

# Uma Ferramenta Web para BI focada no Gestor de Informação

Mikael de Souza Fernandes<sup>1</sup>, Gustavo Zanini Kantorski<sup>12</sup>  
**mikael@cpd.ufsm.br, gustavoz@cpd.ufsm.br**

<sup>1</sup> Curso de Sistemas de Informação, Universidade Luterana do Brasil (ULBRA),  
Campus Santa Maria, RS, Brasil.

<sup>2</sup> Centro de Processamento de Dados, Universidade Federal de Santa Maria (UFSM),  
Santa Maria, RS, Brasil.

## 1. Introdução

O conceito de BI (*Business Intelligence*), de forma mais ampla, pode ser entendido como a utilização de variadas fontes de informação para se definir estratégias de competitividade nos negócios da empresa, afirma Barbieri (2001).

Nos dias atuais, a gestão das instituições está cada vez mais relacionada à sua capacidade de transformar dado em informação, informação em conhecimento e este último em decisões de negócio, o conceito de *Business Intelligence* (BI) exerce um papel fundamental como ferramenta de apoio aos processos de tomada de decisão.

Embora as ferramentas *open source* utilizem recursos sofisticados para Web, estas interfaces não são tão amigáveis para utilização pelo usuário final. A configuração das interfaces e a montagem de consultas necessitam o conhecimento de uma linguagem técnica que é utilizada para manipulação de banco de dados dimensionais. Além disso, as ferramentas não permitem a persistência das consultas após a iteração do usuário.

Neste sentido, foi desenvolvida uma ferramenta *Web open source*, integrada ao servidor OLAP *Mondrian* e a biblioteca *JPivot*, que permite a persistência e criação de consultas sem a necessidade de conhecimento da linguagem técnica para consulta. A ferramenta *Neuro* foi desenvolvida para auxiliar a tomada de decisão através da análise de dados, onde o usuário pode criar as suas consultas e disponibilizá-las à comunidade de interesse.

Este trabalho está organizado da seguinte forma: a seção dois apresenta os conceitos de BI onde são abordados os temas sobre *Data Warehouse* (DW) e *On-Line Analytical Processing* (OLAP). A terceira seção mostra algumas de tecnologias de BI, proprietárias e *open source*, existentes no mercado. Na seção quatro é apresentada a ferramenta desenvolvida, o processo de modelagem e a implementação. E, por fim, são apresentadas as considerações finais e trabalhos futuros.

## 2. *Business Intelligence*

As ferramentas para um ambiente de BI podem ser classificadas como construção, gerência, uso e armazenamento (Barbieri, 2001), sendo, obviamente, o objeto em foco a informação. Tais classificações são concebidas em conceitos como DW e OLAP.

Segundo Inmon (1997), DW é uma coleção de dados orientados ao assunto, integrados, não-voláteis e variantes no tempo, para fornecer suporte às decisões de gerentes. O DW fornece acesso a dados para análise complexa, descoberta de conhecimento e tomada de decisão. Suportam demandas de alto nível de desempenho de

dados e informações em uma organização, vários tipos de aplicações – OLAP, DSS (*Decision Support Systems*) e *Data Mining* (Elmasri & Navathe, 2000).

Para manipular um DW necessita-se de conceitos e tecnologias distintas. Comumente em base de dados tradicionais para operações convencionais, transacionais, utilizam-se tecnologias chamadas de OLTP (*On-Line Transaction Processing*). Entretanto, no caso de DW, utiliza-se de tecnologias denominadas OLAP.

A Tabela 1 adaptada (Thomsen, 2002) que demonstra uma comparação, quanto às atividades desempenhadas, entre os conceitos de OLTP e OLAP, em que é possível visualizar as características opostas.

### 3. Tecnologias BI

#### 3.1 Pentaho

A plataforma Pentaho é um projeto *open source* para BI, compreendendo vários produtos e subprodutos. O projeto também denominado de *Open BI Suite* fornece relatórios abrangentes, análise OLAP, *dashboards*, integração de dados, mineração de dados (Pentaho, 2010).

