

# Resumos

## VII Mostra de estagiários e bolsistas da Embrapa Informática Agropecuária

Campinas, 5 a 7 de dezembro, 2011

Maria Giulia Croce  
Marcos Dias Soares  
Nanci Cassini Manzoti Andrade  
Sílvia Maria Fonseca Silveira Massruhá  
Editores Técnicos



*Empresa Brasileira de Pesquisa Agropecuária  
Embrapa Informática Agropecuária  
Ministério da Agricultura, Pecuária e Abastecimento*

# Resumos

## **VII Mostra de estagiários e bolsistas da Embrapa Informática Agropecuária**

Campinas, 5 a 7 de dezembro, 2011

*Marcos Dias Soares  
Maria Giulia Croce  
Nanci Cassini Manzoti Andrade  
Sílvia Maria Fonseca Silveira Massruhá*  
Editores Técnicos

Embrapa Informática Agropecuária  
Campinas, SP  
2011

## **Embrapa Informática Agropecuária**

Av. André Tosello, 209 - Barão Geraldo  
Caixa Postal 6041 - 13083-886 - Campinas, SP  
Fone: (19) 3211-5700 - Fax: (19) 3211-5754  
www.cnptia.embrapa.br  
sac@cnptia.embrapa.br

## **Comitê de Publicações**

Presidente: *Silvia Maria Fonseca Silveira Massruhá*

Membros: *Poliana Fernanda Giachetto, Roberto Hiroshi Higa, Stanley Robson de Medeiros Oliveira, Maria Goretti Gurgel Praxedes, Adriana Farah Gonzalez, Neide Makiko Furukawa*

Membros suplentes: *Alexandre de Castro, Fernando Attique Máximo, Paula Regina Kuser Falcão*

Supervisor editorial: *Stanley Robson de Medeiros Oliveira, Neide Makiko Furukawa*

Revisor de texto: *Adriana Farah Gonzalez*

Normalização bibliográfica: *Maria Goretti Gurgel Praxedes*

Editoração eletrônica: *Neide Makiko Furukawa*

Secretária: *Carla Cristiane Osawa*

Capa: *Neide Makiko Furukawa*

## **1ª edição on-line 2011**

### **Todos os direitos reservados.**

A reprodução não autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei no 9.610).

### **Dados Internacionais de Catalogação na Publicação (CIP) Embrapa Informática Agropecuária**

---

Mostra de Estagiários e Bolsistas da Embrapa Informática  
Agropecuária (7. : 2011 : Campinas, SP).

Resumos da VII Mostra de Estagiários e Bolsistas da Embrapa  
Informática Agropecuária, Campinas, SP, 5 a 7 de dezembro de  
2011 / Editores técnicos, Marcos Dias Soares... [et al.]. – Campinas :  
Embrapa Informática Agropecuária, 2011.  
98 p.

ISBN 978-85-86168-07-9

1. Agroinformática. 2. Tecnologia da informação 3. Iniciação  
científica. 4. Mudanças climáticas. I. Mostra de Estagiários e  
Bolsistas da Embrapa Informática (7. : Campinas, SP). II. Croce,  
Maria Giulia. III. Embrapa Informática Agropecuária. IV. Título.

CDD 630.285 (21st ed.)

# Apresentação

A Chefia desta unidade tem a satisfação de apresentar os resumos da VII Mostra de Trabalhos de Bolsistas e Estagiários da Embrapa Informática Agropecuária, realizada de 5 a 7 de dezembro de 2011.

A Mostra de Estagiários da Embrapa Informática Agropecuária tem como objetivo complementar o desenvolvimento profissional dos estagiários e bolsistas que participam do Programa de Estágio de Complementação Educacional da Embrapa, junto às equipes de pesquisa e apoio da Unidade e Instituições parceiras, Centro de Pesquisas Meteorológicas e Climáticas Aplicadas à Agricultura (Cepagri/Unicamp), Rede Nacional de Pesquisa - RNP e Escritório de Negócios de Campinas da Embrapa SNT.

Além de promover a experiência de preparação de um pôster utilizando-se de metodologia científica e da apresentação oral de seu conteúdo, o evento também contribui para uma maior integração entre o estudante e seu supervisor nas atividades conjuntas de preparação para a exposição. Acreditamos ainda que ações como essas contribuam para o sucesso profissional desses estudantes.

A Embrapa Informática Agropecuária parabeniza e agradece a dedicação de todos os estudantes e seus orientadores, do Comitê Local de Publicações bem como o apoio da área administrativa, especialmente o Setor de Gestão de Pessoas, da área de Comunicação e Negócios e de todos àqueles que contribuíram para que a VII Mostra de Estagiários e Bolsistas se tornasse realidade.

***Silvia Maria Fonseca Silveira Massruhá***

Chefe de Pesquisa e Desenvolvimento  
Embrapa Informática Agropecuária



# Sumário

<b>Expansão de busca utilizando vocabulário controlado .....</b>	<b>9</b>
Marcos Aparecido Marinho Seixas; Carlos Miguel Tobar Toledo; Maria Fernanda Moura	
<b>Amostragem de indivíduos representativos em uma população de animais da raça Canchim aparentados e genotipados.....</b>	<b>13</b>
Ismael Urbinati; Roberto Hiroshi Higa; Fabiana Barichello Mokry	
<b>Framework de mineração de textos para análise de dados de bioinformática.....</b>	<b>16</b>
Rafael Seiji Tanaka; Rafael Hideo Kashiwara; Roberto Hiroshi Higa; Maria Fernanda Moura	
<b>Uma ferramenta para cadastro e mapeamento dos clientes do website Catálogo de Produtos e Serviços da Embrapa .....</b>	<b>20</b>
Anderson Carlos Ferreira da Silva; Leandro Henrique Mendonça de Oliveira	
<b>A Terminologia aplicada ao projeto INTAGRO nas perspectivas da organização e difusão da informação e da representação do conhecimento .....</b>	<b>22</b>
Carolina Cardoso de Oliveira; Ivo Pierozzi Jr.; Leandro Henrique Mendonça de Oliveira; Dayse Simon L. Kamikawachi	
<b>Uso de ferramenta para construção de questionários em plataforma web: uma experiência no projeto iAgência .....</b>	<b>25</b>
Matheus Silva Miranda; Adriana Delfino dos Santos	
<b>Desenvolvimento de software para a detecção de grupos de genes homólogos sob evidência de seleção positiva .....</b>	<b>28</b>
Jorge Augusto Hongo; Francisco Pereira Lobo	
<b>Ferramenta para gerenciamento de temas em Webgis baseados no i3geo.....</b>	<b>32</b>
Renato Luiz Franceschinelli; Eduardo Antonio Speranza; João dos Santos Vila da Silva	

<b>WikiPragas: plataforma Wiki para fichas de pragas .....</b>	<b>35</b>
Gabrielle Consolini Salazar Alarcon; Carlos Alberto Alves Meira; Anderson Carlos Bueno dos Santos; Bruno Franciscon Mazzotti; Cristiano Borges Cardoso; Isaque Vacari; Marcos Cezar Visoli; Tiago Augusto Pimenta	
<b>Uma ferramenta para cadastro e visualização de metadados em Agricultura de Precisão utilizando GeoNetwork .....</b>	<b>39</b>
Pedro Eduardo Bertelli Corrêa; Eduardo Antonio Speranza; Leonardo Ribeiro Queiros	
<b>JQuery ou Dojo, qual biblioteca JavaScript usar?.....</b>	<b>42</b>
Juliana Pereira de Santana; Luciana Alvim Santos Romani	
<b>Flexibilidade para aplicações web desenvolvidas com o FlexBuilder .....</b>	<b>45</b>
Gabriel Garcia do Nascimento; Luciana Alvim Santos Romani	
<b>Modernizando a apresentação de trabalhos com PREZI.....</b>	<b>49</b>
Marcelo Ceron Massruhá; Luciana Alvim Santos Romani	
<b>Espacialização de notícias de cana-de-açúcar utilizando entidades do IBGE .....</b>	<b>52</b>
Ercília Souza Rodrigues; Eduardo Antonio Speranza; Maria Fernanda Moura; Rosa Nathalie Portugal Vargas; Solange de Oliveira Rezende	
<b>Execução dos ajustes e refinamentos da rede de drenagens do Estado do Mato Grosso do Sul no projeto GeoMS .....</b>	<b>55</b>
Ana Cláudia Olivera Senturião; Júlio César Dalla Mora Esquerdo	
<b>Calibração do modelo cropsyst para cana-de-açúcar: estudo preliminar .....</b>	<b>58</b>
Felipe Ferreira Bocca; Alexandre de Castro	
<b>Atualização da rede de drenagens e inserção de sua nomenclatura no Estado do Mato Grosso do Sul .....</b>	<b>63</b>
Letícia Beatriz de Barros Brito; Júlio César Dalla Mora Esquerdo	
<b>Uso da ferramenta Pentaho BI-Server para obtenção de informações gerenciais do Sistema Aberto e Integrado de Informação em Agricultura (Sabiia).....</b>	<b>66</b>
Anderson Hirata; Isaque Vacari	



<b>Verificação da acurácia do mapeamento de áreas de agricultura localizadas em áreas desflorestadas na Amazônia .....</b>	<b>69</b>
Laís Silva de Oliveira; Júlio César Dalla Mora Esquerdo; Alexandre Camargo Coutinho	
<b>Busca usando sinônimos no Ainfo-Consulta .....</b>	<b>72</b>
Rogério Lecarião Leite; Edmilson José Mangueira Carvalho; Glauber José Vaz	
<b>Utilização da metodologia WGCNA na construção de redes gênicas e identificação de mecanismos envolvidos na resposta de vacas à mastite .....</b>	<b>76</b>
Vinícius Campidelli Gozzo; Poliana Fernanda Giachetto	
<b>Utilização da ferramenta Liferay na construção do portal do Laboratório Multiusuário de Bioinformática da Embrapa - LMB .....</b>	<b>80</b>
Fernanda Cristina de Paiva Pereira; Glauber José Vaz; Francisco Pereira Lobo; Michel Eduardo Beleza Yamagishi; Poliana Fernanda Giachetto	
<b>Análise da rotação de cultura na área de reforma de canavial .....</b>	<b>83</b>
Caroline Nascimento Pereira; Aryeverton Fortes de Oliveira	
<b>Parametrização de modelo compartimental do tipo SEIDR para avaliação quantitativa da HLB em citros.....</b>	<b>87</b>
Guilherme Augusto Braga; Sônia Ternes; Raphael Gustavo D’Almeida Vilamiu	
<b>Análise de custos de tecnologias da Embrapa Informática Agropecuária .....</b>	<b>91</b>
Fernanda da Rocha Fernandes; Aryeverton Fortes de Oliveira	
<b>Relação entre a radiação fotossinteticamente ativa e a radiação global para diferentes classificações de céu .....</b>	<b>96</b>
Fernanda Ferraz Brocardi; Giuliana Amalfi Pinto; Gustavo Costa Rodrigues	



# Expansão de busca utilizando vocabulário controlado

Marcos Aparecido Marinho Seixas<sup>1</sup>

Carlos Miguel Tobar Toledo<sup>1</sup>

Maria Fernanda Moura<sup>2</sup>

## Introdução

No site da Agência de Informação Embrapa, ou simplesmente Agência, para facilitar a localização de informações desejadas pelos usuários, além de outros recursos de navegação, implantou-se um serviço de busca automática (CRUZ, 2003). Esse serviço de busca utiliza uma ferramenta open source chamada swish-e, que realiza a busca e a indexação dos hipertextos (no caso, páginas estáticas). A ferramenta de busca trata de assuntos gerais relacionados à agropecuária e com público variado, tais como publicações de cunho científico, publicitário, jornais, etc. A busca atual é realizada apenas nos metadados (SOUZA et al., 2004) dos recursos disponíveis no site, que são catalogados a partir do padrão Dublin Core. Dessa forma, os metadados obedecem a um padrão definido, e o assunto dos documentos, as palavras-chaves e as categorias agrícolas obedecem constantemente a um thesaurus – especificado junto ao metadado. Porém, toda essa qualificação prévia dos dados nem sempre é explorada pelo usuário final. Muitas vezes, as expressões de busca, definidas pelo usuário, poderiam ser expandidas, de acordo com os thesaurus utilizados para melhorar as respostas obtidas, seja na precisão ou na cobertura destas. Além disso, com uma possível expansão dessas buscas, seria interessante, além de mostrar, individualmente, o percentual de relevância das respostas, mostrá-lo também em uma visualização hierárquica.

---

<sup>1</sup> Pontifícia Universidade Católica de Campinas (PUCAMP),  
marcosams@cnptia.embrapa.br; tobar@puc-campinas.edu.br

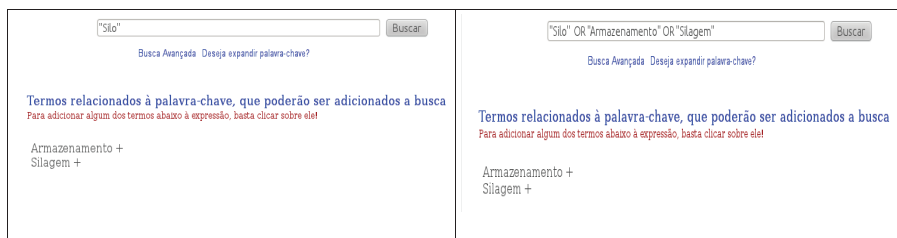
<sup>2</sup> Embrapa Informática Agropecuária, fernanda@cnptia.embrapa.br

Assim, o objetivo dessa proposta é expandir as expressões de busca especificadas pelo usuário, de modo a abrir o leque de resultados, melhorando especialmente a cobertura destes, conferindo uma maior precisão nos resultados das buscas.

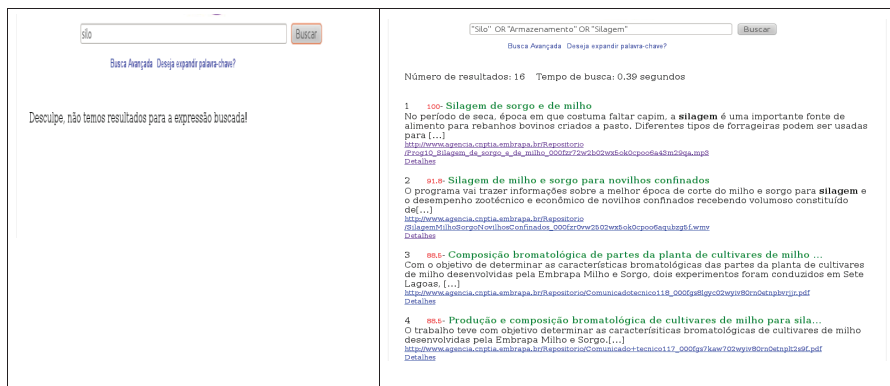
## Material e métodos

A evolução da ferramenta de busca, para incorporar a expansão, foi possibilitada pelo reuso da implementação de um Web Service, que permite acessar o banco de vocabulário controlado da Agência (SOUZA et al., 2010).

Conforme ilustrado na Figura 1 a), o usuário pode escolher se quer expandir a busca. A expansão automática procura no vocabulário controlado (via Web Service), pelos termos que devem ser usados no lugar do termo indicado (“use”) ou dos termos que podem ser usados em seu lugar (“used for”) e já os acrescenta à expressão de busca com o uso do operador “OR”. Além disso, se o usuário desejar pode acrescentar termos relacionados (“related to – RT”) aos termos de sua expressão de busca, podendo escolhê-los na relação apresentada. Supondo que o usuário escolheu o termo “Silo” (Figura 1), então a expressão é expandida para a forma apresentada na parte a) automaticamente, e a parte b) apresenta a expressão com os “related to” incluídos pelo usuário. A Figura 2 ilustra os resultados obtidos no caso de o usuário ter utilizado apenas a expressão “silo” e depois com a expansão escolhida na Figura 1 b). Nota-se, subjetivamente,



**Figura 1.** Expansão semântica de palavras-chave: a) usuário escolhe se quer expandir a busca; b) termos USE UF e RT.



**Figura 2.** Comparativo dos resultados das buscas.

A evolução da ferramenta de busca, para incorporar a expansão, foi possibilitada pelo reuso da implementação de um Web Service, que permite acessar o banco de vocabulário controlado da Agência (SOUZA et al., 2010).

Conforme ilustrado na Figura 1 a), o usuário pode escolher se quer expandir a busca. A expansão automática procura no vocabulário controlado (via Web Service), pelos termos que devem ser usados no lugar do termo indicado (“use”) ou dos termos que podem ser usados em seu lugar (“used for”) e já os acrescenta à expressão de busca com o uso do operador “OR”. Além disso, se o usuário desejar pode acrescentar termos relacionados (“related to – RT”) aos termos de sua expressão de busca, podendo escolhê-los na relação apresentada. Supondo que o usuário escolheu o termo “Silo” (Figura 1), então a expressão é expandida para a forma apresentada na parte a) automaticamente, e a parte b) apresenta a expressão com os “related to” incluídos pelo usuário. A Figura 2 ilustra os resultados obtidos no caso de o usuário ter utilizado apenas a expressão “silo” e depois com a expansão escolhida na Figura 1 b). Nota-se, subjetivamente, uma boa melhora, resultado da busca, utilizando-se efetivamente a qualificação imposta aos metadados.

## Resultados e discussão

Para validar objetivamente o processo, tem-se criado bases de dados artificiais, de modo a reproduzir o uso do vocabulário controlado tanto nos metadados quanto ao longo do textos. Por exemplo, em textos que tratam a cultura de cana-de-açúcar, criam-se novos textos que incorporam, em seus metadados e em seu corpo, os termos relacionados ao termo “cana-de-açúcar” e seus consequentes: cana-de-açúcar, canavieira, *Saccharum officinarum*, canavial, soca, soqueira, cultivo da soca. Dessa forma, tem-se uma base de textos que permite obter resultados de busca, para diferentes expressões de busca, com documentos relevantes e irrelevantes, considerando-se como documentos relevante só conjunto de documentos que contém, pelo menos, um dos termos relacionados ao assunto da expressão de busca, notado por DR.

Como trabalho futuro, será realizada uma análise da revocação e precisão das buscas em relação às expressões supostamente especificadas pelo usuário e às expressões com expansões aleatoriamente selecionadas. A revocação é a relação entre o número de documentos recuperados que efetivamente contém os termos da expressão de busca ( $tp$  – true positive) e o número total de documentos recuperados, e a precisão corresponde à relação entre  $tp$  e a cardinalidade de DR. A hipótese, estatisticamente esperada, é que se tenha em média uma melhor revocação para os resultados de expressões de busca expandidas.

## Referências

- CRUZ, S. A. B. da; CCRUZ, S. A. B. da. **Implantação de um serviço de busca em site da WWW**. Campinas: Embrapa Informática Agropecuária, 2003. 8 p. (Embrapa Informática Agropecuária. Comunicado técnico, 50).
- SOUZA, M. I. F.; MOURA, M. F.; SANTOS, A. D. dos. **Estudo comparativo entre os metadados da Agência de Informação Embrapa e do acervo documental do AINFO**. Campinas: Embrapa Informática Agropecuária, 2004. 10 p. (Embrapa Informática Agropecuária. Comunicado técnico, 66).
- SOUZA, M. I. F.; ALVES, M. das D. R.; QUEIROS, L. R.; SANTOS, A. D. dos; OLIVEIRA, L. H. M. de. Representação descritiva e temática no Sistema Agência de Informação Embrapa: controle de vocabulário. **Transinformação**, Campinas, v. 22, n. 1, p. 61-75, jan./abr. 2010.

