

# Uso da ferramenta Pentaho BI-Server para obtenção de informações gerenciais do Sistema Aberto e Integrado de Informação em Agricultura (Sabiia)

Anderson Hirata<sup>1</sup>  
Isaque Vacari<sup>2</sup>

A base de dados do Sistema Aberto e Integrado de Informação em Agricultura (Sabiia)<sup>3</sup>, com aproximadamente 275.000 registros armazenados, é composta basicamente de metadados bibliográficos de publicações científicas (livros, capítulos de livros, artigos em periódicos, folhetos, teses, anais e proceedings de eventos) coletadas em provedores de dados (repositórios institucionais, repositórios temáticos e periódicos científicos) de acesso aberto na área de agricultura e afins. Também compõem a base de dados do sistema Sabiia os dados cadastrais dos 128 provedores de dados, e os dados de buscas efetuadas no sistema.

De acordo com (PETRINI et al., 2006), Business Intelligence (BI) é um conjunto de tecnologias cujo objetivo é prover e oferecer suporte a um ambiente de informação permitindo às organizações acessar e explorar as informações referentes ao seu negócio que normalmente estão armazenadas em um DataMart e/ou Data Warehouse.

Nesse contexto de ambiente de informação, acesso e extração de informações gerenciais da base de dados do sistema Sabiia, o projeto definiu como objetivo estratégico o uso de soluções livres de BI para obtenção de informações e respostas gerenciais, tais como: Quantos provedores de dados estão sendo coletados? Quais são os provedores de dados coletados?

---

<sup>1</sup> Pontifícia Universidade Católica de Campinas, [hianderson@cnptia.embrapa.br](mailto:hianderson@cnptia.embrapa.br)

<sup>2</sup> Embrapa Informática Agropecuária, [isaque@cnptia.embrapa.br](mailto:isaque@cnptia.embrapa.br)

<sup>3</sup> Sistema Aberto e Integrado de Informação em Agricultura. Disponível em: <<http://www.embrapa.br/sabiia>>.

em cada país ? Qual é a quantidade de consultas efetuadas por ano, mês e país no sistema ? Quais são as buscas mais efetuadas no sistema ?

Devido à maior experiência da Embrapa Informática Agropecuária no uso de soluções livres e arquitetura Java Enterprise Edition (Java EE)<sup>4</sup> em projetos de pesquisa, o projeto Sabiia optou pelo uso da ferramenta Pentaho-BI Server<sup>5</sup> como solução de (BI) para obtenção de informações gerenciais.

A Pentaho Corporation gerencia e lidera o papel de desenvolvimento do Projeto Pentaho-BI - uma iniciativa da comunidade de desenvolvimento Open Source para oferecer ferramentas robustas de BI de código aberto. A suíte Pentaho-BI é composta de ferramentas de consultas e relatórios gerenciais, análise de dados OnLine Analytical Processing (OLAP), tratamento de dados Extraction, Transformation, Load (ETL), mineração de dados (Data Mining) e Workflow.

Após a instalação e configuração da ferramenta Pentaho-BI Server, decidiu-se implementar duas consultas piloto em ambiente de desenvolvimento para obtenção de informações gerenciais relativas às consultas efetuadas no sistema Sabiia e às informações referentes aos provedores de dados coletados. Para construir as consultas piloto e visualizar as informações gerenciais obtidas, os seguintes módulos da ferramenta Pentaho-BI Server foram utilizados:

- Pentaho Data Source – Módulo usado para definição das fontes de dados (conexão com o banco de dados) e consultas Structured Query Language (SQL) - Linguagem de Consulta Estruturada vinculadas ao banco de dados do sistema Sabiia;
- Pentaho Analysis Mondrian - O Mondrian é o servidor OLAP, ele recebe solicitações de um cliente e as submete a um banco de dados tradicional, mapeando as consultas multidimensionais para um banco de dados relacional. O módulo Analisys também oferece recursos para construção da definição de cubos, visualização das consultas retornadas, contendo operações básicas de manipulações de cubos multidimensionais e de apresentação dos resultados em tabelas e gráficos;

---

<sup>4</sup> Java Platform, Enterprise Edition. Disponível em: <<http://java.sun.com/javaee>>.

<sup>5</sup> Pentaho Community. Disponível em: <<http://community.pentaho.com>>.

- Pentaho Reporting Designer - Este módulo oferece ferramentas para criação de relatórios via Web browser (também conhecido como relatórios *Ad-Hoc*).
- Tomcat - O Apache Tomcat é uma solução livre que atua como servidor Web de sistemas desenvolvidos em Java, necessário para executar o Pentaho-BI Server.

O uso do software livre Pentaho-BI Server mostrou-se adequado para obtenção de informações gerenciais do sistema Sabiia, destacando a facilidade de se construir e acessar todas as funcionalidades e resultados gerados dinamicamente pela ferramenta por meio de uma interface web amigável e intuitiva. Vislumbra-se que o domínio tecnológico sobre ferramentas de BI livres contribuirão no atendimento de demandas de análises de dados e extração de informações para seus usuários de forma mais rápida e intuitiva. A próxima etapa do projeto consiste em disponibilizar no ambiente de produção do projeto Sabiia todo ambiente de BI (Pentaho-BI Server, consultas, análises de dados e relatórios) com acesso restrito aos gestores do projeto.

## Referências

- FERREIRA, M.; PIAUHY, C.; CARVALHO, J.; SILVA, R.; VIEIRA, V. **Um estudo de caso com análise comparativa entre ferramentas de BI livre e proprietária**. In: Escola Regional de Banco de Dados (ERBD), 2010, Joinville. Anais... Porto Alegre: SBC, 2010.
- PETRINI, M.; FREITAS, M. T.; POZZEBON, M. (2006). **Inteligência de negócios ou inteligência competitiva: noivo neurótico, noiva nervosa**. Encontro da Associação Nacional de Pós-Graduação e Pesquisa em Administração (EnANPAD).