Inserido no projeto Pentaho *Open BI Suite* está a ferramenta OLAP Mondrian. É uma aplicação Web de análise que permite aplicar as principais operações OLAP em um DW. Ainda possibilita a geração de gráficos, conversão de consultas para visualização nos formatos PDF e Microsoft Excel.

#### 3.2 Microsoft

A Microsoft possui uma solução de BI junto ao banco de dados Microsoft SQL Server, denominada *Analysis Services*. O Microsoft SQL Server *Analysis Services* (SSA) é um servidor baseado em OLAP e mineração de dados (Microsoft, 2008), assim como o banco de dados SQL Server possui custo de licença. Esta solução oferece um pacote completo de recursos para aplicação dos conceitos de Inteligência de Negócios, como ainda ETL e DW.

#### 3.3 IBM

A IBM (*International Business Machines*) possui sua solução de software BI proprietário, isto é, possui custos de licença, denominadas IBM Cognos *Business Intelligence* (atualmente na versão 8). Cognos permite conectar a qualquer fonte de dados e aplicar técnicas de BI, como, por exemplo, emitir relatórios. A solução Cognos possibilita os seguintes itens (Cognos, 2010): relatórios (*Reporting*), análise (*Analysis*), *dashboards* e *scorecarding*.

**Tabela 2 – Tabela comparativa das características das tecnologias de análise**

<b>Tecnologias/ Características de Análise</b>	<b>Pentaho</b>	<b>Microsoft</b>	<b>IBM</b>
<i>Open-source</i>	Sim	Não	Não
Nível de conhecimento dos usuários	Técnico	Técnico	Técnico
Controle de acesso a consultas	Não	Não	Não
Persistência de consultas	Não	Não	Não

Diante da análise das tecnologias BI constatou-se algumas dificuldades (Tabela 1) para o usuário OLAP que, atualmente, deve ser especializado, ou deve possuir

conhecimento fundamentado para manipular essas tecnologias. Ainda constatou-se que dentre as tecnologias estudadas, todas de relevância no mercado de BI, Pentaho é *open-source*, isto é, possui código-fonte aberto e permite manipulação deste. No projeto Pentaho *Open BI Suíte*, na solução de análise (*analysis*) percebeu-se a mesma dificuldade inicial.

O projeto OLAP Mondrian permite sua utilização para a modelagem e o processamento simplificado de consultas em modelos multidimensionais. Sendo o Mondrian um projeto *open-source*, foi selecionado para o processamento de consultas em uma nova ferramenta que satisfaça as necessidades citadas anteriormente: Neuro BI.

#### 4. A Ferramenta Neuro BI

A ferramenta Neuro BI foi desenvolvida para a Web, o que classifica-a como uma ferramenta WOLAP (Web-OLAP). A ferramenta prevê sua utilização por três tipos de usuários (agentes do sistema): usuário do sistema, usuário administrador e usuários simples. O principal recurso da ferramenta é a possibilidade de um usuário (do tipo usuário administrador) visualizar, alterar e salvar (persistir) uma consulta multidimensional sem escrever em uma linguagem técnica. Esta linguagem é denominada MDX (*Multidimensional Expressions*), linguagem padrão de consultas multidimensionais (OLAP).

##### 4.1 Arquitetura

O projeto Neuro engloba a tecnologia Mondrian, entre outras, e também utiliza-se da mesma linguagem de desenvolvimento multiplataforma: Java. Desenvolvido para o ambiente Web, a ferramenta Neuro é processada em um servidor Web que comunica-se com uma base de dados onde encontram-se os dois principais bancos de dados da ferramenta. Um banco de dados diz respeito ao repositório de dados do sistema Neuro BI e o outro banco de dados é o DW onde estão os dados que são processados a cada consulta OLAP, conforme configurado (mapeado) no sistema.

Em Fernandes e Kantorski (2008) é possível se ter uma visão geral do sistema Neuro e seu funcionamento, conforme a figura adaptada (Figura 1).