# Amostragem de indivíduos representativos em uma população de animais da raça Canchim aparentados e genotipados

Ismael Urbinati<sup>1</sup>

Roberto Hiroshi Higa<sup>2</sup>

Fabiana Barichello Mokry<sup>3</sup>

Amostragem consiste em estudar um pequeno grupo, designado amostra, retirado de uma população, de modo que a amostra represente toda a população. Quando a população possui características que permitem a criação de subconjuntos, as amostras extraídas por amostragem simples são menos representativas. Nesse caso, é utilizada a amostragem estratificada (BARBETTA, 2007).

Os estratos podem ser delimitados pela análise de agrupamento, processo pelo qual se agrupa um conjunto de objetos físicos ou abstratos em classes de objetos similares (EVERITT et al., 2001).

O objetivo neste estudo foi avaliar diferentes métodos para selecionar uma amostra de uma população contendo animais aparentados, representando estratos de famílias, de forma que os animais selecionados representem toda a população.

Utilizaram-se duas fontes de informação para o estudo. A primeira foi o arquivo de pedigree (animal, pai e mãe) constituído de 6.801 animais da raça Canchim, sendo 400 animais genotipados. A segunda foi a matriz de parentesco constituída pelos coeficientes de parentesco dos 400 animais genotipados, organizados na forma de matriz, com os animais nas linhas e colunas e os correspondentes coeficientes de parentesco dentro da matriz (BOURDON, 2000).

---

<sup>1</sup> Universidade Estadual Paulista “Júlio de Mesquita Filho”,  
ismaelurbinati@zootecnista.com.br

<sup>2</sup> Embrapa Informática Agropecuária, roberto@cnptia.embrapa.br

<sup>3</sup> Universidade Federal de São Carlos, bolsista Capes, fbmokry@gmail.com

A seleção de animais representativos dessa população foi realizada utilizando três diferentes métodos. Todas as análises foram realizadas utilizando o software estatístico R (R DEVELOPMENT CORE TEAM, 2009).

**Método 1:** O arquivo de pedigree foi utilizado para selecionar os animais mais representativos. Com o auxílio do software estatístico R foram calculadas as distâncias entre cada linha do arquivo, pelo método Manhattan (EVERITT, 2001), e em seguida foi construído um dendrograma com os animais. Foram calculados os coeficientes de parentesco médio para cada animal utilizando a matriz de parentesco. O dendrograma foi particionado em 50 grupos distintos, e dentro de cada grupo foi amostrado o animal que possuía o maior parentesco médio.

**Método 2:** A própria matriz de parentesco foi utilizada para calcular uma matriz de dissimilaridade dada por  $D = \max(D) - D$ . Com essa matriz de dissimilaridade foi construído um dendrograma, posteriormente particionado em 50 grupos, dentro dos quais se amostrou o animal com o maior parentesco médio.

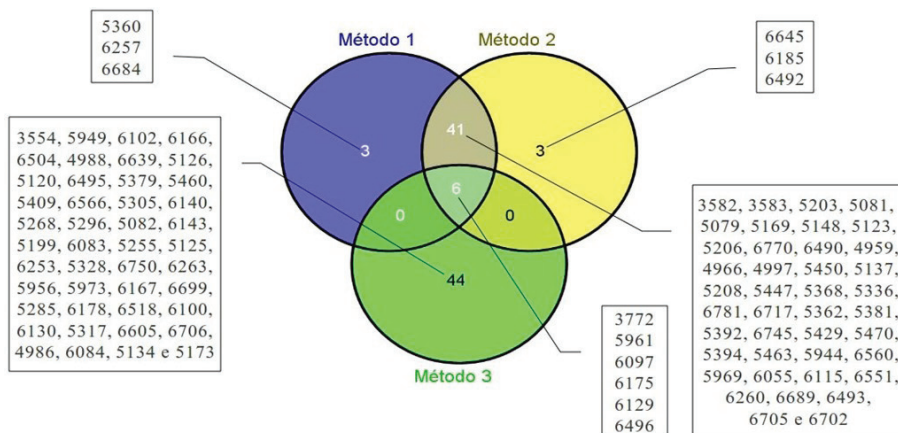
**Método 3:** Os animais foram amostrados de acordo com a seguinte rotina:

- 1º Calculou-se o parentesco médio de cada animal com todos os animais da matriz de parentesco (medida de representatividade);
- 2º Selecionou-se o animal de maior parentesco médio com a população;
- 3º Calculou-se o parentesco médio de cada um dos animais ainda não selecionados com os animais já selecionados;
- 4º Dentre aqueles animais de menor parentesco médio com os animais já selecionados (divergência genética), escolheu-se o animal de maior parentesco médio com os animais ainda não selecionados (representatividade);
- 5º Repetiu-se a rotina até completar o número de animais necessários (50 amostras).

Para comparar os diferentes métodos de amostragem, foi desenhado um diagrama de Venn (Figura 1) (OLIVEROS, 2007), com as quantidades de animais selecionados por cada método dentro de cada círculo, e nas intersecções a quantidade de animais que foram amostrados por mais de um método. Os números em destaque representam a numeração dos animais amostrados.

Os métodos 1 e 2 são mais parecidos entre si pois ambos utilizaram a divisão dos dendrogramas das distâncias entre os animais para a amostra-





**Figura 1.** Comparação entre os conjuntos de animais selecionados por cada um dos 3 métodos implementados.

gem, e o método 3 utilizou diretamente as médias de parentesco. Para trabalhos futuros, sugere-se a comparação dos três métodos utilizando, como fonte de informações, os dados moleculares (genótipos) dos animais.

## Referências

- BARBETTA, P. A. **Estatísticas aplicadas às Ciências Sociais**. 7. ed. Florianópolis: Ed. da UFSC, 2007. 340 p.
- BOURDON, R. M. **Understanding animal breeding**. 2nd ed. Upper Saddle River, NJ : Prentice Hall, 2000. 538 p. ill.
- EVERITT, B. S.; LANDAU, S.; LEESE, M. **Cluster analysis**. 4th ed. LoLondon: Arnold; New York: Oxford University, 2001. 237 p.
- OLIVEROS, J. C. VENNY. **An interactive tool for comparing lists with Venn Diagrams**. Disponível em: <<http://bioinfogp.cnb.csic.es/tools/venny/index.html>>. Acesso em: 1 out. 2011.
- R DEVELOPMENT CORE TEAM. **R: A language and environment for statistical computing**. Vienna, Austria: R Foundation for Statistical Computing, 2009. 409 p. Disponível em: <<http://www.R-project.org>>. Acesso em: 1 out. 2011.

## Framework de mineração de textos para análise de dados de bioinformática

Rafael Seiji Tanaka<sup>1</sup>  
Rafael Hideo Kashiwara<sup>1</sup>  
Roberto Hiroshi Higa<sup>2</sup>  
Maria Fernanda Moura<sup>2</sup>

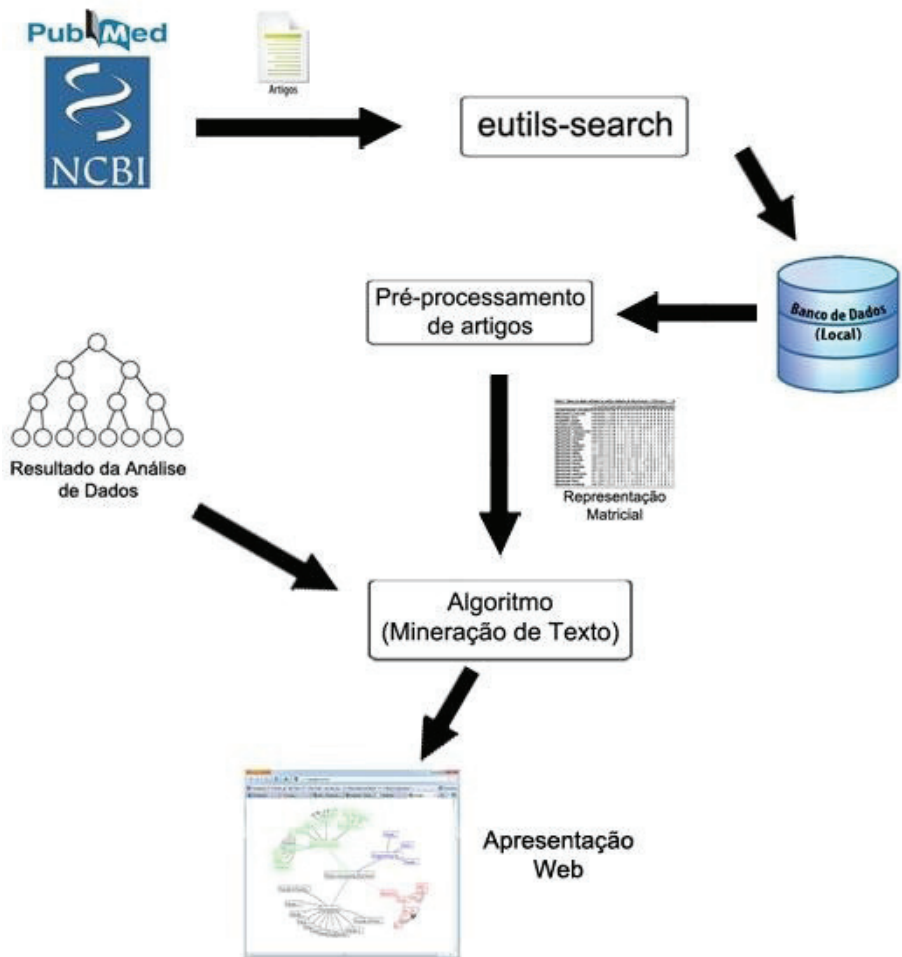
O projeto Prosgen (código SEG 03.09.01.0.25.00.00) se propõe a utilizar informações oriundas do banco de dados de artigos científicos Pubmed (PUBMED, 2011) e técnicas de mineração de textos para apoiar a interpretação biológica de genes candidatos. Seu objetivo é construir um framework de mineração de dados que suporte os processos de priorização e prospecção de genes candidatos associados a características de interesse econômico para a agricultura, para posterior introdução em programas de melhoramento genético. Neste trabalho, apresenta-se o estágio atual de desenvolvimento deste framework, com foco na interpretação de dados de expressão gênica por microarranjos.

Para construção do framework, diferentes programas e bibliotecas de software estão sendo desenvolvidos e/ou integrados (Figura 1). O programa eutils-search foi implementado com a finalidade de fazer o download do site do Pubmed de uma grande quantidade de artigos, relacionados a um organismo específico ou não, e armazená-la em um banco de dados local. Para o seu funcionamento, são necessários 4 arquivos que podem ser obtidos no site do NCBI (NCBI, 2011): (i) o arquivo gene2pubmed, que relaciona os genes de um organismo com um conjunto de publicações do Pubmed ; (ii) o arquivo gene\_info, que contém anotações básicas sobre genes como seu ID, descrição e símbolo; (iii) o arquivo generifs\_basic, que contém informação sumarizada, fornecida por usuários, sobre a função de genes; e (iv) o

---

<sup>1</sup> Universidade Estadual de Campinas, [rafaelst@cnptia.embrapa.br](mailto:rafaelst@cnptia.embrapa.br),  
[rafakashiwara@gmail.com](mailto:rafakashiwara@gmail.com)

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



**Figura 1.** Processo de análise de dados utilizando mineração de textos.

arquivo nodes.dmp, que especifica as relações taxonômicas entre os diversos organismos. Para utilizar o eutils-search, o usuário informa o organismo a que os artigos estão relacionados. O programa, então, verifica quais são os genes relacionados, identifica os correspondentes artigos e faz seu download para a base de dados local PostgreSQL (POSTGRESQL, 2011). Esse procedimento é relativamente demorado, devido às restrições impostas pelo PubMed para acesso ao seu banco de dados.

A partir do banco de dados local, gera-se um conjunto de documentos em formato XML que é submetido a uma série de procedimentos de pré-processamento. Para realizar esse processo foi construída uma biblioteca java, baseada no projeto Apache Lucene (LUCENE, 2011), compreendendo as seguintes funcionalidades: (i) parseamento dos arquivos; (ii) construção de *stems*; (iii) remoção de *stopwords*; (iv) construção e seleção de n-gramas e (v) contagem de frequências de ocorrência. Ao final da fase de pré-processamento, obtém-se uma representação vetorial dos documentos XML, por meio de uma matriz Termo x Documento.

A partir deste ponto, diferentes processamentos são possíveis. O trabalho atualmente em desenvolvimento considera o problema de análise de cluster de dados de expressão gênica por microarranjos. Considera-se uma hierarquia de genes, onde um ou mais artigos pré-processados podem estar associados a cada gene. Está sendo desenvolvido um algoritmo para rotular os ramos da hierarquia com base no conteúdo da matriz termo x documento, isto é, encontrar descritores adequados a cada grupo.

A saída desse processo é um arquivo XML que representa a hierarquia de genes com seus ramos rotulados com os termos estatisticamente associados ao conjunto de genes associados. Para visualização desse resultado, está sendo desenvolvida uma interface web, baseada em um *applet* de árvore hiperbólica (Treebolic) (BOU, 2011).

A utilização eficiente da literatura científica disponível é essencial na pesquisa biológica, tanto na fase de planejamento experimental quanto na interpretação dos resultados obtidos. Em particular, na análise de dados de bioinformática, em que muitas entidades (ex: genes) são analisadas simultaneamente, a agregação automática da literatura ao processo, com uma breve sumarização desta (automaticamente obtida), pretende auxiliar qualitativamente a identificação das funções biológicas compartilhadas por grupos de genes e a comparação entre genes, baseado em sua descrição funcional.

## Referências

BOU, B. Treebolic 2. 2011. <<http://treebolic.sourceforge.net/en/index.html>>. Acesso em: 1 out. 2011.

LUCENE. 2011. Disponível em: <<http://lucene.apache.org/>>. Acesso em: 1 out. 2011.

NCBI. National Center for Biotechnology Information. 2011. Disponível em: <<http://www.ncbi.nlm.nih.gov/>>. Acessado em: 1 out. 2011.

POSTGRESQL. Disponível em: <<http://www.postgresql.org/>>. Acesso em: 1 out. 2011.

PUBMED. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed>>. Acesso em: 1 out. 2011.

# Uma ferramenta para cadastro e mapeamento dos clientes do website Catálogo de Produtos e Serviços da Embrapa

Anderson Carlos Ferreira da Silva<sup>1</sup>  
Leandro Henrique Mendonça de Oliveira<sup>2</sup>

O Catálogo de Produtos e Serviços (CP&S) consiste em um website que reúne informações sobre produtos e serviços da Embrapa. Fundamental para o processo de transferência de tecnologia, esse website atua principalmente na divulgação dos produtos e serviços da empresa e no relacionamento com seu público-alvo (SANTOS, et al., 2007). Disponível desde o final de 2006, o website mantinha somente os registros de acessos de suas páginas, não conhecendo portanto o perfil dos seus clientes ou consumidores do seu conteúdo. Nesse contexto, este trabalho apresenta o desenvolvimento de uma ferramenta de cadastro e mapeamento dos clientes desse website com o intuito de criar e manter um banco de dados organizado com o registro e a identificação dos perfis dos seus visitantes. Desenvolvida no padrão da Web 2.0 e sob as normas da World Wide Web Consortium (W3C) a implementação desse cadastro envolveu a utilização de uma grande variedade de ferramentas e tecnologias, pois sua disponibilização objetivou evitar a abertura de novas janelas ou abas de navegação, justamente para oferecer uma visualização mais dinâmica e interativa do conteúdo, características comuns à natureza e aos propósitos da Web 2.0 (W3C, 2008). Assim, foi desenvolvido um aplicativo dinâmico, de modo que o usuário não necessitasse sair do contexto de navegação para realizar seu cadastro. Escrita em HTML5 (W3C e W3Schools, 2011), CSS3 (W3SCHOOLS, 2011) e o framework jQuery (SOFTWARE FREEDOM CONSERVANCY, 2010) para visualização do cliente e PHP5 (PHP5, 2011)

---

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

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

e PostGreSQL (POSTGRESQL, 2011) no lado servidor, essa ferramenta tem se mostrado muito flexível para adaptação e utilização em websites de informação tecnológica. Como os resultados de sua utilização, espera-se estreitar o relacionamento com os clientes do CP&S, ampliando a interação entre eles e a Embrapa, já que essa ferramenta converge nas duas direções. A primeira está sob o ponto de vista do usuário, pois este poderá se apresentar para a Embrapa, conhecer mais sobre a empresa e suas ações por meio dos informativos recebidos, além de poderem informar seus principais interesses, para melhoria do próprio website e do conteúdo disponibilizado. A segunda está sob o ponto de vista da Embrapa, já que, com a base de dados dos clientes, será possível realizar uma prospecção de novos conteúdos e demandas baseada nos interesses dos usuários registrados. Como trabalhos futuros, espera-se implementar ainda uma ferramenta de consulta e análise dos dados armazenados, bem como realizar a geração de relatórios e estatísticas do cadastro que subsidiem a tomada de decisões gerenciais.

## Referências

PHP5. Disponível em: <<http://www.php.net/>> Acesso em: 18 out. 2011.

POSTGRESQL. Disponível em: <<http://www.postgresql.org/>> Acesso em: 18 out. 2011.

SANTOS, A. D.; MATSUURA, F. A. U.; HIGA, M.; BOLSON, E. A.; OLIVEIRA, L. H. M.; PEREIRA, A. de T. M.; QUEIROS, L. R. Catálogo de produtos e serviços da Embrapa: ferramenta virtual facilita acesso aos produtos desenvolvidos. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 6., 2007, São Pedro, SP. **Anais...** Campinas: Embrapa Informática Agropecuária, 2007. p. 32-33.

SOFTWARE FREEDOM CONSERVANCY. jQuery. Disponível em: <<http://jquery.com/>>. Acesso em: 18 out. 2011.

W3C. **Web wontent accessibility guidelines:** (WCAG) 2.0. Disponível em: <<http://www.w3.org/TR/WCAG/>> Acesso em: 17 out. 2011.

\_\_\_\_\_. **HTML5.** Disponível em: <<http://www.w3.org/TR/html5/>> Acesso em 17 out. 2011.

W3SCHOOLS. **HTML5 tutorial.** Disponível em: <<http://www.w3schools.com/html5/>> Acesso em: 17 out. 2011.

\_\_\_\_\_. **CSS3 tutorial.** Disponível em: <<http://www.w3schools.com/css3/>> Acesso em: 17 out. 2011.

# A Terminologia aplicada ao projeto INTAGRO nas perspectivas da organização e difusão da informação e da representação do conhecimento

Carolina Cardoso de Oliveira<sup>1</sup>

Ivo Pierozzi Jr.<sup>2</sup>

Leandro Henrique Mendonça de Oliveira<sup>2</sup>

Dayse Simon L. Kamikawachi<sup>1</sup>

No contexto da agropecuária brasileira, o projeto Intagro (Intensificação Agropecuária em Pólos de Produção de Soja e Cana de Açúcar: Territorialidade, Sustentabilidade e Competitividade), desenvolvido pela Embrapa Informática Agropecuária, trata de questões referentes ao processo de intensificação agropecuária e suas relações com territorialidade, sustentabilidade e competitividade. Para esse fim, o projeto inclui estudos de diferentes abordagens que englobam aspectos agroambientais e socioeconômicos, bem como aqueles relacionados à disseminação e organização da informação, e à representação do conhecimento desse domínio. Nesse último aspecto, vêm sendo desenvolvidas atividades de cunho terminológico para subsidiar o processo de integração e apropriação dos conhecimentos gerados, variando desde a extração semiautomática até a definição do repertório de termos. Este trabalho apresenta a tarefa de descrição terminológica do conjunto dos termos, coletados no escopo desse projeto para servir de recurso terminológico complementar em diferentes iniciativas do setor agropecuário brasileiro.

