; [tradução: Patrícia Azeredo e Fausto Magalhães].

# TELA DA FERRAMENTA GERENCIAL



# TELA DE PARAMETRIZAÇÕES

**Parametrização do Sistema**

**Definir o percentual de confiança mínima da regra**

A confiança da regra é o percentual mínimo que deve ser atendido entre as correlações das áreas para gerar as regras de associação. (Min. 0 e Max. 100)

**Definir a quantidade de dias retroativos para novas aquisições**

A quantidade de dias, é o parametro para a busca das novas aquisições incorporadas no acervo. Por exemplo: as obras incorporadas ao acervo nos últimos 7 dias.

**Salvar** **Fechar**

# TELA DE CONSULTA ÁREAS CADASTRADAS

Consulta Áreas Cadastradas

Digite o código da área:

Código da Área	Descrição da Área	Sub-categoria? (Sim / Não)	Código da Categoria Pai
▶ 658	Administração	S	650
650	Administração e Serviços Auxiliares	N	650
350	Administração Pública	N	350
630	Agricultura	N	630
005.12	Algoritmos	S	004
611	Anatomia Humana	S	610
636	Animais Domésticos	S	636.089
639	Animais Selvagens	S	636.089
004.25	Arquitetura de Computadores	S	004
710	Arquitetura e Urbanismo	N	710
700	Artes	N	700
005.74	Banco de Dados	S	004
020	Biblioteconomia	N	020

# TELA PARA CONSULTA DE REGRAS

Consulta Regras de Associação

Digite o código da área:

341.2

Confiança da regra definida em: 30%

Código	Descrição	Cód. Área 2	Descrição Área 2	Confiança (%)
▶ 341.2	Direito Constitucional	341.46	Direito Processual Civil	49
341.2	Direito Constitucional	341.5	Direito Penal	48
341.2	Direito Constitucional	342.1	Direito Civil	39
341.2	Direito Constitucional	341.6	Direito do Trabalho	35
341.2	Direito Constitucional	342.14	Obrigações e Contratos	35
341.2	Direito Constitucional	341.3	Direito Administrativo	33

# RESULTADOS E DISCUSSÕES

- Informações filtradas ao usuário através de seu *e-mail* poupando esforços;
- Regras de associação para gerar sugestões agregando um diferencial ao trabalho;
- Permitir agregar valor aos serviços oferecidos pela BU;
- Resultados interessantes nas regras geradas em relação às áreas.

## EXEMPLO DE REGRAS GERADAS

Definindo a confiança mínima da regra em 30%, resultaram as seguintes regras, para as áreas de Artes e Literatura Inglesa:

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%

# CONCLUSÕES

- Gerenciar a informação na BU/FURB mostrou-se importante para a criação de um canal de comunicação eficaz com os usuários;
- Os objetivos propostos foram atingidos;
- A utilização da tecnologia de mineração de dados foi útil tanto na ferramenta de envio como na ferramenta gerencial;
- O emprego das regras de associação permitiu criar um sistema mais elaborado;
- O sistema proposto irá proporcionar que a BU/FURB disponibilize um serviço de disseminação da informação mais condizente com as necessidades dos usuários.

# EXTENSÕES

- Incluir novas áreas ao módulo *web* para que o usuário tenham mais opções;
- Trabalhar com as áreas mais específicas ao buscar informações das novas aquisições referente ao perfil selecionado;
- Gerar as regras de associação para encontrar áreas correlatas das obras emprestadas por usuários de acordo com cursos de graduação, pós-graduação e departamentos;
- Implementar o uso da técnica de regras de associação para encontrar outras correlações entre os empréstimos de obras realizados na BU/FURB.



# REFERÊNCIAS

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](http://polo1.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/dias_sl_me_mar.pdf)>. Acesso em 01 jun. 2010.

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

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

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](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.

SCHMITT, Sidnei. **Gestão do conhecimento**: aplicação em data mining utilizando a teoria dos conjuntos aproximativos para geração do capital intelectual. 2005. 61 f. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação) - Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau, 2005. Disponível em: <[http://www.bc.furb.br/docs/MO/2008/329590\\_1\\_1.pdf](http://www.bc.furb.br/docs/MO/2008/329590_1_1.pdf)>. Acesso em: 2 jun. 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.



**OBRIGADO!**