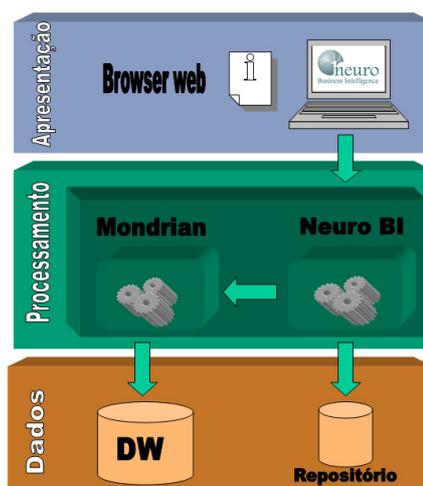


Figura 1. Arquitetura da ferramenta Neuro

Conforme a Figura 1, em um nível superior há a camada de apresentação do sistema e sua interação com o usuário através de um *browser* Web (necessário para

qualquer aplicação Web). Abaixo está a camada intermediária, mais complexa. A camada intermediária é a camada do servidor, onde é feito o processamento da lógica do sistema e a solicitação de consultas a camada inferior: camada de base de dados. Na camada de base dados está o repositório com as informações do sistema e o DW onde estão os dados para consultas OLAP.

Existem dois bancos de dados sob a base de dados manipulada pela ferramenta Neuro BI. O primeiro, o repositório de dados da ferramenta, possui informação sobre os usuários, cubos, consultas, entre outras entidades persistidas, conforme mostrado na Figura 2, enquanto o segundo banco é o DW modelado conforme a aplicação da ferramenta. O repositório de informações da ferramenta Neuro permite o armazenamento das informações referentes ao controle de acesso e às informações das consultas MDX.