Este trabalho vem sendo realizado, sobretudo no e-Termos (<http://www.etermos.cnptia.embrapa.br/index.php>), um ambiente colaborativo que auxilia o trabalho terminológico. Assumindo a orientação de uma teoria

---

<sup>1</sup> Universidade Federal de São Carlos, [carolinaco@cnptia.embrapa.br](mailto:carolinaco@cnptia.embrapa.br), [daysesimon@gmail.com](mailto:daysesimon@gmail.com)

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



descritiva de base linguística, o e-Termos organiza o trabalho terminológico em 6 etapas: iniciando com a compilação do corpus; seguida pela extração automática de termos; criação do mapa conceitual; gerenciamento da base de dados terminológicos; e, finalizando com a edição do verbete e publicação do produto terminológico.

A partir do corpus compilado e de uma taxonomia contendo 639 entradas (entre termos e conceitos) este trabalho se aloja primordialmente na quinta etapa, ou seja, no preenchimento de Bases Definicionais e Fichas Terminológicas. Para a concretização dessa tarefa, foram pesquisados contextos que apresentem o termo em ocorrências reais e que possibilitem a sua compreensão, auxiliando, posteriormente, na redação da definição terminológica. Nesse processo definicional, são consultados o corpus, websites específicos da área e trabalhos de pesquisa sobre o assunto que apliquem o termo em cenários técnicos e científicos.

Com esse banco de dados terminológicos, é possível preencher a ficha terminológica do termo que contém os seguintes campos: Termo, Código do Termo, Morfologia, Definição, Variantes, Equivalências, Glosa, Informações Enciclopédicas, Termos Relacionados, Data de Criação, Data de Revisão, Número de Ocorrências, Responsável e Validador. A definição terminológica é a fase mais morosa e complexa do processo, pois além de exigir do terminólogo um domínio sobre múltiplos conhecimentos e habilidades, essa tarefa também segue critérios rigorosos como: adequação do texto definitório conforme o público-alvo e a finalidade da obra terminográfica em questão; os traços semânticos estabelecidos; os contextos inseridos na Base Definicional e as convenções quanto à estrutura textual.

Os principais resultados desse trabalho se referem à sedimentação do conhecimento sobre a Terminologia de viés descritivo-linguístico, no contexto do projeto Intagro, ao ambiente e-Termos e ao domínio do conhecimento em questão, por meio da pesquisa de termos em diversas fontes. Além disso, foram produzidas 49 Fichas Terminológicas e aproximadamente 270 fragmentos textuais foram armazenados na Base Definicional.

Foram verificadas nessas 49 definições certas regularidades na presença de fatores como: origem, objetivo, método, causa, consequência e exemplos. Essas regularidades permitem a caracterização da área no que diz respeito aos seus traços semânticos recorrentes nos contextos explicativos da Base Definicional e que possibilitaram a redação das definições. Tais fatores servem como guia para os trabalhos futuros dessa pesquisa, na

continuação da produção das fichas terminológicas seguida pela validação dos especialistas, para culminar na publicação de um glossário sobre a intensificação agropecuária.

## Literatura recomendada

BARROS, L. A. **Curso básico de terminologia**. São Paulo: Editora da Universidade de São Paulo, 2004.

CABRÉ, M. T. **La terminología**: teoría, metodología, aplicaciones. Tradução de Carles Tebé. Barcelona: Editorial Antártida: Empúries, 1993.

KRIEGER, M. da G.; FINATTO, M. J. B. **Introdução à terminologia**. teoria & prática. São Paulo: Contexto, 2004.

OLIVEIRA, L. H. M. **e-Termos**: um ambiente colaborativo web de gestão terminológica. 2009. Tese (Doutorado) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, SP.

PIEROZZI JUNIOR, I.; OLIVEIRA, L. H. M. de; SOUZA, K. X. S. de. Construindo ontologias de domínio: o (re)conhecimento da intensificação agropecuária no Brasil. In: SEMINÁRIO DE PESQUISA EM ONTOLOGIA NO BRASIL, 3., 2010, Florianópolis. Glossários, taxonomias e tesouros enriquecendo as ontologias: **anais...** Florianópolis: UFSC, 2010. p. 100-108. ONTOBRASIL 2010.

# Uso de ferramenta para construção de questionários em plataforma web: uma experiência no projeto iAgência

Matheus Silva Miranda<sup>1</sup>  
Adriana Delfino dos Santos<sup>2</sup>

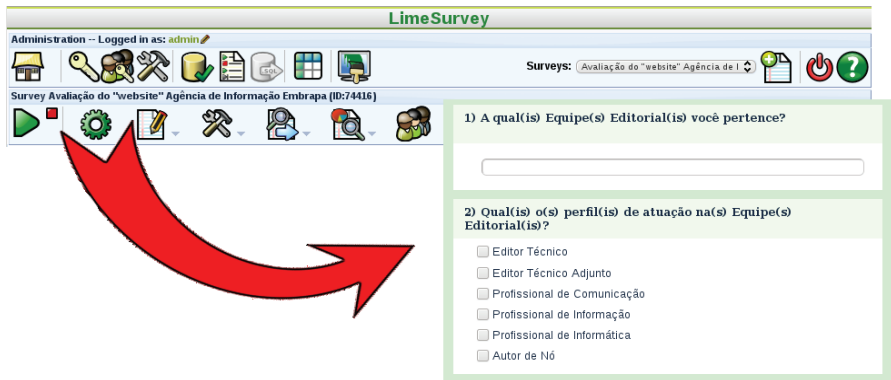
O projeto intitulado “*Websites* de informação tecnológica: ampliação do relacionamento com o cliente” (iAgência), em desenvolvimento pela Embrapa Informática Agropecuária, visa incorporar conceitos e ferramentas aos *websites* Agência de Informação Embrapa ([www.embrapa.br/agencia](http://www.embrapa.br/agencia)) e Catálogo de Produtos e Serviços da Embrapa ([www.embrapa.br/catalogo](http://www.embrapa.br/catalogo)) para conferir a estes maior acessibilidade, navegabilidade e interatividade (SANTOS, 2009). No âmbito do Plano de Ação “Desenvolvimento de um novo modelo de arquitetura da informação e de design para *website* de conteúdo de informação tecnológica”, está definida uma atividade para avaliação daqueles *websites* pelos seus gestores de conteúdo (distribuídos geograficamente no território nacional em quarenta unidades de pesquisa descentralizadas) e este trabalho deu apoio tecnológico. O objetivo deste trabalho é compartilhar a experiência na identificação e no uso de uma ferramenta para construção e aplicação de questionários em plataforma web. Para tal, elaborou-se as questões junto à equipe do projeto. Na sequência, definiu-se os seguintes atributos para a ferramenta: (a) software livre, (b) código aberto (devido à necessidade de customização), (c) armazenamento dos dados em servidor interno à Empresa, (d) interface de administração da ferramenta amigável, (e) formato das questões compatível com o definido no questionário. O levantamento das principais ferramentas foi feito por meio de busca na internet. Uma análise foi realizada confrontando os atributos com a especificação das ferramentas identificadas e a ferramenta mais adequada foi selecionada. Então, foi possível realizar a construção

---

<sup>1</sup> PUC-Campinas, Sistemas de Informação, [matheussm@cnptia.embrapa.br](mailto:matheussm@cnptia.embrapa.br)

<sup>1</sup> Embrapa Informática Agropecuária, [adriana@cnptia.embrapa.br](mailto:adriana@cnptia.embrapa.br)

do questionário nessa ferramenta selecionada, o mesmo foi validado pela equipe do projeto e, posteriormente, aplicado aos gestores de conteúdo. As ferramentas analisadas foram *Google Spreadsheets* ferramenta do *Google Docs* (GOOGLE DOCS, 2011), *SurveyMonkey* (SURVEYMONKEY, 2011), *Zoomerang* (MARKETTOOLS, 2011) e *LimeSurvey* (SCHMITZ, 2011). A *Google Spreadsheets* está disponível no ambiente do Google Docs e as informações são armazenadas (hospedadas) naquele ambiente, o que não atende ao atributo “c”. A *SurveyMonkey* e *Zoomerang* são ferramentas proprietárias e suas versões gratuitas possuem muitas limitações, portanto não atendem aos atributos “a”, “b” e “e”. A *LimeSurvey* versão 1.91 atende a todos os atributos e, conseqüentemente, foi selecionada para uso no projeto. Procedeu-se à sua instalação em um servidor Apache (THE APACHE SOFTWARE FOUNDATION, 2011) com PHP versão 5 (THE PHP GROUP, 2011) e banco de dados PostgreSQL 8.4 (POSTGRESQL, 2011). Como apoio técnico, o *website* da ferramenta provê o manual online da *LimeSurvey* (cobrindo a maioria dos casos de dúvidas desde a instalação até a customização de questões) e também o fórum de discussão. Com a ferramenta instalada, o questionário foi construído e validado. A ferramenta oferece possibilidade de criação de senhas de acesso, que são enviadas automaticamente nos e-mails. Para tal, foram configuradas as opções de acesso a partir da lista de e-mails dos gestores de conteúdo dos *websites* avaliados e o questionário foi aplicado por um período de três semanas. As respostas podem ser analisadas graficamente no próprio navegador web ou então exportadas para diversos formatos de arquivos, como “xls”, “pdf”, “doc”, entre outros. A ferramenta *LimeSurvey* proporciona a construção e aplicação rápida e eficiente de questionários em plataforma web, por meio de uma interface de administração fácil e intuitiva (Figura 1), proporcionando maior segurança às informações e aos acessos dos participantes. Também é um software livre e de código aberto (o que possibilitou, rapidamente, criar um novo tipo de resposta) além de atender a todos os atributos requeridos pelo projeto. Devido a essa experiência de sucesso, a *LimeSurvey* está disponível no servidor interno da Embrapa Informática Agropecuária, para que qualquer outro setor possa criar e aplicar seus próprios questionários.



**Figura 1.** Interface de administração da LimeSurvey e um exemplo do questionário gerado.

## Referências

GOOGLE Docs. 2011. Disponível em: <<http://www.docs.google.com>>. Acesso em: 21.out.2011.

MARKETTOOLS. Online survey software: create online surveys: zoomerang. 2011. Disponível em : <<http://www.zoomerang.com/>>. Acesso em: 21.out.2011.

POSTGRESQL PostgreSQL: the world's most advanced open source database. 2007. Disponível em: <<http://www.postgresql.org/>>. Acesso em: 21 out. 2011.

SANTOS, A. D. dos; QUEIROS, L. R.; OLIVEIRA, L. H. M. de; SOUZA, M. I. F.; MATSUURA, F. C. A. U.; MOURA, M. F.; FERREIRA, V. R.; HANASHIRO, M. M.; TELLES, M. A.; NASCIMENTO NETO, F.; LEITÃO, A. P. da S. D. M.; MATTOS, C. B.; BATISTA, J. A. O.; LEITE, F. C. L.; CORDEIRO, R.G.; PEREIRA, R. M.; BERTIN, P. R. B.; HENRIQUES, B. F. R.; BERNARDES, R. M.; BAMBINI, M. D. Websites de informação tecnológica: ampliação do relacionamento com o cliente. Brasília, DF: Embrapa, 2009. Embrapa Macroprograma 5 – Desenvolvimento Institucional. Projeto em Andamento.

SCHMITZ, C. LimeSurvey - the free & open source survey software tool. 2011. Disponível em : <<http://www.limesurvey.org/>>. Acesso em: 21 out. 2011.

SURVEYMONKEY SurveyMonkey: free online survey software & questionnaire tool. 2011. Disponível em : <<http://www.surveymonkey.com/>>. Acesso em: 21 out. 2011.

THE APACHE SOFTWARE FOUNDATION. The Apache Software Foundation. 2011. Disponível em : <<http://apache.org>>. Acesso em: 21.out.2011.

THE PHP GROUP. PHP: hypertext preprocessor. 2011. Disponível em : <<http://www.php.net/>>. Acesso em: 21 out.2011.

## Desenvolvimento de software para a detecção de grupos de genes homólogos sob evidência de seleção positiva

Jorge Augusto Hongo<sup>1</sup>  
Francisco Pereira Lobo<sup>2</sup>

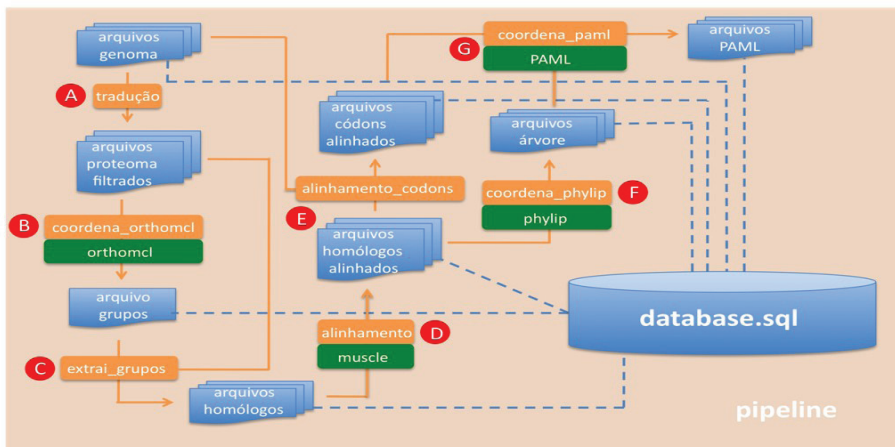
Uma fração considerável dos genes encontrados em projetos genoma não possui função biológica conhecida (REICHARDT, 2007). Esse vasto universo de genes desconhecidos constitui um campo fértil para a busca de genes interessantes, visando aplicações de biotecnologia. No caso de espécies de interesse agropecuário, esses genes desconhecidos constituem um vasto campo de buscas para localização de genes de interesse para ganhos de produção (CORBI et al., 2011; FAN et al., 2009; GU et al., 2009). Nesse cenário, é de extrema importância que novos métodos computacionais sejam desenvolvidos para a detecção de genes desconhecidos que apresentem potencial de contribuir para traços fenotípicos interessantes em espécies animais e vegetais estudadas pela Empresa Brasileira de Pesquisa Agropecuária (Embrapa). Uma estratégia ainda não explorada pela Embrapa para a detecção de genes potencialmente interessantes é a busca por grupos de genes homólogos (grupos de genes encontrados em espécies diferentes) sob evidência de seleção positiva (AGUILETA et al., 2009). É fato amplamente conhecido que a vasta maioria dos genes homólogos é conservada. Isso ocorre porque usualmente mutações não-sinônimas diminuem a eficiência funcional da proteína, o que diminui a aptidão evolutiva do indivíduo e impede a fixação do novo alelo quando comparado ao alelo ancestral (HARTWELL, 2011). Entretanto, alguns poucos grupos de genes homólogos evoluem apresentando uma forte pressão seletiva para a variação, ao invés da conservação (YANG, 2007). Uma vez que as espécies estudadas pela Embrapa têm sido alvo de seleção artifi-

---

<sup>1</sup> UNICAMP, Instituto de Computação, [jorgeahongo@gmail.com](mailto:jorgeahongo@gmail.com)

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

cial para alguns poucos fenótipos de interesse, visando à produtividade, é razoável supor que os genes sob evidência de seleção positiva nessas espécies serão, possivelmente, associados a fenótipos de produtividade (CORBI et al., 2011; FAN et al., 2009; GU et al., 2009). Nesse contexto, a busca por genes, sob evidência de seleção positiva em genomas de espécies de interesse da Embrapa, constitui uma importante ferramenta para indicar possíveis genes associados a um maior ganho de produção nessas espécies. Assim, o presente trabalho descreve o desenvolvimento de um software para a busca por grupos de genes homólogos sob evidência de seleção positiva. Nesse trabalho, desenvolveu-se um software em perl para integrar diversos softwares de terceiros capazes de realizar as tarefas individuais para a detecção de genes sob evidência de seleção positiva, conforme resumido na Figura 1. Nessa figura as caixas em azul indicam os arquivos produzidos sequencialmente durante a análise; as caixas em laranja indicam os módulos perl desenvolvidos (os quais controlam a execução coordenada dos softwares de terceiros utilizados); as setas em laranja indicam o fluxo sequencial da análise; as caixas em verde indicam softwares desenvolvidos por terceiros utilizados neste trabalho. O cilindro em azul indica o banco de dados desenvolvido para organizar e armazenar os dados produzidos. As letras sequenciais em vermelho indicam as etapas necessárias para que uma análise seja feita. Todas as etapas da análise



**Figura 1.** Arquitetura do sistema.

foram desenvolvidas de maneira modular, de modo a permitir o futuro uso de programas alternativos em cada uma das etapas. Para o desenvolvimento do software para a detecção de seleção positiva em grupos de genes homólogos, empregou-se os softwares OrthoMCL para a detecção dos grupos de homólogos (CHEN et al., 2006), MUSCLE para o alinhamento dos homólogos (EDGAR, 2004; RETIEF, 2000) phylip para a construção de árvores filogenéticas (RETIEF, 2000) e PAML para a localização de seleção positiva (YANG, 2007). Foram desenvolvidos procedimentos para adequar cada um dos arquivos de saída de cada um dos softwares listados acima para o próximo software da *pipeline*, conforme ilustrado na Figura 1 (caixas laranja). O programa final produzido possui aproximadamente 1000 linhas de código e utiliza diversos módulos sofisticados de bioinformática previamente desenvolvidos para perl (bioperl, [http://www.bioperl.org/wiki/Main\\_Page](http://www.bioperl.org/wiki/Main_Page)). O usuário pode controlar o comportamento de todos os softwares de terceiros utilizados através de parâmetros globais definidos no início da execução da *pipeline*. Diversos conjuntos de dados na literatura científica são compostos por buscas manuais, por genes sob evidência de seleção positiva, e serão utilizados como provas de conceito do software desenvolvido (CORBI et al., 2011; ESTEBAN; HUTCHINSON, 2011; FAN et al., 2009).

Os autores agradecem à Embrapa por fornecer a bolsa, para que o estagiário desenvolvesse o trabalho em questão.

## Referências

- AGUILETA, G.; REFREGIER, G.; YOCKTENG, R.; FOURNIER, E.; GIRAUD, T. "Rapidly evolving genes in pathogens: methods for detecting positive selection and examples among fungi, bacteria, viruses and protists." **Infection, Genetics and Evolution**, v. 9, n. 4, p. 656-670, 2009.
- CHEN, F.; MACKEY, A. J.; STOECKERT JUNIOR, C. J.; ROOS, D. S. "OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups." **Nucleic Acids Research**, v. 34 (Database issue), p. D363-368, Jan. 2006.
- CORBI, J.; DEBIEU, M.; ROUSSELET, A.; MONTALENT, P.; LE GUILLOUX, M.; MANICACCI, D.; TENAILLON, M. I. "Contrasted patterns of selection since maize domestication on duplicated genes encoding a starch pathway enzyme." **Theoretical and Applied Genetics**, v. 122, n. 4, p. 705-722, 2011.



EDGAR, R. C. "MUSCLE: a multiple sequence alignment method with reduced time and space complexity." **BMC Bioinformatics**, v. 5, p. 113, 2004.

ESTEBAN, D. J.; HUTCHINSON, A. P. "Genes in the terminal regions of orthopoxvirus genomes experience adaptive molecular evolution." **BMC Genomics**, v. 12, p. 26, 2011.

FAN, L.; BAO, J.; WANG, Y.; YAO, J.; GUI, Y.; HU, W.; J.; ZHU, J.; ZENG, M.; LI, Y.; XU, Y. "Post-domestication selection in the maize starch pathway." **PLoS ONE**, v. 4, n.10, p. e7612, 2009.

GU, J.; ORR, N.; PARK, S. D.; KATZ, L. M.; SULIMOVA, G.; MACHUGH, D. E.; HILL, E. W. "A genome scan for positive selection in thoroughbred horses." **PLoS ONE**, v. 4, n. 6, p. e5767, 2009.