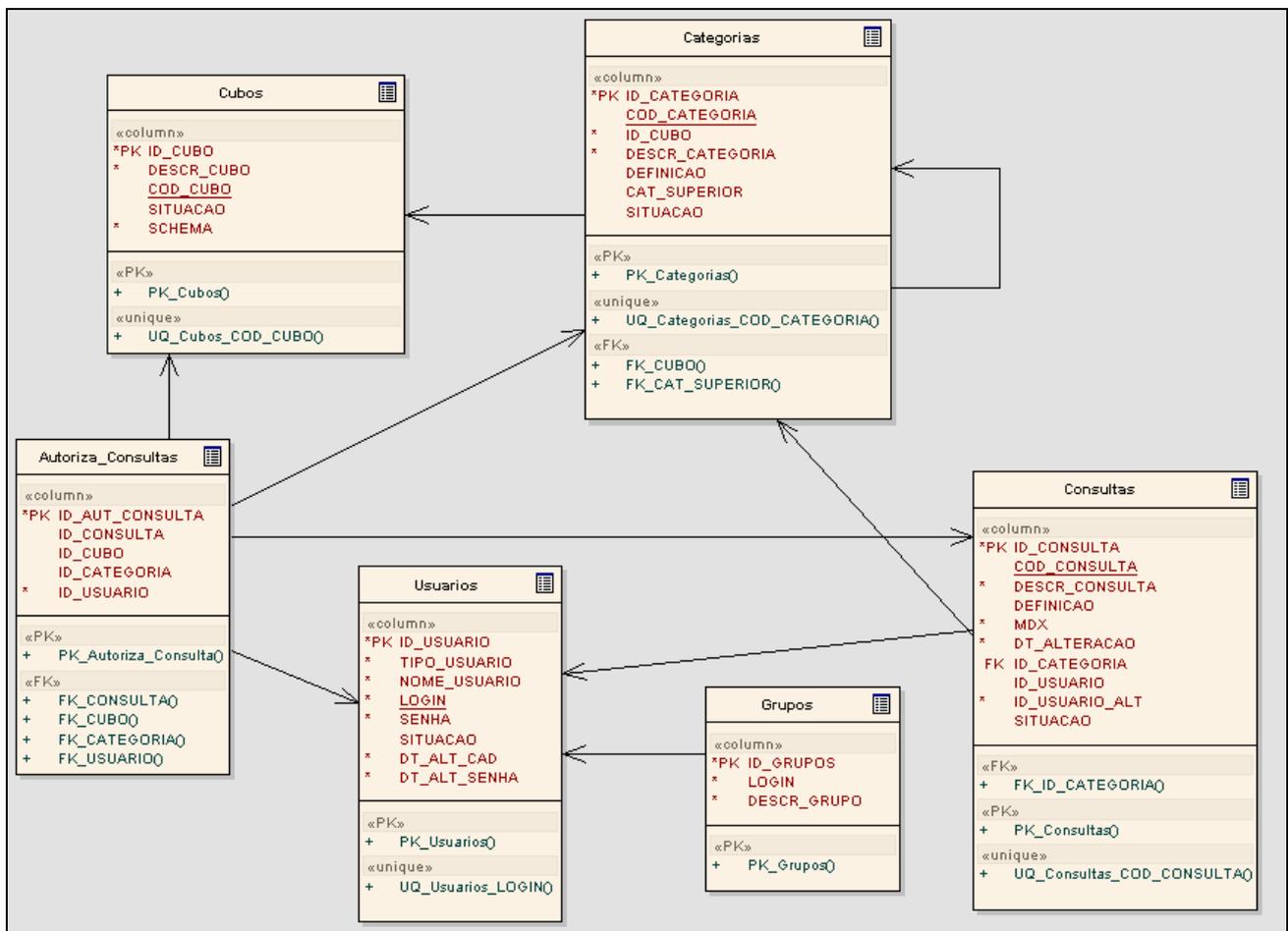


Figura 2. Modelo de dados

O Mondrian realiza as operações OLAP em um DW, previamente configurado. A integração das tecnologias (Neuro e Mondrian) exige configurações em nível de codificação. Primeiramente, o Mondrian manipula arquivos XMLs (*eXtensible Markup Language*) denominados *Schemas* que possuem o mapeamento do DW para o modelo multidimensional (cubos, dimensões, medidas etc.). Estes arquivos devem estar em diretórios que serão visualizados pelas páginas JSP que realizarão as consultas. No Neuro, estes arquivos estão no diretório denominado *schemas*. A figura 3 ilustra um exemplo de mapeamento do DW.

```

<?xml version="1.0" encoding="UTF-8" ?>
- <Schema name="Vestibular">
- <Cube name="Vestibular">
  <Table schema="osbi" name="bi_fat_vestibular_2008" />
  + <Dimension name="inscritos" foreignKey="id_inscricao">
  + <Dimension name="situacao" foreignKey="cod_situacao">
  - <Dimension name="cursos" foreignKey="id_opcao">
    - <Hierarchy name="cursos" hasAll="true" allMemberName="All cursos" primaryKey="id_curso">
      <Table schema="osbi" name="bi_dim_cursos" />
      - <Level name="Nome do Curso" schema="osbi" table="bi_dim_cursos" column="descr_curso" nameColumn="descr_curso"
        uniqueMembers="false">
        <Property name="codigo" column="cod_curso" />
        <Property name="identificado" column="id_curso" />
      </Level>
    </Hierarchy>
  </Dimension>
  + <Dimension name="cotas" foreignKey="id_cota">
  + <Dimension name="tempo" foreignKey="id_tempo">
  + <Measure name="Cotistas" column="ind_cota" aggregator="sum" datatype="Integer" formatString="###0" />
  + <Measure name="pontuacao" column="total_pontos" aggregator="sum" datatype="Numeric" formatString="###0.###" />
  + <Measure name="inscritos" column="id_inscricao" aggregator="count" datatype="Integer" formatString="###0" />
  + <Measure name="Menor pontuacao" column="total_pontos" aggregator="min" datatype="Integer" formatString="###0" />
  + <Measure name="Maior pontuacao" column="total_pontos" aggregator="max" datatype="Integer" formatString="###0" />
</Cube>
</Schema>

```

Figura 3. Mapeamento Relacional - Dimensional

## 4.2 Funcionamento

O projeto Neuro BI foi desenvolvido sobre a tecnologia Mondrian, através de uma configuração de integração, mas possui uma identidade própria, isto é, agrega funcionalidades que é justamente sua proposta. Um dos desafios no projeto de aplicações OLAP refere-se a como aliar a simplicidade para o usuário à expressividade nas consultas e análises. Parte da resposta é obtida através da maneira como os dados são organizados no modelo dimensional (Donald, 1997; Kimball; Ross, 2002).

Após a identificação na ferramenta Neuro, o usuário visualiza a tela inicial (Figura 4) que possui um menu lateral (à esquerda) e as opções principais: acessar os cubos existentes (autorizados) ou acessar as consultas salvas.

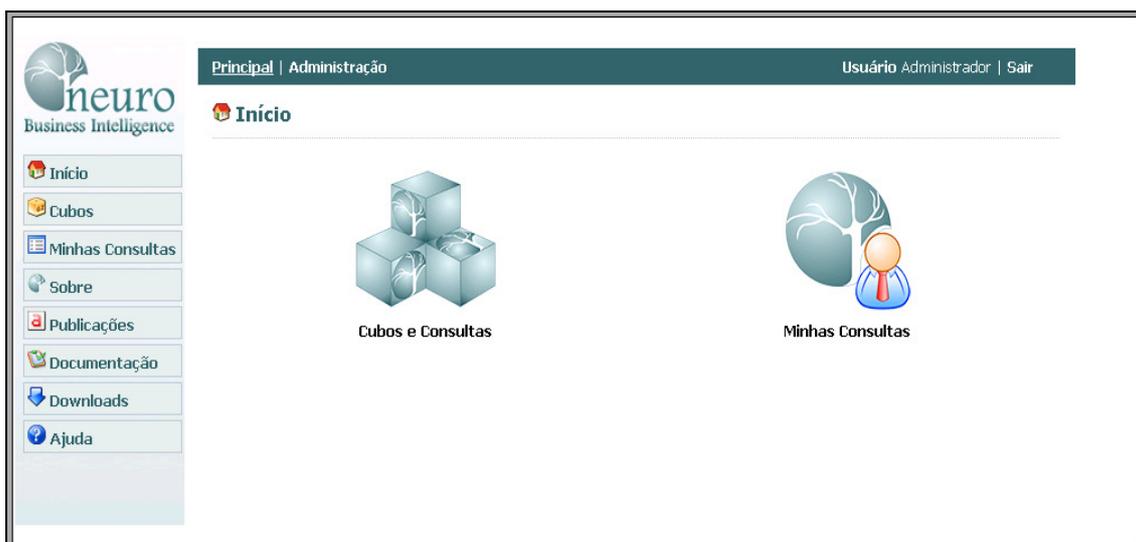


Figura 4. Tela inicial do Neuro

Navegando nos cubos, categorias e consultas o usuário chegará até a visualização da consulta (Figura 5). Nesta etapa são mostrados o título da consulta, sua descrição e o conteúdo da consulta. Além disto, existe o menu lateral (que poderá ser personalizado pelo usuário) e barra de botões, na parte superior, com as operações OLAP tradicionais.

Alguns botões agregam características adicionais como o botão de visualização do comando MDX:

- Linguagem MDX é ocultada do usuário (podendo ser visualizada, caso o usuário desejar);
- Possui um novo botão “Salvar Alterações” que permite ao usuário, ao alterar a consulta atual, salvá-la com as respectivas alterações;
- Possui outro botão “Copiar consulta” que possibilita o usuário alterar a consulta atual e gerar uma nova consulta baseada nesta.