HARTWELL, L. **Genetics: from genes to genomes**. 4th ed. New York: McGraw-Hill, 2011. v. 1.

REICHARDT, J. K. "Quo vadis, genoma? A call to pipettes for biochemists." **Trends Biochem SciENCE**, v. 32, n.12, p. 529-530, 2007.

RETIEF, J. D. "Phylogenetic analysis using PHYLIP." **Methods in Molecular Biology**, n. 132: 243-258, 2000.

YANG, Z. "PAML 4: phylogenetic analysis by maximum likelihood." **Molecular Biology and Evolution**, v. 24, n. 8, p.1586-1591, May, 2007.

## Ferramenta para gerenciamento de temas em Webgis baseados no i3geo

Renato Luiz Franceschinelli<sup>1</sup>  
Eduardo Antonio Speranza<sup>2</sup>  
João dos Santos Vila da Silva<sup>2</sup>

Em um WebGIS, temas são filtros que podem ser adicionados em mapas, tornando-os mais dinâmicos e específicos, de acordo com as informações que o usuário necessita. Exemplos de temas são: tipos de solos, hidrografia, divisão política, etc. No caso do software i3geo (I3GEO, 2011), o tema é representado por meio de um arquivo do tipo mapfile, que pode ser construído a partir de informações de arquivos georreferenciados vetoriais(shapefile) ou matriciais(imagens de satélite).

A ferramenta de gerenciamento de temas em questão será utilizada no SISLA - Sistema Interativo de Suporte ao Licenciamento Ambiental (SPERANZA, et al., 2011), que utiliza como base o software i3geo. Em um primeiro momento, trabalhará apenas sobre arquivos vetoriais e tem, como função, realizar três operações: inserir novos temas, editar e excluir temas já existentes.

A inserção requer do usuário um nome para o tema e três arquivos(.shp, .shx, .dbf) que juntos formam um shapefile. A edição conta com duas possibilidades: em sua primeira opção, o usuário poderá realizar alterações apenas no mapfile, ou seja, apenas na forma como os dados serão exibidos, não sendo alterada a fonte de dados; no segundo caso, deverá realizar o upload de um novo shapefile no banco de dados. A exclusão consiste em apagar as informações contidas no banco de dados e excluir o arquivo mapfile referente ao tema.

O objetivo desse trabalho foi criar uma ferramenta visual para administração de temas para WebGIS baseados no i3Geo, possibilitando ao usuário esco-

---

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

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

lher classificações e legendas apropriadas para cada tema.

O desenvolvimento da ferramenta utilizou apenas tecnologias livres, sua interface foi desenvolvida em HTML e PHP para torná-la mais dinâmica e realizar a interação com o banco de dados. O SGBD utilizado foi o PostgreSQL que possui a extensão geográfica PostGIS, uma vez que as informações extraídas de shapefiles contém campos geométricos.

A validação dos arquivos enviados pelo usuário(shapefile) é feita tanto client-side, utilizando Javascript, como server-side, com PHP, evitando dados inconsistentes no banco de dados.

Para manipulação de mapfile, como criação, edição e leitura, foi utilizado o servidor de mapas Mapserver (MAPSERVER, 2011) pelo módulo PHPMapscript, que oferece métodos que facilitam todas as operações necessárias em um mapfile.

Toda vez que um tema for inserido, será criada uma tabela no banco de dados exclusivamente para guardar informações e os dados geométricos encontrados no shapefile enviado. A partir disso é possível montar o mapfile do tema.

Na edição de temas, a primeira ação a se realizar é verificar se o usuário inseriu ou não um novo shapefile. Caso contrário, a tela irá exibir apenas opções para alterar o mapfile, como nome e cor de classes, gerar novas classificações, entre outras. Posteriormente, o novo shapefile é comparado com as informações contidas no banco de dados, e se forem encontradas diferenças, o primeiro passo é criar um backup da tabela com as informações antigas, e os novos dados são armazenados no banco.

No processo de exclusão também é criado um backup contendo data e horário da exclusão do tema.

A Figura 1, ilustra um exemplo de uso da ferramenta.

Ainda que em fase de testes, a ferramenta já apresenta bons resultados, mostrando grande eficiência, facilitando e tornando mais ágil o gerenciamento de temas. A ideia inicial de possibilitar aos próprios técnicos do IMASUL a edição de temas, sem necessitar da intervenção de um profissional de computação, deverá ser atendida com essa ferramenta.

O desenvolvimento dessa ferramenta teve apoio financeiro do projeto GeoMS, e os autores agradecem à EMBRAPA Informática Agropecuária pelo auxílio de infraestrutura.

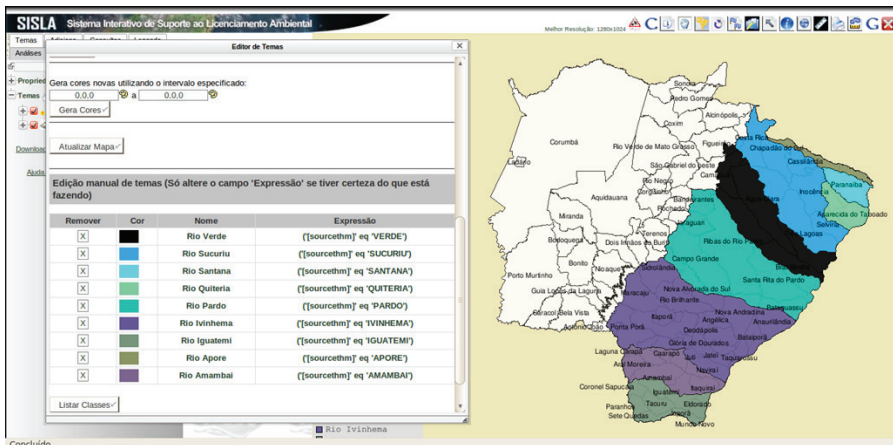


Figura 1. Tela para edição dos temas, possibilitando a escolha de classes e cores.

## Referências

I3GEO. **I3geo Documentação**. 2011. Disponível em: <<http://mapas.mma.gov.br/i3geo/documentacao/>> Acesso em: 17 out. 2011.

MAPSERVER. **MapServer Documentation**. 2011. Disponível em: <<http://mapserver.org/documentation.html>> Acesso em: 17 out. 2011.

SPERANZA, E. A.; ESQUERDO, J. C. D. M.; SILVA, J. dos S. V.; ANTUNES, J. F.; LOURENÇO, F. V.; CEZAR, V. M. SISLA - Interactive system for environmental licensing support. **Geografia**, Rio Claro, v. 6, p. 57-72, jun. 2011.

## WikiPragas: plataforma Wiki para fichas de pragas

Gabrielle Consolini Salazar Alarcon<sup>1</sup>  
Carlos Alberto Alves Meira<sup>2</sup>  
Anderson Carlos Bueno dos Santos<sup>3</sup>  
Bruno Franciscan Mazzotti<sup>4</sup>  
Cristiano Borges Cardoso<sup>5</sup>  
Isaque Vacari<sup>2</sup>  
Marcos Cezar Visoli<sup>2</sup>  
Tiago Augusto Pimenta<sup>6</sup>

O agronegócio brasileiro é um dos principais setores da economia nacional. Eficiente e competitivo, responde por cerca de  $\frac{1}{3}$  do produto interno bruto do Brasil e por 42% das nossas exportações. Entretanto, o crescimento da participação brasileira no comércio internacional implica em uma grande movimentação de produtos agropecuários e favorece a dispersão e introdução de pragas agrícolas. Estas, quando introduzidas em áreas produtivas do país, podem causar forte impacto econômico e social.

Com o objetivo de mitigar o risco de introdução de pragas exóticas, a Organização das Nações Unidas para Agricultura e Alimentação (FAO) e a Organização Mundial de Comércio (OMC) desenvolveram normas e obrigações para disciplinar o comércio de produtos vegetais. O Brasil é um dos países signatários dessas normas. A importação de vegetais, suas partes, produtos ou subprodutos, com potencial de abrigar pragas, deve ser realizada baseando-se nos resultados de Análises de Risco de Pragas

---

<sup>1</sup> Pontifícia Universidade Católica de Campinas-SP, [gabi\\_alarcon@hotmail.com](mailto:gabi_alarcon@hotmail.com)

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

<sup>3</sup> Bolsista CNPq, [anderson.carlosb@gmail.com](mailto:anderson.carlosb@gmail.com)

<sup>4</sup> Bolsista CNPq, [bruno.mazzotti@gmail.com](mailto:bruno.mazzotti@gmail.com)

<sup>5</sup> Bolsista CNPq, [cristianoborgescardoso@gmail.com](mailto:cristianoborgescardoso@gmail.com)

<sup>6</sup> Bolsista CNPq, [tiagoapimenta@gmail.com](mailto:tiagoapimenta@gmail.com)

(ARPs), instrumentos regulatórios do comércio que visam proteger a sanidade vegetal (CNPq, 2009).

A atuação do Brasil no negócio de *commodities* agrícolas fez aumentar, consideravelmente, a demanda de ARPs sob a Organização Nacional de Proteção Fitossanitária (ONPF) brasileira, o DSV (Departamento de Sanidade Vegetal). Identificou-se, prontamente, a necessidade de um sistema de informação que disponibilizasse, de forma facilitada, dados sobre pragas com o propósito de agilizar a elaboração de ARPs.

O Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), por meio de recursos do Fundo Setorial do Agronegócio, tornou público o Edital de Análise de Riscos de pragas (CNPq, 2009), alinhado aos objetivos e metas do Ministério da Agricultura, Pecuária e Abastecimento (Mapa) e do DSV. Dividiu-se o edital em duas frentes: a primeira selecionou projetos para levantamentos em nível mundial, das pragas associadas às principais culturas geradoras de divisas para o agronegócio brasileiro, bem como os respectivos estudos de ARP; e, na segunda frente, foi selecionado um projeto para o desenvolvimento e implantação de um banco de dados com as informações coletadas. Coube à Embrapa Informática Agropecuária o desenvolvimento desse sistema de informação.

Dando início ao projeto, foi possível identificar que o sistema precisaria de duas principais funcionalidades: catalogação das pragas e elaboração de fichas de pragas. As fichas de pragas são documentos com conteúdo descritivo e elaborado, necessitando da composição de texto com tabelas, figuras e mapas. A estrutura segue um modelo, mas o preenchimento é livre; a edição caracteriza-se por ser incremental e colaborativa; a informação está mais sujeita a atualizações; as consultas ao conteúdo são indeterminadas, geralmente por alguma palavra-chave. Dadas essas características, decidiu-se implementar o módulo de elaboração de fichas por meio de um sistema wiki, denominado WikiPragas, integrando-o com o módulo web de catalogação das pragas que gerencia um banco de dados relacional (MEIRA et al., 2011).

O WikiPragas ([www.wikipragas.cnptia.embrapa.br](http://www.wikipragas.cnptia.embrapa.br)) foi construído com o MediaWiki (WIKIPEDIA, 2011). Essa ferramenta foi escolhida, pois é um software livre, disponibilizado sem custos, pode ser personalizado e reaproveitado conforme necessidades específicas, possibilita liberdade de escolha de plataformas e sistemas, permite independência de fornecedores de software, liberdade para ler, escrever e reaproveitar o software, e

também porque é uma ferramenta madura e estável. Além disso, possui editor de texto integrado com suporte multimídia, o que permite elaborar o texto de forma contínua, unificada e no formato final que será visualizado pelo usuário; também, é registrado um histórico das alterações, com possibilidade de comparar versões e recuperar versões anteriores da ficha.

A Figura 1 exibe a ficha da praga *Zonocerus variegatus* no WikiPragas. A ficha foi incluída no WikiPragas pela equipe de desenvolvimento do sistema, baseando-se em informações coletadas por outro projeto do edital. No sumário do conteúdo, pode-se notar todos os itens encontrados em uma ficha de praga: 1 – Sinonímias (outros nomes científicos da praga), 2 – Hospedeiros, 3 – Distribuição geográfica (países onde há ocorrência da praga), 4 – Código internacional da praga, 5 – Posição taxonômica (reino, sub reino, filo, sub filo, classe, ordem, família, gênero, espécie), 6 – Nomes comuns, 7- Biologia (caracterização morfológica da praga, formas reprodutivas, condições edafo-climáticas ideais para o desenvolvimento), 8 – Vias de ingresso (parte de plantas, embalagens ou substratos que podem servir como transporte para a praga), 9 – Inspeção e detecção (instruções para dirigir a inspeção e diagnóstico de sintomas), 10 – Impactos da praga (econômicos, ambientais ou sobre a produtividade), 11 – Medidas de controle e mitigação (mitigação do risco da praga) e 12 – Referências.

O WikiPragas adequou-se muito bem às necessidades de organização de informação das fichas de pragas utilizadas pelo DSV no processo de elaboração de ARPs. A ferramenta encontra-se em fase de produção e atualmente disponibiliza 245 fichas de pragas. O uso do wiki para a pes-

The image shows a screenshot of the WikiPragas website. At the top, there is a header with the logo 'WikiPragas' and navigation links like 'wikipeer', 'minha discussão', etc. Below the header, there is a search bar and a list of actions: 'Página', 'Discussão', 'Editar', 'Histórico', 'Eliminar', 'Mover', 'Proteger', 'Vigiar'. The main content area is titled 'Zonocerus variegatus'. On the left, there are two sidebars: 'Navegação' with links like 'Página principal', 'Portal comunitário', and 'Eventos atuais'; and 'Ferramentas' with links like 'Páginas afilientes', 'Alterações relacionadas', and 'Enviar arquivo'. The main content area has a 'Tabela de conteúdo' with a list of 12 items: 1 Sinonímias, 2 Hospedeiros, 3 Distribuição Geográfica, 4 Código EPPO, 5 Posição taxonômica, 6 Nomes comuns, 7 Biologia, 8 Vias de ingresso, 9 Inspeção e Detecção, 10 Impactos da praga, 11 Medidas de Controle e Mitigação, 12 Referências. To the right of the table of contents is a box titled 'Zonocerus variegatus' containing a photograph of the insect on a green leaf, with the text 'Grupo: Inseto' below it.

Figura 1. Ficha do inseto *Zonocerus variegatus* no WikiPragas.

quisa e a edição de fichas agiliza os processos de ARP e aumenta a qualidade das decisões, colaborando com o desenvolvimento do setor agrícola nacional.

## Referências

CNPq. **Editais MCT/CNPq/CT-AGRO No 032/2009 – análise de risco de pragas.** Brasília, DF, 2009. 11 p.

MEIRA, C. A. A.; VISOLI, M. C.; VACARI, I.; SANTOS, A. C. B. dos; MAZZOTTI, B. F.; CARDOSO, C. B.; RODRIGUES, L. de C. F.; PIMENTA, T. A. Banco de dados e wiki integrados para análise de risco de pragas. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 8., 2011, Bento Gonçalves. **Anais...** Florianópolis: UFSC; Pelotas: UFPel, 2011. Não paginado. SBIAgro 2011.

WIKIPEDIA **MediaWiki**. Disponível em: <<http://en.wikipedia.org/wiki/MediaWiki>>. Acesso em: 24 de outubro de 2011.



# Uma ferramenta para cadastro e visualização de metadados em Agricultura de Precisão utilizando GeoNetwork

Pedro Eduardo Bertelli Corrêa<sup>1</sup>

Eduardo Antonio Speranza<sup>2</sup>

Leonardo Ribeiro Queiros<sup>2</sup>

A Rede Agricultura de Precisão (RedeAP) consiste em um sistema que pode ser descrito como uma estratégia de gestão dos dados produzidos nas Unidades Piloto, e é fundamentalmente constituído de três componentes inter-relacionados entre si, sendo eles: Bancos de Dados, Política de Acesso e uso dos Dados, e Aplicação. O componente de Aplicação se trata de um Sistema de Informações Geográficas para a Web baseado no software livre i3Geo com o aplicativo de catalogação de dados geoespaciais GeoNetwork (QUEIROS et al., 2011).

A adição de novos registros no catálogo do GeoNetwork pode ser feita utilizando o próprio editor de metadados ou pela ferramenta de importação dos dados no formato XML. Em ambos os casos, o usuário deve ser, no mínimo, registrado como editor no grupo correspondente. Contudo, o controle de usuários, perfis e níveis de acesso deve ser feito pelo próprio sistema RedeAP, utilizando, assim, o GeoNetwork somente para a visualização dos metadados catalogados.

Baseado nesse contexto, surgiu a necessidade da criação de uma ferramenta que fizesse a catalogação dos metadados com o objetivo de mantê-los armazenados no banco de dados do sistema e também possibilitar sua visualização nos padrões de exibição do GeoNetwork.

O sistema de cadastro em questão foi desenvolvido baseado nos padrões da web e, durante o seu processo de implementação, foram utilizadas uma

---

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

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

variedade de tecnologias, todas OpenSource (Código Livre). A interface foi desenvolvida em XHTML (W3SCHOOLS, 2011a) e seu design desenvolvido em CSS2 (W3SCHOOLS, 2011b) utilizando as técnicas de “tableless” (forma de desenvolvimento web que não utiliza tabelas para disposição de conteúdo na página sugerido pela W3C). Desse modo, a interface foi estruturada sobre a plataforma PHP (PHP5, 2011) e o Servidor de Gerenciamento de Banco de Dados (SGBD) PostgreSQL foi utilizado para armazenamento dos dados.

Foi construído o formulário de inserção dos metadados, utilizando o framework de Javascript, jQuery (SOFTWARE FREEDOM CONSERVANCY, 2010), tanto na sua validação client-side como para as interações em Ajax do navegador com o banco de dados.

Após a submissão do formulário, é realizado um parseamento de suas informações no servidor PHP. Em seguida é montado um arquivo XML, no formato interpretado pelo GeoNetwork, que é transcrito para uma tabela no SGBD destinada a guardar metadados nesse formato.

Caso a transação tenha ocorrido com sucesso, os metadados já podem ser acessados sob a interface do GeoNetwork, permitindo aos usuários visualizá-los de forma mais fácil e dinâmica. Em caso de erro, o metadado não é inserido e o sistema informa a eventual falha ao usuário.

A ferramenta apresentou resultados satisfatórios e se mostrou bem estável apesar de estar em fase de testes. Durante as simulações tanto de inserção de dados, quanto nas tentativas de induzir um erro ou mal funcionamento ao sistema, a ferramenta se comportou como o esperado, tornando o cadastro de metadados mais simples e eficiente.

A ferramenta deverá ser adotada pelo repositório da Rede AP para cadastro de metadados referentes aos dados obtidos nas Unidades Piloto. Para tanto, deverão ser realizadas modificações na estrutura desse repositório, de forma a permitir o acesso a esse formulário logo após o upload e à pré-catalogação dos dados realizados pelo usuário.

## Referências

QUEIROS, L. R.; SPERANZA, E. A.; COARACY, G. M. B.; FILIPPINI ALBA, J. M.; BERNARDI, A. C. de C.; INAMASU, R. Y.; GREGO, C. R.; RABELLO, L. M.. **Gestão de recursos de informação em agricultura de precisão**. Campinas: Embrapa Informática Agropecuária, 2011. 41 p. il. (Embrapa Informática Agropecuária. Documentos, 112).

OPEN SOURCE GEOSPATIAL FOUNDATION. **GeoNetwork opensource**. 2011. Disponível em: < <http://geonetwork-opensource.org/> >. Acesso em: 21 out. 2011.

PHP5. 2011. Disponível em: <<http://www.php.net/>> Acesso em: 21 out. 2011.

POSTGRESQL. 2011. Disponível em: <<http://www.postgresql.org/>> Acesso em: 21 out. 2011.