The screenshot shows the Neuro Business Intelligence interface. The main content area displays a table titled 'Maior e menor pontuação pos Sistema de Cotas'. The table has columns for 'cotas', 'cursos', and 'Measures' (with sub-columns for 'inscritos', 'Menor pontuacao', and 'Maior pontuacao'). The data is as follows:

cotas	cursos	inscritos	Menor pontuacao	Maior pontuacao
-All cotas	+All cursos	16.902	0	136
Não Cotista	+All cursos	10.392	0	136
Sistema Cidadao Presente C - escolas publicas	+All cursos	6.042	0	135
Sistema Cidadao Presente A - afro-brasileiros	+All cursos	394	0	122
Sistema Cidadao Presente B - necessidades especiais	+All cursos	51	0	89
Sistema Cidadao Presente D - indigenas	+All cursos	23	0	0

**Figura 5. Tela de visualização de consulta**

Para fins de gerenciamento dos *schemas*, cubos, categorias e consultas da ferramenta Neuro existe uma seção de configurações (Figura 6) em que é possível listar, visualizar, editar, cadastrar ou excluir um dos itens anteriores.

The screenshot shows the 'Configurações: Cubo' page in the Neuro Business Intelligence interface. It includes a navigation menu on the left and a main content area with the following sections:

- Cubo**
  - Novo Cubo
  - Listar
- Categorias**
  - Nova Categoria
  - Listar
- Consultas**
  - Novo Consulta
  - Listar
- Esquemas (Schemas)**
  - Novo/Listar

**Figura 6. Configurações da ferramenta**

A ferramenta possui controle de usuários e a delegação de autorizações de visualização de cubos a estes. A premissa é que nem todas as consultas dizem respeito a todos os administradores (supostos usuários do Neuro) da instituição.

Primeiramente na ferramenta Neuro é realizado o cadastro com as informações básicas do usuário (Figura 7). Este usuário inicialmente ao acessar a ferramenta não visualizará nenhum cubo, consulta ou categoria, pois precisa ser vinculado a ele, pelo usuário administrador, o devido cubo, categoria ou consulta para visualização.

The screenshot shows the 'Cadastrar: Usuário' page. At the top, there is a navigation bar with 'Principal | Administração' and 'Usuário Administrador | Sair'. The left sidebar contains a menu with items: Início, Configurações, Feedback, Sobre, Publicações, Documentação, Downloads, and Ajuda. The main content area has the title 'Cadastrar: Usuário' and a description: 'Nesta seção é possível cadastrar e alterar uma novo Usuário, também listar os Usuários já cadastrados.' Below this is a form titled 'Dados do Usuário' with the following fields: 'Nome \*' (text input), 'Login \*' (text input), 'Senha \*' (password input), 'Confirmação \*' (password input), 'Situação' (dropdown menu with 'Ativo' selected), and 'Tipo de Usuário \*' (dropdown menu with 'Visualização' selected). There are three buttons at the bottom: 'Salvar', 'Configurações', and 'Listar'. A note at the bottom right says '(\*) Dado obrigatório.'

Figura 7. Cadastro de Usuários.

Uma vez cadastrado, os usuários podem ser listados, editados, excluídos ou podem ter autorizações cadastradas. O procedimento de cadastro de autorizações corresponde a listagem de todas as cubos, categorias e consultas em quem são selecionados os itens desejados ao devido usuário (Figura 9). Esta autorização poderá ser visualizada futuramente, assim como poderá ser excluída.

The screenshot shows the 'Cadastrar: Autorização' page. At the top, there is a navigation bar with 'Principal | Administração' and 'Usuário Administrador | Sair'. The left sidebar is the same as in Figure 7. The main content area has the title 'Cadastrar: Autorização' and a description: 'Nesta seção é possível cadastrar e alterar uma nova permissão de visualização, aos Usuários já cadastrados.' Below this is a form titled 'Dados da Autorização' with two expandable sections: 'Usuários' and 'Cubos'. Under 'Usuários', 'Mikael (usuario)' is checked. Under 'Cubos', there are two main categories: 'Vestibular 2008 - Dados Estatísticos Provas' (checked) and 'Vestibular' (unchecked). The 'Vestibular 2008' category has sub-items: '[+] Administração do Cubo e Consultas' (checked) and '[+] Informações sobre as Provas' (checked). The 'Vestibular' category has sub-items: '[+] Vestibular 2008' (unchecked) and '[+] Administração do Cubo' (unchecked). There are three buttons at the bottom: 'Salvar', 'Configurações: Autorizações', and 'Listar'. A note at the bottom right says '(\*) Dado obrigatório.'