SOFTWARE FREEDOM CONSERVANCY. **JQuery**. 2011. Disponível em: <<http://jquery.com/>> Acesso em: 21 out. 2011.

W3SCHOOLS. **HTML Tutorial**. 2011a. Disponível em: <<http://www.w3schools.com/html/>> Acesso em: 22 out. 2011.

W3SCHOOLS. **CSS Tutorial**. 2011b. Disponível em: <<http://www.w3schools.com/css/>> Acesso em: 22 out. 2011.

## JQuery ou Dojo, qual biblioteca JavaScript usar?

Juliana Pereira de Santana<sup>1</sup>  
Luciana Alvim Santos Romani<sup>2</sup>

O desenvolvimento de sistemas web apoia-se no uso de bibliotecas JavaScript que já implementam uma série de funções e de componentes utilizados nas interfaces gráficas. Dessa forma, é importante conhecer as bibliotecas disponíveis a fim de escolher a mais adequada a cada projeto. Para se obter um melhor desempenho no desenvolvimento do sistema, deve-se levar em consideração alguns aspectos importantes na escolha da biblioteca: documentação, qualidade e compatibilidade entre os navegadores. Este trabalho apresenta uma análise comparativa das bibliotecas jQuery (JQUERY, 2011) e Dojo (DOJO, 2011) que possuem códigos de funções, classes, manipulação do Document Object Model (DOM), entre outros, o que facilita o desenvolvimento de aplicações.

Durante o desenvolvimento deste trabalho que apoiou importantes decisões de projeto do desenvolvimento da nova versão do Agritempo (EVANGELISTA et al., 2003; COLTRI et al., 2007), foram utilizados tutoriais e informações disponíveis nos *websites* das bibliotecas Dojo e jQuery. A fim de se obter parâmetros para avaliar melhor as duas bibliotecas, foram desenvolvidos protótipos usando diversos componentes de ambas as bibliotecas. Essas interfaces web criadas permitiram observar o desempenho das duas bibliotecas e esclarecer dúvidas que não estavam claras no material bibliográfico consultado.

Diferentes aspectos foram avaliados na comparação das duas bibliotecas. A documentação é um ponto muito importante, pois fornece informações de como usar a ferramenta, juntamente com os tutoriais e os exemplos de pequenas aplicações que mostram o uso de funções e classes. O jQuery pos-

---

<sup>1</sup> COTUCA, Colégio Técnico de Campinas, [julips.hp@gmail.com](mailto:julips.hp@gmail.com)

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

sui uma documentação completa com eventos, códigos para manipulação do DOM, seletores, efeitos de animação, *plugins* e utilitários interessantes.

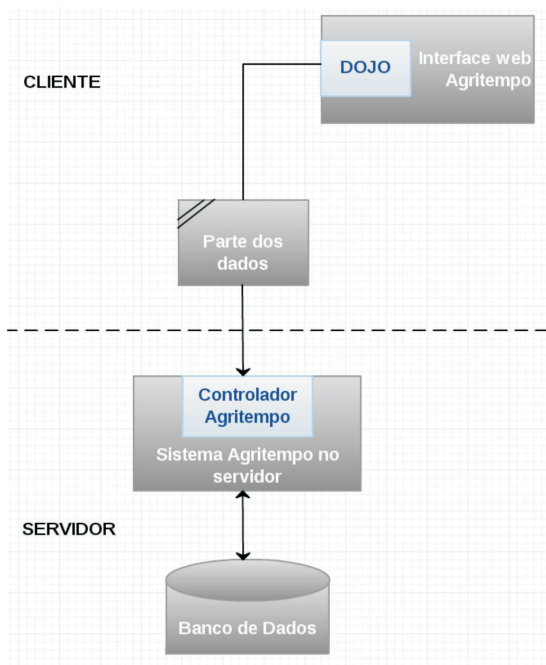
Por outro lado, o Dojo, além de eventos, de códigos para manipulação DOM, de seletores e efeitos de animação, possui uma documentação mais abrangente com as suas extensões Dijit, uma camada acima do Dojo, que fornece *widgets* para criar *websites* e aplicações AJAX com temas prontos. Também apresenta informações sobre o Dojox, área para o desenvolvimento de extensões para o Dojo, um lugar para novas ideias. No entanto, no que se refere ao volume dos tutoriais disponíveis na web, o jQuery possui uma quantidade bem maior se comparado ao Dojo.

Uma das qualidades de um *framework* está na velocidade em que é executado e na sua leveza. Nesse aspecto, embora o jQuery seja bastante rápido, o Dojo é mais eficiente, sendo mais rápido e leve.

As duas bibliotecas foram desenvolvidas para os navegadores mais usados: Internet Explorer, Mozilla Firefox, Google Chrome e, também, o Opera e o Safari. Além desses, o Dojo também funciona em plataformas móveis, como o iPhone, iPod Touch, iPad *smartphones* e Android. Apesar de o jQuery suportar quase todos os navegadores, existem alguns problemas com algumas versões, como a versão 1.0 do Explorer e 1.0.x do Firefox, porém ele não gera nenhum erro grave. Em geral, são compatíveis com quase todos os navegadores.

Especificamente, no projeto do novo sistema Agritempo, optou-se pela biblioteca Dojo, uma vez que suas classes e funções assemelham-se mais à estrutura da linguagem JAVA usada no projeto. O uso do Dojo tem permitido um desenvolvimento mais ágil da interface do sistema uma vez que a biblioteca já possui funcionalidades que permitem criar interfaces com interação semelhante às da Web 2.0, o que é um dos objetivos do projeto. A Figura 1 mostra a integração da biblioteca Dojo na estrutura de desenvolvimento do sistema Agritempo, usado para apoiar o desenvolvimento de *tags* e funções *javascript* que são usadas para a criação dos elementos gráficos que compõem a interface web do sistema

De forma geral, as bibliotecas possuem vantagens e desvantagens. Em relação ao JQuery e ao Dojo, ambas possuem uma vasta documentação, com funções semelhantes e outras, diferentes. Dependendo do navegador, a biblioteca permite uma execução mais rápida, ou mais lenta, com erros ou não. Dessa forma, a escolha da biblioteca deve considerar a análise completa da ferramenta e as necessidades da aplicação que se deseja projetar.



**Figura 1.** Esquema da integração da biblioteca Dojo no sistema Agritempo.

## Referências

COLTRI, P. P.; PINTO, H. S.; FERREIRA, N. J.; CECCARELLI, M.; CORAL, G. Sistema de monitoramento e previsão agrometeorológica: Agritempo. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA, 15., 2007, Aracaju. **Anais...** Aracaju: SB Agro, 2007. 5 p.1 CD-ROM

EVANGELISTA, S. R. M.; TERNES, S.; SANTOS, E. H. dos; ASSAD, E. D.; ROMANI, L. A. S.; FRANZONI, A. Agroclima - Sistema de Monitoramento Agroclimatológico. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA, 13., 2003, Santa Maria. Situação atual e perspectivas da agrometeorologia: **anais**. Santa Maria: UNIFRA: SBA: UFSM, 2003. p. 603-604.

JQUERY: the write less, do more, JavaScript library. Disponível em: <<http://jquery.com/>> Acesso em: 15 mar. 2011

UNBEATABLE JavaScript tools – The dojo toolkit. Disponível em: <<http://dojotoolkit.org/>> Acesso em: 15 mar. 2011

# Flexibilidade para aplicações web desenvolvidas com o FlexBuilder

Gabriel Garcia do Nascimento<sup>1</sup>

Luciana Alvim Santos Romani<sup>2</sup>

O Agritempo é um sistema web de monitoramento agrometeorológico que fornece informações para o usuário por meio de mapas e gráficos. Em sua nova versão, o Agritempo pretende prover mais flexibilidade e autonomia para seus usuários. Dessa forma, está sendo desenvolvida uma nova ferramenta para geração de gráficos mais flexíveis, integrados à base de dados do sistema.

Diversas soluções foram analisadas, como por exemplo a biblioteca javascript do Dojo (DOJO, 2011). Embora essa biblioteca esteja sendo utilizada no desenvolvimento do sistema Agritempo, seu uso na geração de gráficos não se mostrou adequado por possuir uma interface pouco flexível com codificação complexa. Com isso, buscou-se outra alternativa e a solução adotada foi a ferramenta FlexBuilder (ADOBE, 2011).

A usabilidade da interface do software de desenvolvimento, seus gráficos com design moderno e flexível, somados a uma linguagem declarativa e fácil de usar, bem como a existência de uma comunidade grande de desenvolvedores e total compatibilidade tanto com os computadores clientes como o servidor, fizeram do Flex a melhor escolha para o projeto Agritempo.

Durante o desenvolvimento do novo módulo de geração de gráficos do Agritempo (COLTRI et al., 2007; EVANGELISTA et al., 2003), foram utilizados tutoriais e informações disponíveis nos websites da biblioteca FlexBuilder. O Adobe FlexBuilder é um framework de desenvolvimento baseado no Eclipse, disponível para MacOS e Windows, que permite a

---

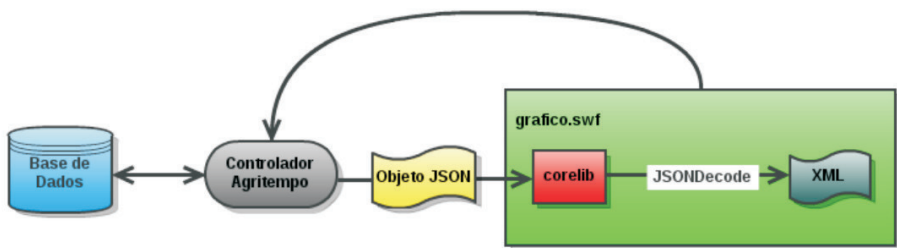
<sup>1</sup> PUC Campinas, [mjga@terra.com.br](mailto:mjga@terra.com.br)

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

codificação e edição para MXML e ActionScript, além de possibilitar a projeção visual do layout de interface, aparência e comportamento dos componentes. Utiliza uma ampla biblioteca de componentes incorporados que podem ser modificados ou mesmo criados de acordo com a necessidade da aplicação.

O maior diferencial do FlexBuilder está no modo de exibição de interface. Nela é possível modelar visualmente o layout da aplicação, criar novos objetos como botões, canvases, adicionar imagens, gráficos, datagrids, caixas de texto, e ainda editar os eventos de cada elemento. Todas essas funcionalidade podem ser usadas sem que haja a necessidade de editar o código fonte do programa, uma vez que o Flex se encarrega de sua geração conforme a interface é modificada (What You See Is What You Get - WYSWYG).

No contexto do Agritempo, o componente dos gráficos elaborado no FlexBuilder, requisita ao servidor um objeto JavaScript Object Notation (JSON) do tipo 'ClimaDiario', porém, para que seja utilizado em linguagem MXML, o objeto JSON tem de ser decodificado para XML, que é o formato padrão de entrada de dados da linguagem MXML. A biblioteca 'corelib' do Flex possui todas as funções de codificação e decodificação de objetos JSON, sendo necessário apenas importá-la e utilizar o método JSONDecoder para que o objeto seja convertido para XML, esse esquema pode ser visto na Figura 1.

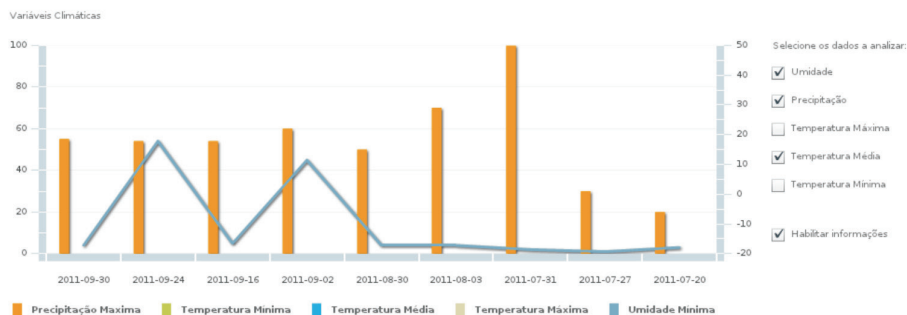


**Figura 1.** Como funciona a aplicação.

O módulo do gráfico foi gerado com rapidez, leveza e design moderno e flexível. O modo de edição visual do FlexBuilder, somado à sua vasta biblioteca de funções, tudo isso reduziu muito o número de alterações no sis-



tema Agritempo durante sua implantação. Foi necessária apenas a implementação do controlador para selecionar as estações a serem analisadas nos gráficos, e os objetos que são requisitados e enviados pelo servidor permaneceram os mesmos. O resultado final pode ser visto na Figura 2.



**Figura 2.** Resultado final da aplicação dos gráficos.

Dessa forma, a adoção de uma ferramenta moderna como o FlexBuilder permite que o novo Agritempo tenha uma interface mais próxima do conceito web 2.0. Esse novo modelo de interface objetiva dar mais flexibilidade e poder ao usuário durante sua interação com o sistema; dessa forma, o Agritempo consolida-se como um sistema de monitoramento agrometeorológico que objetiva atingir de forma adequada uma vasta gama de clientes.

## Referências

- ADOBE flash builder 4.5 premium. 2011. Disponível em: <<http://www.adobe.com/products/flash-builder.html>>. Acesso em: 21 out. 2011.
- COLTRI, P. P.; PINTO, H. S.; FERREIRA, N. J.; CECCARELLI, M.; CORAL, G. Sistema de monitoramento e previsão agrometeorológica: Agritempo. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA, 15., 2007, Aracaju. **Anais...** Aracaju: SBAGro, 2007. 5 p.1 CD-ROM.
- DOJO toolkit. 2011. Disponível em: <<http://dojotoolkit.org/>>. Acesso em: 21 out. 2011.

EVANGELISTA, S. R. M.; TERNES, S.; SANTOS, E. H. dos; ASSAD, E. D.; ROMANI, L. A. S. ; OTAVIAN, A. F. Agroclima sistema de monitoramento agroclimatológico.. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA, 2003, Santa Maria. **Anais...** Santa Maria: UNIFRA, SBA, UFSM, 2003. v. 1. p. 603-604.

TOUR de Flex. 2011. Disponível em: <<http://www.adobe.com/devnet/flex/tourdeflex.html>>. Acesso em: 21 out. 2011.

## Modernizando a apresentação de trabalhos com PREZI

Marcelo Ceron Massruhá<sup>1</sup>  
Luciana Alvim Santos Romani<sup>2</sup>

Uma das formas de disseminação dos resultados do trabalho científico é a apresentação oral em palestras, cursos, seminários, congressos, entre outros. Para que se tenha o sucesso na comunicação do trabalho científico, é fundamental preparar-se bem, explorando adequadamente os recursos audiovisuais disponíveis, bem como dominando o conteúdo e treinando a apresentação. Portanto, é fundamental entender não só do conteúdo a ser apresentado, mas também conhecer o perfil da audiência e ainda produzir um bom material de apoio.

Um dos maiores desafios ao se preparar uma apresentação é como produzir um material interessante, capaz de prender a atenção e despertar o interesse dos ouvintes, de forma que não os canse. Não existe uma fórmula mágica; na verdade, para cada ideia, existem diferentes técnicas que podem ser empregadas.

O princípio básico para que a apresentação não se torne cansativa é, certamente, evitar ao máximo, a utilização de muito texto, o excesso de informação, a ausência de cores, etc. Deve-se, sempre que possível, utilizar-se de imagens e de animações, e trocar os longos textos por palavras-chave. Além disso, com o advento da tecnologia, é possível usar os diversos recursos inovadores, a fim de despertar o interesse e prender a atenção da audiência. No entanto, excessos devem ser evitados para não comprometer a apresentação. A objetividade na apresentação das ideias é fundamental para o sucesso na comunicação.

Atualmente, além dos tradicionais slides, existem recursos computacionais mais atrativos, como a ferramenta *Prezi* (PREZI, 2011), um novo paradigma

---

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

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

ma para desenvolvimento de apresentações baseado em nuvem (*cloud-based presentation*). Buscando criar apresentações modernas e bem produzidas para o projeto Agritempo, este trabalho mostra os recursos e facilidades da ferramenta Prezi.

A fim de explorar os recursos dessa nova plataforma para geração de apresentações, foram usados tutoriais no website da ferramenta <<http://prezi.com/>>. Após estudar diferentes exemplos disponíveis no website, uma nova apresentação do projeto Agritempo (EVANGELISTA et al., 2003) foi produzida. Esta apresentação teve por objetivo validar a nova plataforma e mostrar a evolução do sistema Agritempo atual para a nova versão que está sendo desenvolvida.

Para gerar uma apresentação na ferramenta Prezi, basta criar uma conta no website que, na opção gratuita, dá direito a 100Mb de espaço para armazenar os seus arquivos de apresentação gerados.

*Prezi* é uma ferramenta online, totalmente diferente dos programas para a criação de apresentações em slide, a começar pelo simples fato de que o aplicativo não se limita ao espaço retangular dos slides. É similar ao “pptPlex”, um projeto da Microsoft Office Labs que traz esse tipo de funcionalidade para o Power Point (PPTFLEX, 2011). Com *Prezi* é possível gerar uma apresentação que utiliza-se da mudança de foco e zoom para conectar diferentes ideias no mesmo espaço visual. Assim, o apresentador pode levar o público a acompanhar seu raciocínio de uma forma diferente trilhando um caminho de descobertas (PREZI, 2011). Dessa forma, a ferramenta permite substituir os tradicionais slides. No entanto, devido à grande quantidade de diferentes possibilidades, é preciso cautela para não sobrecarregar a apresentação com efeitos visuais excessivos e não desviar o foco do conteúdo.

O usuário pode organizar o conteúdo em um mapa visual, criando, assim, apresentações não-lineares. Na ferramenta, escreve-se as palavras e colocam-se os links, as imagens e os vídeos em uma única tela, explorando formatos e tamanhos que serão visualizados quando o zoom é aproximado ou afastado da imagem principal.

O efeito visual resultante dessa nova proposta de organização não-linear de apresentação provoca um impacto positivo nos expectadores, prendendo sua atenção. Mais uma vez, a tecnologia da informação apoia a pesquisa científica fornecendo recursos para melhorar e facilitar o trabalho dos pesquisadores.



Figura 1. Exemplo de apresentação com Prezi

## Referências

EVANGELISTA, S. R. M.; TERNES, S.; SANTOS, E. H. dos; ASSAD, E. D.; ROMANI, L. A. S. ; OTAVIAN, A. F. Agroclima sistema de monitoramento agroclimatológico. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA, 2003, Santa Maria. Anais... Santa Maria: UNIFRA: SBA: UFSM, 2003. v. 1. p. 603-604.

PPTFLEX. 2011. Disponível em: <<http://www.officelabs.com/projects/pptPlex/Pages/default.aspx>>. Acesso em: 25 out. 2011.

PREZI. 2011. Disponível em: <<http://prezi.com/>>. Acesso em: 25 out. 2011.

WIKIPÉDIA. 2011. Disponível em:< <http://en.wikipedia.org/wiki/Prezi>>. Acesso em: 20 out. 2011.

## Espacialização de notícias de cana-de-açúcar utilizando entidades do IBGE

Ercilia Souza Rodrigues<sup>1</sup>

Eduardo Antonio Speranza<sup>2</sup>

Maria Fernanda Moura<sup>2</sup>

Rosa Nathalie Portugal Vargas<sup>3</sup>

Solange de Oliveira Rezende<sup>3</sup>

Os estudos da Divisão Regional do Instituto Brasileiro de Geografia e Estatística (IBGE) tiveram início em 1941, tendo, como principal objetivo, sistematizar divisões regionais que vinham surgindo no país, organizando, assim, uma única divisão regional para a divulgação das estatísticas do país. Atualmente, o Brasil é dividido em cinco regiões segundo a classificação do IBGE (IBGE, 2011), sendo elas: Norte, Nordeste, Centro-Oeste, Sul e Sudeste. Porém, existe outra maneira de classificar o país, captando melhor a situação socioeconômica e as relações entre sociedade e o espaço natural. Trata-se da divisão do país em três grandes complexos regionais, as macrorregiões: Centro Sul, Nordeste e Amazônia.