Figura 8. Cadastro de Autorizações em Cubos e Consultas

## 5. Considerações Finais

Este artigo apresentou uma ferramenta OLAP *open source* para apoio a tomada de decisões pelos usuários. A principal contribuição da ferramenta desenvolvida é a possibilidade do usuário criar novas consultas a partir da visualização de todas as dimensões, da tabela de fatos e poder executar as operações OLAP sem a necessidade de conhecer a linguagem mdx. As novas consultas criadas podem ser persistidas em um repositório e serem disponibilizadas para usuários autorizados. A ferramenta desenvolvida permite que usuários do nível gerencial das instituições possam manipular os cubos de dados sem a necessidade de intervenção técnica. A tabela 2 apresenta uma síntese das contribuições da ferramenta em comparação com a tabela 1.

**Tabela 2 – Tabela comparativa das características das tecnologias de análise e a ferramenta Neuro**

<b>Tecnologias/ Características de Análise</b>	<b>Pentaho + Neuro</b>	<b>Microsoft</b>	<b>IBM</b>
Open-source	Sim	Não	Não
Nível de conhecimento dos usuários	Gerencial	Técnico	Técnico
Controle de acesso a consultas	Sim	Não	Não
Persistência de consultas	Sim	Não	Não

A ferramenta Neuro é um projeto em constante desenvolvimento, paralelamente a pesquisa envolvendo outros acadêmicos. Destaca-se como trabalhos futuros a criação de um agente para assessorar a criação de consultas MDX e suas métricas. A ferramenta está disponível em <http://portal.ufsm.br/neuro> para consulta.

## Referências

- Barbieri, Carlos. BI – *Business Intelligence* – Modelagem e Tecnologia. Rio de Janeiro: Axcel Books do Brasil, 2001.
- \_\_\_\_\_. *Cognos*. Site Web acessado em março de 2010. <http://www.cognos.com/>
- Donald, B. *High Performance Oracle Data Warehousing*. USA: The Coriolis Group, 1997.
- Fernandes, Mikael S.; Kantorski, Gustavo Z. Uma Ferramenta WEB para apoio à Decisão em Ambiente Hospitalar. Espanha, Universidad de Vigo: *3ª Conferencia Ibérica de Sistemas y Tecnologías de la Información*, 2008.
- Fernandes, Mikael S.; Kantorski, Gustavo Z. Neuro BI: Uma Ferramenta *Web open source* para apoio à Tomada de Decisão. In: XIII Simpósio de Informática, Uruguaiana, RS, 2008. ISSN: 1983-6511.
- Inmon, W. H. *Como construir o Data Warehouse*. 2 ed. Rio de Janeiro: Campus, 1997.
- Kimball, Ralph; Ross, Margy. *The Data Warehouse Toolkit: guia completo para modelagem dimensional*. 2 ed. Rio de Janeiro: Campus, 2002.
- Navathe, S. e Elmasri, R. (2000) *Fundamentals of Database Systems*, Addison Wesley, 3<sup>th</sup> edition. \_\_\_\_\_, *Microsoft: IT Library*. Site Web acessado em outubro de 2008. <http://www.windowstlibrary.com/Content/1646/01/toc.html>
- Pentaho. (2010) “Open Source Business Intelligence”, <http://www.pentaho.org>, December.
- Thomsen, E. (2002), *OLAP Solutions: Building Multidimensional Information Systems*, 2<sup>th</sup> edition, John Wiley & Sons, New York, NY .