Outras classificações estabelecidas pelo IBGE que dividem o território são as mesorregiões, que são áreas com características socioeconômicas comuns, utilizadas para fins estatísticos. As Microrregiões são agrupamentos de municípios limítrofes com finalidade de integrar a organização e o planejamento de municípios com similaridades econômicas e sociais.

O foco deste trabalho é gerar a classificação manual de notícias agrícolas, de acordo com essas divisões geográficas, para serem utilizadas no âmbito do projeto “Tecnologias Inovadoras em mineração de textos para a Espacialização de Notícias Agrícolas” (TIENA). O objetivo dessa classifica-

---

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

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

<sup>3</sup> Universidade de São Paulo, [{nathalie, solange}@icmc.usp.br](mailto:{nathalie, solange}@icmc.usp.br)

ção é comparar resultados de classificação automática obtida por classificadores desenvolvidos no âmbito do projeto com os resultados manuais, a fim de validar os resultados automaticamente obtidos e encontrar formas de aprimorá-los

Para realizar a classificação de notícias agrícolas, o primeiro passo é formar um *Corpus* de um domínio fixo de conhecimento; nesse caso, foram utilizadas notícias sobre cana-de-açúcar. Sardinha (2004) descreve *Corpus* como uma abordagem que se ocupa da coleta e exploração de *corpora* ou conjunto de dados linguísticos textuais que foram coletados criteriosamente, com o propósito de servir de ferramenta para uma pesquisa.

Após a montagem do *Corpus* e com um programa, desenvolvido no âmbito do projeto, é possível ler cada notícia do *Corpus* e associá-la a uma classificação, segundo a divisão regional proposta pelo IBGE, em macrorregião, região, mesorregião, microrregião, município, ou usina à qual a notícia pertence. A decisão de qual ou quais classificações devam ser aplicadas a cada notícia é realizada pelo especialista do domínio, no caso um geógrafo com algum conhecimento de produção de cana-de-açúcar, no Brasil. Conferida, subjetivamente, a classificação aplicada, esta é anotada no *Corpus*. Quando todo o *Corpus*, ou uma janela do dele, é, subjetivamente, considerado corretamente marcado, é repassado aos processos de classificação automática.

Após a classificação automática da janela do *Corpus*, os resultados são comparados com os manualmente obtidos. A comparação pode ser completamente subjetiva ou automática. Então, as diferenças consideradas significativas são estudadas caso a caso, a fim de verificar a origem do erro e como tratá-lo.

Os primeiros experimentos foram realizados sobre uma janela de 237 notícias. O objetivo foi validar o uso da ferramenta linguística Rembrandt que utiliza um vocabulário controlado para reconhecer automaticamente todas as entidades geo-gráficas em cada notícia. O foco foi apenas nas entidades reconhecidas como de "localização geográfica". Porém, como exemplificado na Figura 1, Cosan aparece como empresa, mas é uma importante dica de localização geográfica da notícia. Logo, nesses primeiros experimentos, verificamos que o especialista deveria marcar localizações com base em todas as que aparecessem explicitamente na notícia ou que fossem possíveis de localizar via alguma instituição. Além disso, após a identificação automática das entidades, há um processo de desambigua-

```

<!-- Rembrandt by v.1.3-b1841 -->
<DOC DOCID="stdin-1" LANG="pt">
<TITLE>
</TITLE>
<BODY>
{<EM ID="0" S="0" T="0" C1="TEMPO" C2="TEMPO CALEND" C3="DATA"
TG="!:Y+2010M12D13">[13/12/2010]</EM>[10:37:00]
<EM ID="1" S="0" T="2" C1="ORGANIZACAO" C2="EMPRESA"
RI="14;27;45;48" RT="sameAs;sameAs;sameAs;sameAs"
WK="Petrobras" DB="Petrobras">[Petrobras]</EM>[planeja][conter]
["][estrangeiros]["][no][alcool]
<EM ID="2" S="0" T="10" C1="LOCAL" C2="VIRTUAL"
C3="COMSOCIAL" RI="1;14;27;45;48"
RT="sameAs;sameAs;sameAs;sameAs">[Folha][de][S.][Paulo]
[-][SP][A][Petrobras]</EM>
[tenta][conter]<EM ID="3" S="0" T="20" C1="NUMERO" C2="TEXTUAL"
[uma]</EM> {[Antes][da]<EM ID="33" S="8" T="2"
C1="ORGANIZACAO" C2="EMPRESA" DB="Cosan">[Cosan]</EM>
[,][al][americana]<EM ID="34" S="3" T="6"
C1="ORGANIZACAO" C2="EMPRESA" DB="Bunge_Limited">[Bunge]</EM>
.....
.....
....
</BODY>
</DOC>

```

**Figura 1.** Exemplo de classificação por software, onde são encontradas entidades como tempo, organização (empresa) e local.

ção das classificações, implementado no âmbito do projeto. Por exemplo, para decidir se Belém está aqui ou em algum outro estado cujo nome da cidade também seja Belém, essa de-sambiguação é feita pelo uso das coordenadas geográficas de cada entidade.

Nos trabalhos futuros prevê-se melhorar a classificação manual produzida na janela de 237 notícias, corrigindo erros que foram surgindo com o decorrer da classificação, como a inserção do estado do Mato Grosso na macrorregião Centro Sul. Nesse caso, apenas uma parte do estado se localiza nela, pois a maioria de sua área está na Amazônia. Com isso, espera-se verificar se o Rembrandt identifica as mesmas categorias e se o processo de desambiguação aplicado após essa identificação permite melhor aproximação dos resultados manuais dos automáticos.

## Referências

SARDINHA, T. B. **Linguística de Corpus**. Barueri: Manole, 2004. 2004. 410 p.

IBGE. **Instituto Brasileiro de Geografia e Estatística**. 2011. Disponível em: <<http://www.ibge.gov.br/home>>. Acesso em: 20 out. 2011.



# Execução dos ajustes e refinamentos da rede de drenagens do Estado do Mato Grosso do Sul no projeto GeoMS

Ana Cláudia Olivera Senturião<sup>1</sup>  
Júlio César Dalla Mora Esquerdo<sup>2</sup>

O projeto de pesquisa GeoMS, executado pela Embrapa Informática Agropecuária, em parceria com o Instituto de Meio Ambiente do Mato Grosso do Sul (Imasul), teve início no ano de 2007 e tem gerado informações georreferenciadas do Estado do Mato Grosso do Sul, baseadas em sensoriamento remoto e outras bases cartográficas (PROJETO GEOMS, 2011). Para disponibilização dos dados gerados ou organizados pelo projeto, foi desenvolvido, a partir de ferramentas livres, o SISLA – Sistema Interativo de Suporte ao Licenciamento Ambiental (SISLA, 2011), uma aplicação web que permite a visualização, aquisição e análise das informações georreferenciadas, servindo como ferramenta fundamental no processo de licenciamento ambiental no Estado. Entre as informações disponibilizadas, estão o mapeamento da cobertura vegetal, a rede de drenagens, as informações sobre o relevo, os mapas com a localização de territórios indígenas, as unidades de conservação, entre outras.

Uma das informações geradas no projeto GeoMS é o mapa das drenagens do Estado do Mato Grosso do Sul. Esse trabalho envolve 161 cartas georreferenciadas e digitalizadas, na escala 1:100.000, contendo informações de recursos hídricos do estado, como rios perenes e temporários. O principal objetivo dessa etapa é o ajuste e o refinamento das informações já existentes, geradas em levantamentos realizados na década de 60 pelo Exército Brasileiro (IBGE, 2011). Nesse processo, são utilizadas como referência as imagens do satélite CBERS-2B do ano de 2007, com resolução espacial de 20m, presentes no banco de dados do projeto GeoMS.

---

<sup>1</sup> Faculdade de Engenharia Agrícola/Unicamp, [anasenturiao@gmail.com](mailto:anasenturiao@gmail.com)

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

Para iniciar os ajustes e atualizações do mapeamento das drenagens, a equipe do projeto GeoMS recebeu da Agência de Desenvolvimento Agrário e Extensão Rural (AGRAER/MS) os dados hidrográficos digitalizados a partir dos levantamentos do Exército Brasileiro, no formato DWG (AutoDesk). Os dados foram, então, convertidos para o formato vetorial shapefile, compatível com os pacotes Spring e ArcGIS. A drenagem foi recebida nas seguintes classes: “rios\_perenes” (linhas referentes aos rios perenes contidos na carta), “rios\_temporários” (linhas referentes aos rios temporários), “linha\_drenagem” (rede de drenagem sem identificação como temporário ou permanente) e “área\_drenagem” (áreas alagadas mapeadas para o ano da confecção das cartas 1965/66, lagos e lagoas).

Após a conversão, os vetores da drenagem foram importados para o banco de dados do software Spring. No entanto, para ajuste e complementação da rede de drenagem, foi utilizado o software ArcGIS, dada sua facilidade de operação. Para tanto, foi necessário inserir nesse SIG as informações necessárias oriundas do banco de dados do Spring, além da elaboração de um projeto georreferenciado, de modo a integrar as informações geoespaciais.

Por meio da interpretação das imagens CBERS, as drenagens vêm sendo ajustadas de forma manual, sendo movidas e/ou redesenhadas sobre as feições de drenagem identificadas nas imagens. Além disso, quando o intérprete identifica linhas de drenagens que não constam na base cartográfica original, elas são digitalizadas e inseridas na classe denominada “rios\_cnptia”.

Ao longo desse processo de ajustes, tem-se verificado que muitas linhas de drenagens originais contidas nos arquivos digitais recebidos têm apresentado deslocamentos em relação às imagens de satélite atuais, devido, em alguns casos, às próprias mudanças fluviomorfológicas nos últimos 50 anos. No caso dos grandes rios, digitalizados por meio de linhas duplas, tem-se utilizado como referência o mapa da cobertura vegetal e uso da terra, ora em elaboração no âmbito do projeto GeoMS, tomando-se como referência os limites das matas ciliares mapeadas ao longo dos grandes rios.

Após o ajuste das drenagens, o processo seguinte é a inclusão do nome dos rios, inseridos nos campos de atributos dos arquivos vetoriais. Tais nomes são obtidos a partir das cartas originais do Exército. Por fim, é realizada a articulação entre as diversas cartas, que consiste na conexão

entre as linhas de drenagem de uma carta nos limites com suas cartas vizinhas.

Quando as cartas topográficas na escala 1:100.000 são finalizadas pela equipe do projeto GeoMS, elas são encaminhadas para uma equipe de auditoria do Imasul, que revisa e retorna uma avaliação sobre os resultados. Quando são encontrados erros, estes são corrigidos e reenviados ao Instituto. As dificuldades mais comuns que provocam erros na interpretação das imagens são devidas à resolução espacial da imagem CBERS, onde cada pixel equivale a 20 metros, tornando a identificação mais difícil no caso dos rios de menor largura. Outra dificuldade ocorre nas regiões da planície do Pantanal, onde o relevo plano faz com que as drenagens não assumam padrões claros de escoamento superficial.

Os resultados desse processo de ajustes e refinamento da rede de drenagens serão disponibilizados no SISLA, sendo de grande importância na tomada de decisão para a implantação de projetos que requeiram licença ambiental no Estado e para a definição das áreas de preservação permanente ao redor da rede hidrográfica.

## Referências

IBGE. **Instituto Brasileiro de Geografia e Estatística**. Disponível em <<http://www.ibge.gov.br>>. Acesso em 30 out. 2011.

PROJETO GeoMS. Disponível em: <<http://www.macroprograma1.cnptia.embrapa.br/projeto/geoms>>. Acesso em: 21 out. 2011.

SISLA - **Sistema Interativo de Suporte ao Licenciamento Ambiental**. Disponível em: <<http://sisla.imasul.ms.gov.br>>. Acesso em: 10 nov. 2011.

# Calibração do modelo cropsyst para cana-de-açúcar: estudo preliminar

Felipe Ferreira Bocca<sup>1</sup>  
Alexandre de Castro<sup>2</sup>

## Introdução

O cultivo da cana-de-açúcar possui grande importância no Brasil, e sua produção deve aumentar nos próximos anos. Nesse contexto, a gerência da área agrícola demandará ferramentas de suporte à tomada de decisão, uma vez que o setor sucroenergético possui uma demanda por um fluxo de matéria-prima para atender a unidade industrial ao longo da safra, sendo necessário prever o quanto será produzido, e quando ocorrerá essa produção. Os métodos estatísticos disponíveis, atualmente, apresentam limitações quanto à extrapolação de resultados, enquanto os modelos ecofisiológicos permitem simular os efeitos das condições de cultivo na cana-de-açúcar, permitindo estimativas coerentes para a curva de crescimento. Dentre os diversos pacotes computacionais disponíveis para simular o crescimento da cana-de-açúcar, a plataforma CropSyst teve em 2009, o modelo para cana-de-açúcar implementado. Foi realizado um estudo preliminar da calibração do modelo CropSyst para cultivares de cana do Brasil com objetivo de melhorar o entendimento da ferramenta de calibração. O objetivo da calibração é de estabelecer os parâmetros ligados à acumulação de tempo termal, particionamento de biomassa e utilização de água e radiação pela cultura.

---

<sup>1</sup> Universidade Estadual de Campinas, felipefbca@gmail.com

<sup>2</sup> Embrapa Informática Agropecuária, acastro@cnptia.embrapa.br

## Material e métodos

Foi utilizado no desenvolvimento desse trabalho, o software CropSyst (versão 4.15.07). Os dados utilizados são referentes à curva de crescimento da cultivar CB 41-76 plantada em fevereiro de 1974 e colhida em setembro de 1975, em Piracicaba, sob irrigação plena. Foram utilizados os dados meteorológicos diários da estação meteorológica da Escola Superior de Agricultura “Luiz de Queiroz” (Esalq), Piracicaba.

## Resultados e discussão

Para computo da acumulação de tempo termal para os eventos fenológicos, foram utilizadas as datas de 02/02/74 para o plantio, 28/02/74 para emergência, 01/11/74 para pico do índice de área foliar (IAF) e 02/08/75 para colheita. Embora tenha sido selecionada a opção de não fornecer data para os eventos de “início do florescimento” e “início do enchimento de grãos”, o software preencheu os campos com a data do fim do desenvolvimento do IAF. Como parâmetros de acumulação de tempo termal, foi utilizada uma temperatura base de 18 °C (BARBIERI et al., 2010) e uma temperatura máxima que limita crescimento de 34 °C, valor também utilizado por Tatsch et al. (2009).

Baseado nas temperaturas de base (mínima e máxima) fornecidas, o software computou o acúmulo de tempo termal para os eventos fenológicos. O valor encontrado para emergência de 222 graus-dia e para o fim do crescimento de novas folhas de 953 graus-dia foram considerados adequados, uma vez que Ferreira Junior (2010) encontrou a estabilização do IAF na faixa de 750~1000 [GD] para variedades RB cultivadas em Alagoas (TATCH et al., 2009) cita para emergência da cana o valor de 100 graus-dia.

Após a calibração da fenologia, procedeu-se a calibração da evolução de biomassa e IAF onde a série de dados de biomassa e IAF foi utilizada como entrada nessa etapa. Os parâmetros gerados pelo ajuste foram a área foliar específica de 12,06 m<sup>2</sup>.kg<sup>-1</sup>, e o coeficiente de particionamento de 3,06. O coeficiente de particionamento é calculado pela razão entre a

massa de colmos e a massa de folhas. Os valores determinados se encontram dentro da faixa esperada. Pinto et al. (2006) encontraram valores de área específica para o dossel da cana-de-açúcar entre 13, 48 e 15,73  $\text{m}^2 \cdot \text{kg}^{-1}$ . Por outro lado, o coeficiente de particionamento de 3,06 equivale a uma proporção de 75 % da biomassa em colmos, valor próximo ao esperado tipicamente de 70 %.

A última etapa da calibração consiste em calibrar os parâmetros de crescimento da biomassa e produtividade. Os parâmetros de temperatura média que limitam o crescimento, duração do dossel e eficiência do uso de radiação devem ser ajustados manualmente pelo usuário na calibração visando o melhor ajuste na comparação da curva de crescimento do modelo com os dados experimentais. A tela do programa, onde é feita a comparação, pode ser observada na Figura 1. Foram obtidos os valores de 34 °C, 1800 GD, e 6  $\text{g} \cdot \text{MJ}^{-1}$  para a temperatura média que limita o crescimento inicial, duração do dossel e eficiência do uso de radiação, respectivamente.

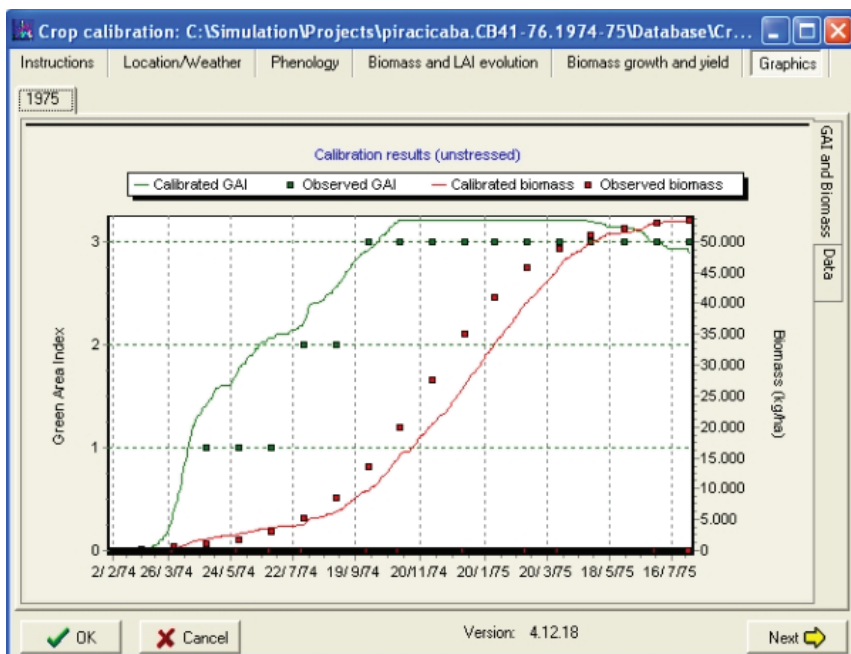


Figura 1. Comparação entre o modelo calibrado e os dados originais.

Embora os valores do IAF (GAI, Green Area Index) estejam truncados no gráfico, os valores correspondem aos fornecidos para obtenção dos parâmetros de área foliar específica e coeficiente de particionamento.

Na comparação, a curva de crescimento do modelo apresentou divergência quanto ao desenvolvimento intermediário da cultura, porém o comportamento das curvas experimentais e modelada foi similar. O modelo estimou uma produtividade de 53.617,4 kg de matéria seca para biomassa aérea, enquanto foram observados 53.584 kg experimentalmente. Os valores obtidos na calibração para o coeficiente de transpiração da biomassa [kPa], eficiência de uso da radiação [g.MJ<sup>-1</sup>] e área específica do dossel [m<sup>2</sup>.kg<sup>-1</sup>] estão compatíveis com o disponível na literatura, sendo 8,82, 6 e 12, enquanto os utilizados por Tatsch et al. (2009) foram 10, 4 e 14 respectivamente.

## Conclusão

Foi possível utilizar a ferramenta de calibração da plataforma CropSyst para calibrar a Cana-de-açúcar a partir de dados de experimentos de crescimento.

Os parâmetros obtidos na calibração foram considerados adequados quando comparados com os disponíveis na literatura.

## Referências

BARBIERI, V.; SILVA, F. C. da; DIAS-AMBRONA, C. G. H. Modelagem de cana de açúcar para previsão de produtividade de canaviais no Brasil e na Austrália. In: CONGRESO DE AGROINFORMÁTICA, 2., 2010, Buenos Aires. **Anales...** Buenos Aires: Sociedad Argentina de Informática, 2010. p. 745-762. JAIIO - CAI 2010.

FERREIRA JUNIOR, R. A. **Crescimento de variedades RB de cana-de-açúcar irrigadas e fotossíntese modelada pela radiação solar.** 2010. 68 f. Dissertação (Mestrado em Proteção de Plantas) – Centro de Ciências Agrárias, Universidade Federal de Alagoas, Rio Largo, AL.

PINTO, L. F. G.; BERNARDES, M. S.; PEREIRA, A. R. Yield and performance of sugarcane in on-farm interface with rubber in Brazil. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 41, n. 2, p. 251-255, fev. 2006.

TATSCH, J. D. ; BINDI, M. ; MORIONDO, M. A Preliminary Evaluation of the Cropsyst Model for Sugarcane in the Southeast of Brazil. In: BINDI, M. I; BRANDANI, G.; DIBARI, C.; DESSÌ, A.; FERRISE, R.; MORIONDO, M.; TROMBI, G. (Org.). **Impact of climate change on agricultural and natural ecosystems.** Florença: Firenze University, 2009. p. 75-84.



## **Atualização da rede de drenagens e inserção de sua nomenclatura no Estado do Mato Grosso do Sul**

Letícia Beatriz de Barros Brito<sup>1</sup>  
Júlio César Dalla Mora Esquerdo<sup>2</sup>

O Projeto GeoMS visa à identificação, à caracterização e ao mapeamento das áreas desmatadas, à cobertura vegetal e ao uso da terra, para auxiliar a implantação de projetos estratégicos, promover a preservação e a conservação ambiental, além de um desenvolvimento sustentável do Estado do Mato Grosso do Sul (PROJETO GEOMS, 2011). O projeto é dividido basicamente em três áreas: o desenvolvimento do Sistema Interativo de Suporte ao Licenciamento Ambiental (Sisla) (SISLA, 2011); o levantamento e mapeamento de áreas degradadas e cobertura vegetal e uso da terra; e o ajuste da drenagem e delimitação das bacias hidrográficas. Um dos objetivos dessa terceira parte é a atualização das drenagens do Estado do Mato Grosso do Sul.

O mapa das drenagens, atualmente disponível no estado, foi realizado na década de 60 pelo Exército Brasileiro (IBGE, 2011), por meio de aerofotogrametria, sendo necessário um ajuste no delineamento dos rios que sofreram mudanças ao longo do tempo. Além do ajuste das drenagens contidas na base original, tem sido realizada também a inserção de novas drenagens observadas nas imagens de satélite (CBERS), tomadas como referência nesse processo.

As drenagens antigas foram disponibilizadas ao projeto em formato digital vetorial e inseridas em ambiente ArcGIS, depois de importadas do software SPRING.

---

<sup>1</sup> Universidade Estadual de Campinas, [leticia\\_brito@hotmail.com](mailto:leticia_brito@hotmail.com)

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

Para o ajuste das drenagens são utilizados como referência as imagens do satélite CBERS do ano de 2007 (resolução espacial de 20m), os mapas de classificação da vegetação (gerados pelo projeto GeoMS) e o software Google Earth. O Estado do Mato Grosso do Sul está dividido em 34 cartas na escala 1:250.000, que possuem 6 subdivisões na escala 1:100.000. O resultado final será a divisão das drenagens de acordo com as cartas 1:100.000, separadas nas seguintes classes: “linha\_drenagem”, “rios\_perenes”, “rios\_temporarios” e “rios\_cnptia”. Essa classificação das drenagens segue a nomenclatura original da base do Exército, com exceção da classe “rios\_cnptia”, contendo as drenagens novas que não estavam presente nas cartas originais.

Como as drenagens extraídas das cartas topográficas são muito antigas, muitas vezes elas não se ajustam aos limites das matas ciliares. Os mapas de classificação da vegetação auxiliam nessa delimitação das matas ciliares, indicando onde há possíveis cursos d’água e drenagens. Pelos mapas da vegetação, também são extraídos os “rios duplos”, que são rios muito largos que não são vetorizados apenas como uma linha simples, mas como um polígono.

Já o Google Earth auxilia na inserção de novas drenagens, pois a resolução espacial das imagens é maior, além de ser possível observar o relevo da região. Quando alguma feição é observada na imagem CBERS indicando a existência de uma drenagem não presente na base original, a região é checada nas imagens de alta definição presentes no Google Earth, além do relevo, que também pode ser observado.

Para a inserção da nomenclatura das drenagens, são utilizadas as cartas topográficas originais escaneadas, que possuem informações a respeito das drenagens e o nome delas.

Até o momento, foram ajustadas 116 cartas na escala 1:100.000 contidas em 25 cartas na escala 1:250.000, e inserida a nomenclatura delas. Após o término das atividades, as cartas são enviadas para o Instituto de Meio Ambiente do Mato Grosso do Sul (Imasul), onde é feita uma avaliação mais detalhada e, então, é gerado um relatório a respeito dos erros encontrados.

Com o desenvolvimento das atividades foi possível notar algumas dificuldades, principalmente na região do Pantanal, onde o relevo muito plano não permite o correto delineamento das drenagens, que mudam

com bastante frequência. Outra dificuldade se deve ao fato das imagens CBERS possuírem resolução espacial de 20m, o que é insuficiente para a delimitação exata de rios de menor largura. Muitas vezes é possível ver com clareza a drenagem na imagem presente no Google Earth, mas na imagem CBERS isso se torna muito mais difícil, perdendo um pouco a exatidão do trabalho.

Apesar dessas dificuldades, os resultados tem se mostrado cada vez melhor, pois ao longo dos meses foram inseridas novas ferramentas para auxiliar no ajuste e extração das drenagens, tornando o trabalho mais fácil e mais exato.

## Referências

IBGE. **Instituto Brasileiro de Geografia**. Disponível em <<http://www.ibge.gov.br>>. Acesso em: 30 out. 2011.

PROJETO GeoMS. Disponível em: <<http://www.macroprograma1.cnpia.embrapa.br/projeto/geoms>>. Acesso em: 21 out. 2011.

SISLA (**Sistema Interativo de Suporte ao Licenciamento Ambiental**). Disponível em <<http://sisla.imasul.ms.gov.br>>. Acesso em: 10 nov. 2011.

# 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).

## **Verificação da acurácia do mapeamento de áreas de agricultura localizadas em áreas desflorestadas na Amazônia**

Laís Silva de Oliveira<sup>1</sup>

Júlio César Dalla Mora Esquerdo<sup>2</sup>

Alexandre Camargo Coutinho<sup>2</sup>

Há muito tempo, a região que compõe a Amazônia Legal vem sendo alvo de intensas atividades de desflorestamento, caracterizando um dos mais graves problemas ambientais enfrentados pelo Brasil. Muitas atividades relacionadas à ocorrência desse fenômeno ocorrem pela busca incessante de crescimento econômico na região, geralmente promovido à custa do aumento das atividades relacionadas ao agronegócio brasileiro (CASTRO, 2005; COSTA, 2005). Dentre elas, pode-se caracterizar a expansão da área agrícola, na qual porções territoriais, que antes eram cobertas por florestas, atualmente transformaram-se em áreas de cultivo agrícola e de pecuária. A compreensão do processo de ocupação de áreas desflorestadas pela agricultura depende, fundamentalmente, da capacidade de promover sua delimitação espacial e seu monitoramento sistemático. Métodos de qualificação e quantificação dessas áreas, a partir do uso de imagens orbitais, são muito adequados e frequentemente adotados no cumprimento dessa tarefa (IPPOLITI-RAMILO, 1999). Contudo, para que os resultados obtidos por esses métodos de classificação possam ser utilizados e divulgados, faz-se necessária a sua validação e a sua verificação da acurácia dos mapeamentos efetuados. Este trabalho teve por objetivo avaliar a acurácia de um método de mapeamento da agricultura em áreas desflorestadas da Amazônia, baseado nos valores máximos e mínimos de biomassa, expressos pelas imagens do índice de vegetação Normalized Difference Vegetation Index (NDVI),

---

<sup>1</sup> Faculdade de Engenharia Agrícola/Unicamp, [la.silvaoliveira@gmail.com](mailto:la.silvaoliveira@gmail.com)

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

obtidas do sensor Moderate Resolution Imaging Spectroradiometer (Modis).

A região de estudo envolveu as áreas desflorestadas da Amazônia em nove estados brasileiros: Acre, Amapá, Amazonas, Mato Grosso, Pará, Rondônia, Roraima, Tocantins e parte do Maranhão. O mapeamento da agricultura nas áreas desflorestadas desses estados foi realizado dentro do âmbito do projeto TerraClass, com a participação da Embrapa Informática Agropecuária. A partir desses mapeamentos estaduais, foi efetuada uma verificação de sua acurácia, por meio do uso de imagens Landsat-5/TM, georreferenciadas e mosaicadas, além de imagens de NDVI quinzenais empilhadas do satélite Modis, compreendendo o período entre agosto de 2009 e julho de 2010.

Inicialmente, a verificação foi realizada com base em pontos distribuídos randomicamente nos estados ou partes deles. Com o auxílio do software ArcGIS, foram sobrepostos o mosaico de imagens Landsat-5/TM e a imagem classificada de cada estado. Para consolidar essa abordagem, foi criada uma tabela de atributos dos pontos gerados contendo, além da classe temática “agricultura anual”, as seguintes classes: “vegetação secundária”, “pastagem” e “outros”. Essas classes foram utilizadas apenas com o intuito de qualificar os erros encontrados nas áreas de desflorestamento, inicialmente mapeadas como sendo utilizadas pela agricultura anual. A tabela de atributos foi preenchida com os números 0 ou 1, de acordo com o erro ou acerto da classificação, sendo o número 1 utilizado para caracterizar os pontos em que a classificação foi bem sucedida e 0 para caracterizar os pontos em que a classificação foi equivocada. A quantidade de pontos aleatórios gerados variou de acordo com as características da ocorrência e distribuição da agricultura em cada estado.

Como resultado parcial das atividades, pode-se citar o estado do Mato Grosso, no qual foram verificados 531 pontos distribuídos aleatoriamente dentro das áreas desflorestadas. Desses pontos, 457 foram corretamente identificados na classe agricultura, o que representou cerca de 86% de acerto. As maiores confusões ocorreram com a classe “pastagem”, que em alguns locais apresentou comportamento espectro-temporal semelhante ao da agricultura anual. Os números mostraram que o classificador apresentou bons resultados, sendo uma alternativa viável, de baixo custo e de maior velocidade de execução, quando comparado com o desempenho dos métodos tradicionais, baseados em classificadores supervisionados.



Entretanto, novas avaliações de acurácia serão conduzidas em outros estados, para verificação dos resultados do mapeamento em cada um deles e consolidação da validação global do projeto.

Os resultados parciais alcançados demonstraram que a técnica de mapeamento da atividade agrícola empregada pelo projeto, baseada na avaliação do comportamento espectro-temporal do índice de vegetação NDVI, tem sido boa, sendo uma alternativa viável, de baixo custo e de rápida execução.

## Referências

CASTRO, E. Dinâmica socioeconômica e desmatamento na Amazônia. **Novos cadernos do NAEA**, Belém, v. 8, n.2, p. 5-40, 2005.

COSTA, F. A. Questão agrária e macropolíticas para a Amazônia. **Estudos Avançados**, São Paulo, v. 19, n. 53, p. 131-156, 2005.

IPPOLITI-RAMILO, G. A.; EPIPHANIO, J. C. N.; SHIMABUKURO, Y. E. FORMAGGIO, A. R. Sensoriamento remoto orbital como meio auxiliar na previsão de safras. **Agricultura em São Paulo**, São Paulo, v. 46, n. 1, p. 89-101, 1999.

## Busca usando sinônimos no Ainfo-Consulta

Rogério Lecarião Leite<sup>1</sup>

Edmilson José Mangueira Carvalho<sup>1</sup>

Glauber José Vaz<sup>2</sup>

O AINFO é um sistema para automação de bibliotecas e recuperação de informações que permite a gestão da informação técnico-científica. É utilizado pelas bibliotecas da Embrapa, por organizações estaduais de pesquisa agropecuária e por outras instituições públicas e privadas (EMBRAPA INFORMÁTICA AGROPECUÁRIA, 2011a). Atualmente, o AINFO 6 é composto por três módulos: Ainfo-Gestor, responsável pelo registro e pelo controle de acervos, Ainfo-Digital, que armazena publicações digitais, e Ainfo-Consulta, que permite a realização de buscas de informações textuais baseadas nos registros do sistema e pode ser acessado sem necessidade de cadastro por qualquer terminal que possua conexão com a internet. A atual versão do Ainfo-Consulta não possui recurso de expansão de consulta por sinônimos, mas é interessante que, em uma busca com a utilização de determinada palavra, o resultado seja próximo àquele obtido em uma busca realizada com termo equivalente ou relacionado a esse termo. O foco deste trabalho é na expansão de consultas com termos obtidos por meio do Thesaurus Brasileiro de Agricultura - Thesagro (BRASIL, 1999) para melhorar os resultados do Ainfo-Consulta.

Para ilustrar essa situação, foram realizadas consultas na atual versão do Ainfo-Consulta (EMBRAPA INFORMÁTICA AGROPECUÁRIA, 2011b) com os termos 'mandioca', 'manihot esculenta', nome científico da mandioca, e 'aipim', outro nome usado para a palavra 'mandioca'. Na Tabela 1, são exibidas as consultas realizadas por um usuário, associando os termos e operadores utilizados à quantidade de documentos recuperados.

---

<sup>1</sup> Universidade Estadual de Campinas - Faculdade de Tecnologia, {edmilsonjmc, rogerioll}@cnptia.embrapa.br

<sup>2</sup> Embrapa Informática Agropecuária, glauber@cnptia.embrapa.br

**Tabela 1.** Busca por sinônimos no Ainfo-Consulta

Termo e operadores utilizados	Recuperados
mandioca	39324
“manihot esculenta”	4117
aipim	333
mandioca OU “manihot esculenta” OU aipim	39447

Observa-se claramente uma grande perda de informação que pode ser relevante. Com a exploração das relações entre termos apresentadas no Thesagro, o resultado de uma consulta utilizando individualmente qualquer um dos termos considerados traz 39.447 registros, equivalente ao resultado obtido pela consulta ‘mandioca OU “manihot esculenta” OU aipim’. No entanto, na última versão do Ainfo-Consulta, a quantidade de registros recuperados por consultas individuais aos termos ‘mandioca’, ‘manihot esculenta’ e ‘aipim’ são 39.324, 4.117 e 333, respectivamente.

Portanto, se são considerados relevantes todos os documentos que apresentam sinônimos de palavras pesquisadas, apenas 10,44% dos documentos relevantes são recuperados na consulta a ‘manihot esculenta’ e, pior ainda, 0,84% na consulta a ‘aipim’.

O Thesagro especifica relações de equivalência USE (utilize-se) e UF (utilize-se para) entre termos justamente para indicar sinônimos. E essas relações são remissivas. Na Figura 1, são ilustradas as entradas no *thesaurus* para ‘MANDIOCA’, ‘AIPIM’ e ‘MANIHOT ESCULENTA’. Além das relações USE e UF, há outras que podem ser exploradas

	<b>MANDIOCA</b>
AIPIM	UF AIPIM
USE <b>MANDIOCA</b>	UF MACAXEIRA
	UF MANIHOT ESCULENTA
	UF MANIHOT UTILISSIMA
MANIHOT ESCULENTA	<b>BT</b> TUBERCULO
USE <b>MANDIOCA</b>	NT MANDIOCA BRAVA
	NT MANDIOCA MANSA
	RT TAPIOCA
	RT MANIVA

**Figura 1.** Exemplos de descritores no Thesagro.

futuramente, como as hierárquicas BT e NT, que indicam generalizações e especializações, e as de associação RT.

De acordo com o Thesagro, o termo AIPIM é direcionado para MANDIOCA, assim como MANIHOT ESCULENTA. O termo MANDIOCA, por sua vez, é utilizado para representar os termos AIPIM, MACAXEIRA, MANIHOT ESCULENTA e MANIHOT UTILISSIMA. Uma consulta a 'aipim', portanto, pode ser facilmente expandida para uma consulta que envolve os termos 'mandioca', 'macaxeira', 'manihot esculenta' e 'manihot utilissima'.

Para sua implementação, desde as primeiras versões, o Ainfo-Consulta é construído com o Apache Solr, uma plataforma de busca de código aberto (THE APACHE SOFTWARE FOUNDATION, 2011; SMILEY; PUGH, 2009). O Solr utiliza a biblioteca de recuperação de informação de alto desempenho e escalável Apache Lucene (GOSPODNETIC; HATCHER, 2005). Essas tecnologias já possuem recursos que permitem a utilização de sinônimos nas consultas e também ordenam os documentos recuperados de acordo com sua relevância.

Portanto, o Ainfo-Consulta, em sua versão atual, trata apenas os termos fornecidos explicitamente pelos usuários nas consultas, não oferecendo suporte para uma busca envolvendo os sinônimos desses termos, o que pode representar a perda de muitos documentos relevantes. O Solr e o Lucene fornecem recursos que possibilitam a exploração de sinônimos nas buscas e o Thesagro fornece um excelente conjunto de termos do domínio agrícola, identificando sinonímias que são úteis para a expansão das consultas no Ainfo-Consulta. Posteriormente, será possível trabalhar com pesos diferentes de relevância para cada termo acrescentado à consulta inicial, conforme suas relações, que podem envolver, além das equivalências, as hierárquicas e as de associação. Também será possível acrescentar descritores de outros thesaurus e novos recursos, como sugestões de termos parecidos e buscas de acordo com os perfis dos usuários.

## Referências

THE APACHE SOFTWARE FOUNDATION. **Apache Solr**. Disponível em: <<http://lucene.apache.org/solr>>. Acesso em: 24 out. 2011.

BRASIL. Secretaria de Desenvolvimento Rural. **Thesagro**: thesaurus agrícola nacional. Brasília, DF: SDR, Cenagri, 1999. 242 p.

EMBRAPA INFORMÁTICA AGROPECUÁRIA. **Ainfo**. 2011a. Disponível em: <<http://www.ainfo.cnptia.embrapa.br>>. Acesso em: 24 out. 2011.

EMBRAPA INFORMÁTICA AGROPECUÁRIA. **Ainfo-Consulta**. 2011b. Disponível em: <<http://ainfo.cnptia.embrapa.br/consulta>>. Acesso em: 24 out. 2011.

GOSPODNETIC, O.; HATCHER, E. **Lucene in action**. Greenwich: Manning, 2005. 421 p.

SMILEY, D.; PUGH, E. **Solr 1.4 Enterprise Search Server**. Birmingham, UK: Packt Publishing, 2009. 317 p.

## **Utilização da metodologia WGCNA na construção de redes gênicas e identificação de mecanismos envolvidos na resposta de vacas à mastite**

Vinícius Campidelli Gozzo<sup>1</sup>  
Poliana Fernanda Giachetto<sup>2</sup>

A técnica de microarranjos é uma metodologia de análise da expressão gênica, que permite investigar o nível de expressão de milhares de transcritos em uma amostra, de maneira simultânea. Por meio dessa tecnologia, pode-se examinar comparativamente a atividade transcricional global que ocorre entre diferentes amostras, sejam elas diferentes tipos celulares, tecidos ou condição experimental, por exemplo. No âmbito da Embrapa, a tecnologia de microarranjos tem sido utilizada na investigação de mecanismos que levam à determinação de fenótipos de interesse, como resistência a doenças e parasitas, melhor qualidade de carne ou maior produção de leite.

Dentre as análises que podem ser realizadas a partir dos dados gerados pelos microarranjos, a análise de expressão diferencial tem sido bastante aplicada e se mostrado importante para descoberta dos genes que influenciam o processo estudado. Apesar de tratar cada gene como uma entidade independente, e não considerando que haja interação entre eles, a identificação dos genes diferencialmente expressos tem trazido resultados importantes na busca pelos principais genes envolvidos na manifestação dos fenótipos de interesse na agricultura, pecuária e saúde. Por outro lado, as células exibem um comportamento de interação complexo, que geralmente não pode ser completamente predito a partir das propriedades de componentes individuais do sistema, uma vez que a regulação da expressão

---

<sup>1</sup> Universidade Estadual de Campinas, [campidelligozzo@gmail.com](mailto:campidelligozzo@gmail.com)

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

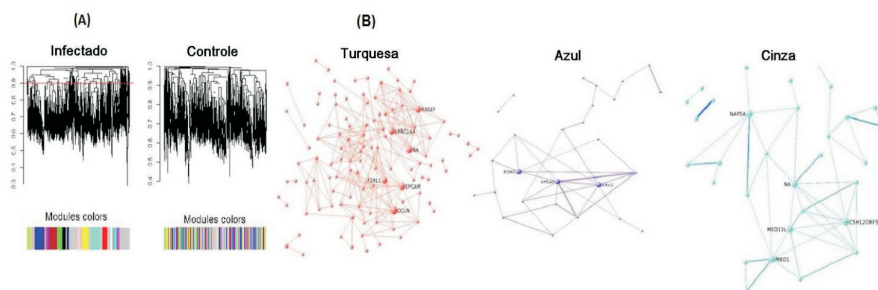
gênica e, essencialmente, todas as funções biológicas são resultantes do efeito coordenado de múltiplos genes. Assim, para se compreender como uma característica se manifesta, muitas vezes é necessária a identificação das vias biológicas envolvidas, o que inclui identificar os genes participantes e explicitar as relações entre eles. Várias metodologias podem ser utilizadas na construção de redes gênicas a partir dos dados de microarranjos. Abordagens que utilizam co-expressão na identificação de redes gênicas baseiam-se na premissa de que genes com perfis de expressão similares podem possuir funções similares (EISEN et al., 1998) e, portanto, participar de um mesmo processo biológico. Podem ser empregadas quando o objetivo é a identificação de grupos de genes envolvidos no fenótipo estudado. Além disso, com base no princípio conhecido como “princípio da culpa por associação”, que afirma que se um gene desconhecido é co-expresso com genes conhecidos, envolvidos em um mecanismo biológico particular, o gene desconhecido em questão está provavelmente envolvido no mesmo processo (SAITO et al., 2008), esse tipo de análise permite a predição da função de genes ainda não caracterizados.

Assim, tendo em vista a importância da predição das funções biológicas de genes ainda não caracterizados e da identificação de redes gênicas, esse trabalho teve por objetivo implantar uma metodologia baseada em co-expressão para análise dos dados gerados nos experimentos de microarranjos da Rede Genômica Animal. Nesse trabalho, dados de um experimento com microarranjos da Embrapa Gado de Leite, realizado para se identificar genes e mecanismos biológicos envolvidos na resposta de vacas leiteiras a infecção por microorganismos causadores da mastite foram utilizados. A mastite é uma das principais doenças que acometem os rebanhos leiteiros no Brasil e no mundo e leva a sérios prejuízos, decorrentes da queda da qualidade e da produção de leite.

Foram utilizados dois conjuntos de dados, gerados pela hibridização de chips com RNA extraído de amostras de leite de animais não infectados (n=5, grupo CONTROLE) e amostras de leite de animais colhidas 3 horas após infecção artificial com *Staphylococcus agalactiae* (n=5, grupo INFECTADO). Os dados foram pré-processados utilizando-se os pacotes *affy* e *affyQCReport* do R/Bioconductor e as redes de co-expressão construídas utilizando a metodologia WGCNA - *Weighted Gene Co-expression Network Analysis* que, a partir de medidas de correlação (Pearson), calcula a conectividade dos genes e a sobreposição topológica, possibilitando a

determinação de módulos pelo agrupamento (*clustering*) dos genes mais similares. Uma vez construídas as redes, os módulos (grupos de genes fortemente correlacionados) identificados foram comparados entre elas, a fim de se descobrir quais grupos não eram conservados, e que podem evidenciar os mecanismos biológicos distintos entre os dois grupos. Tais análises foram realizadas por meio dos pacotes do R/Bioconductor, Biobase, cluster, gcrma e WGCNA. A função molecular mais representativa dos genes de cada módulo foi identificada pela ferramenta Blast2GO (<http://www.blast2go.com/b2ghome>).

Foram construídas as 2 redes e identificados 17 módulos, sendo 3 deles não conservados entre os grupos (Figura 1 e Tabela 1). Esses módulos representam mecanismos alterados entre os 2 grupos, e que podem estar



**Figura 1.** Dendogramas com os respectivos módulos identificados (A) e visualização dos módulos não conservados entre as redes (B).

**Tabela 1.** Módulos não conservados, genes componentes e funções moleculares.

Módulo	Genes	Função molecular
Cinza	BTG3, CD3E, MBD1, CHIC2, PLXNA3, MOCS3, NEIL1, VPS45, BCL2	Apoptose e reconhecimento de antígenos
Turquesa	FGL1, GJA1, F2RL1, PTPRF, S100A2, TGFB2	Mediadores de inflamação e genes marcadores de mastite
Azul	CD97, MAD2L1, ZFP106, CDKN2C, LOC514364, NOP14, PCBD1, LOC100139798, AP1S1, EDN1, IL1B, ANXA11	Divisão celular e resposta inflamatória



envolvidos na resposta dos animais à infecção inicial por *S. agalactiae*. No momento, outras análises estão sendo conduzidas, a partir desses resultados, com o objetivo de se aprofundar o conhecimento acerca dos genes identificados, que podem levar a um melhor entendimento da resposta dos animais à mastite e identificar candidatos que serão usados em programas de melhoramento animal.

## Referências

EISEN, M. B.; SPELLMAN, P. T.; BOTSTEIN, D. Cluster analysis and display of genome-wide expression pattern. **PNAS**, v. 95, p.14863-14868, 1998.

SAITO, K; HIRAI, M. Y.; SAKAKIBARA, K. Y. Decoding genes with coexpression networks and metabolomics – majority report by precogs. **Trends in Plant Science**, v. 13, n.1, p.36-43, 2008.

# Utilização da ferramenta Liferay na construção do portal do Laboratório Multiusuário de Bioinformática da Embrapa - LMB

Fernanda Cristina de Paiva Pereira<sup>1</sup>

Glauber José Vaz<sup>2</sup>

Francisco Pereira Lobo<sup>2</sup>

Michel Eduardo Beleza Yamagishi<sup>2</sup>

Poliana Fernanda Giachetto<sup>2</sup>

O Laboratório Multiusuário de Bioinformática (LMB) da Embrapa, implantado na Embrapa Informática Agropecuária em outubro de 2011, é uma iniciativa que se propôs a integrar competências multidisciplinares e elevada capacidade computacional para atender as demandas em Bioinformática dos projetos da Embrapa e parceiros. Em função do caráter colaborativo e da necessidade de organização e controle das informações entre os usuários e a equipe do LMB, surgiu a necessidade de uma ferramenta de gerenciamento que atendesse a esses requisitos e que, ao mesmo tempo, permitisse agilidade e dinamismo na execução das ações.

Dentre as opções de ferramentas existentes, uma alternativa é o Portal Liferay, uma plataforma baseada em software livre e de código aberto, desenvolvida em Java com tecnologia Web 2.0 (SARANG, 2009), idealizado para a construção de portais web corporativos. Permite a interação com várias linguagens de programação e suporta os principais bancos de dados, servidores de aplicações e sistemas operacionais. Usuários dos portais construídos por meio dessa ferramenta podem ser agrupados hierarquicamente em organizações ou comunidades interorganizacionais, o que confere flexibilidade e permite um dinâmico e adequado compartilhamento de informações, gerenciadas por níveis de permissões. Assim, o

---

<sup>1</sup> Pontifícia Universidade Católica de Campinas, [fernandapaiva.pucc@gmail.com](mailto:fernandapaiva.pucc@gmail.com)

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

objetivo desse trabalho foi utilizar a ferramenta Liferay para a construção do portal do LMB da Embrapa.

Na construção das páginas do portal do LMB foram usadas as linguagens HTML, CSS e JavaScript, o que permitiu acessibilidade, facilidade de manutenção e ganho de tempo no desenvolvimento dos códigos. A linguagem HTML foi utilizada para a marcação estrutural do documento e a CSS utilizada para aspectos relacionados à sua apresentação, como posicionamento do texto, cores e tamanhos de fontes e outros conteúdos de estilo. A JavaScript, uma linguagem de *script*, considerada uma linguagem leve de programação com orientação a objetos, foi utilizada por apresentar recursos como fechamentos e funções de alta ordem, comumente indisponíveis em linguagens populares como Java e C++.

A ferramenta Liferay mostrou-se bastante eficiente na construção do portal do LMB (Figura 1). De simples implementação, tem permitido aos usuários



**Figura 1.** Portal do Laboratório Multiusuário de Bioinformática da Embrapa – LMB criado com o uso da ferramenta Liferay.

Fonte: Embrapa (2011).

internos e externos o acesso a páginas específicas, exigindo autorização para atualizar conteúdos ou agregar serviços existentes. Os usuários podem ter suas próprias páginas e aplicações, e facilmente trabalhar com elas em um ambiente público ou restrito. Além disso, permite executar buscas inteligentes no banco de dados que filtra itens, baseado em permissões individuais. Facilitou bastante a comunicação, funcionando como uma rede social e permitindo a interação por meio de fóruns, podendo, ainda, ser melhorada por meio da inclusão de blogs e wikis. Por meio desse portal, os laboratórios parceiros poderão fazer solicitações de acesso ao parque computacional ou de consultoria para definição das estratégias de trabalho e realização de análises de dados. Além disso, acreditamos que o portal desenvolvido irá proporcionar uma melhor organização dos dados e das informações, contribuindo, também, para melhor construção do conhecimento gerado pelos projetos de pesquisa da Embrapa e assegurando a excelência na pesquisa, desenvolvimento e inovação.

Agradeço a Embrapa Informática Agropecuária pela oportunidade concedida.

## Referências

EMBRAPA. **Laboratório Multiusuário de Bioinformática**. Disponível em: <<http://www.lmb.cnptia.embrapa.br>>. Acesso em: 1 dez. 2011.

SARANG, P. **Practical liferay**: Java-based Portal Applications Development. NewYork :Apress, 2009. 384 p.

## **Análise da rotação de cultura na área de reforma de canavial**

Caroline Nascimento Pereira<sup>1</sup>  
Aryeverton Fortes de Oliveira<sup>2</sup>

O setor sucroenergético vem sendo pressionado para atender a crescente demanda por etanol, consequência das vendas de carros flex, de açúcar e bioeletricidade, gerada a partir da queima do bagaço e da palha, além dos incipientes investimentos na indústria alcoolquímica, como futura substituta da petroquímica. Assim sendo, o estudo da eficiência produtiva no setor sucroenergético ganha importância pela necessidade de ampliar a produção de cana-de-açúcar, como procura garantir a competitividade das unidades produtivas, com o aumento da rentabilidade da usina pelas economias de escopo e de escala, a qual é possível, entre outros motivos, pelo melhor uso dos recursos, tanto na área industrial como na agrícola.

Na área agrícola percebe-se a necessidade do melhor uso do solo, o qual pode ser conseguido pela rotação de culturas na área de reforma do canavial. O “descanso” do solo ocorre compulsoriamente após cerca de seis cortes do canavial, suprimindo cerca de 20% da área cultivada. Assim, a ocupação das áreas de reforma de cana com outras espécies, além de desejável do ponto de vista da conservação de solo, representa uma oportunidade para geração de renda (SALOMÉ et al., 2007). Ademais, a reforma do canavial é um processo natural de descanso do solo após ininterruptos cortes, os quais devido ao processo de exaustão dos nutrientes, resultam em produtividade decrescente ao longo das safras. (SANTIAGO; ROSSETTO, 1999).

A partir do problema exposto, o trabalho analisará o processo de rotação de culturas em área de reforma do canavial, enfatizando a capacidade das usinas na coordenação dos processos e recursos, considerando a estrutu-

---

<sup>1</sup> Universidade Estadual de Campinas, [carolinenasper@gmail.com](mailto:carolinenasper@gmail.com)

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

ra de governança e os custos de transação, como fatores importantes para o desenvolvimento da atividade. A análise terá um caráter empírico pelo estudo de caso de três usinas em regiões heterogêneas na produção de cana-de-açúcar, baseando-se na Teoria dos Custos de Transação (TCT) (FIANI, 2002). Cada usina representará um tipo de agente no processo, sendo um agente que não realiza a rotação, um agente que realiza a rotação de forma autônoma e um agente que realiza a rotação de forma terceirizada.

A amostra utilizada no trabalho para caracterizar a rotação de culturas com grãos em áreas de reforma é composta por três usinas localizadas no Estado de São Paulo, Usina Rio Pardo, em Cerqueira César, Paraíso Bioenergia, em Brotas, e a Usina Guaíra, em Guaíra, as quais são heterogêneas tanto na produção de cana-de-açúcar, como na adoção, prática e governança da rotação. Foram aplicados questionários modelo nas três unidades visitadas, os quais seguem uma estrutura dividida em duas partes: a primeira com a caracterização do agente, da estrutura econômica-financeira e do processo inovativo, e a segunda caracterizando o processo de reforma do canavial, rotação e arranjos contratuais estabelecidos.

A análise da rotação considerou as usinas como representantes dos três casos analisados: as usinas que adotaram a rotação e realizaram de forma autônoma, as usinas que adotaram e terceirizaram e, por fim, as usinas que não adotaram a rotação.

Observou-se que as usinas apresentam interesse na adoção da rotação de culturas, pois é de amplo conhecimento os seus benefícios agrônômicos e financeiros. Entretanto, a adoção somente ocorreu quando a usina certificou-se que a nova prática não afetaria o planejamento da produção de cana-de-açúcar. E a forma de controlar a rotação dentro do planejamento da safra dependeu da estrutura de governança adotada, como se verá em cada caso.

A Paraíso Bioenergia representa o modelo de adoção pela terceirização. A usina realizou rotação na área de reforma com o amendoim e cobertura verde em cerca de 20% da área total. Todo o processo foi terceirizado, o que resultou na falta de informações completas por parte do gerente agrícola. Esse padrão caracteriza o agente que sabe das vantagens da rotação e pretende realizá-la, entretanto, opta pela terceirização do serviço para que a rotação não desvie o foco da atividade principal e também não seja necessário investir em ativos específicos para a atividade.

A Usina Guaíra adotou a rotação pela autoprodução em cerca de 10% da área total, com rotação de soja. A opção pela autoprodução resulta em controle total da produção (plantio e trato), além de possuir domínio na gestão. A desvantagem no processo próprio refere-se ao investimento em ativos específicos, como plantadeiras e colheitadeiras de soja. Esse agente representa o produtor que possui a tradição da cultura da soja e condições econômico-financeiras favoráveis para o investimento na rotação. Ademais, esse produtor se encontra em uma região tradicional de rotação cana-soja, em que esta última era a atividade principal da região há muitos anos.

E para o caso da não adoção da rotação em área de reforma, temos como exemplo a Usina Rio Pardo, a qual não teve ainda área de reforma pelo pouco tempo de operações, mas planeja realizar rotação com soja, contanto que essa não atrapalhe o cultivo de cana-de-açúcar, ou seja, a principal restrição para a rotação é a janela de plantio. Assim sendo, esse agente pode ser caracterizado pela não-adoção, com intenção de realização, contanto que essa não atrapalhe sua atividade principal e não represente custos adicionais, além de vislumbrar a possibilidade de ganhos agrônômicos e financeiros.

A análise dos três casos possibilitou inferir que as usinas apresentam interesse na rotação, dadas as vantagens agrônômicas e financeiras observadas. Esse fato é comprovado pela adoção e a intenção de adotar, pelas três usinas entrevistadas, contanto que a nova atividade não interfira no planejamento da produção de cana-de-açúcar e não represente custos adicionais para a usina.

## **Agradecimentos**

Meus agradecimentos à Embrapa pelo suporte concedido para a realização do trabalho. Ao pesquisador Pedro Abel Vieira, da Embrapa-SNT, e ao professor José Maria F. J. da Silveira, Instituto de Economia-UNICAMP, pelo auxílio e pelas contribuições que muito enriqueceram o trabalho. E, por fim, agradeço aos profissionais das usinas visitadas, Paraíso Bionergia, Usina Guaíra e Usina Rio Pardo, pela disponibilidade e atenção dedicadas.

## Referências

FIANI, R. Teoria dos custos de transação. In: KUPFER, D.; HASENCLEVER, L. (Org.).

**Economia Industrial**: fundamentos teóricos e práticas no Brasil. Rio de Janeiro:

Elsevier, 2002. 640 p.

SALOMÉ, J. R.; SAKAI, R. H.; AMBROSANO, E.; BUENO, J. Viabilidade econômica da rotação de adubos verdes com cana-de-açúcar. Resumos do V CBA – Desenvolvimento Rural. **Revista Brasileira de Agroecologia**, v. 2, n. 2, out. 2007.

SANTIAGO, A. D.; ROSSETTO, R. Produção manejo: rotação e reforma. In: **Agência de Informação Embrapa**: Cana-de-açúcar. Disponível em: <[http://www.agencia.cnptia.embrapa.br/gestor/cana-de-acucar/arvore/CONTAG01\\_75\\_22122006154841.html](http://www.agencia.cnptia.embrapa.br/gestor/cana-de-acucar/arvore/CONTAG01_75_22122006154841.html)>.

Acesso em: 28 nov. 2011.



# Parametrização de modelo compartimental do tipo SEIDR para avaliação quantitativa da HLB em citros

Guilherme Augusto Braga<sup>1</sup>

Sônia Ternes<sup>2</sup>

Raphael Gustavo D'Almeida Vilamiu<sup>3</sup>

O Brasil é um dos maiores produtores de citros do mundo, negócio que gera mais de 400 mil empregos, sendo o Estado de São Paulo responsável por 80% da produção nacional de laranja (BELASQUE JUNIOR et al., 2010). Devido a essa importância econômica, a presença de doenças como o Huanglongbing ou HLB (ex greening) em plantações gera enormes preocupações. O HLB é caracterizado pela presença da bactéria *Candidatus Liberibacter* spp., que é transmitida para a planta pelo psilídeo (*Diaphorina Citri*), inseto presente nos pomares brasileiros desde a década de 40 (FUNDECITRUS, 2011).

Com base nos trabalhos de Ross (BACAËR, 2010) e Vandermeer e Power (1990) para estudos da transmissão da malária em humanos e de doenças na cultura do milho, respectivamente, desenvolveu-se um modelo matemático compartimental determinístico para descrever a dinâmica da propagação do HLB no sistema citros - psilídeo *Diaphorina Citri*. Para tal, o modelo considera a fase de incubação da doença na planta, período desde a infecção da planta pelo psilídeo infectado, até o aparecimento de sintomas.

A Figura 1 apresenta os compartimentos considerados no modelo, onde  $C_0$  representa o compartimento de plantas saudáveis,  $C_1$  o de plantas em período de incubação,  $C_2$  o de plantas sintomáticas,  $I_0$  o de insetos saudáveis

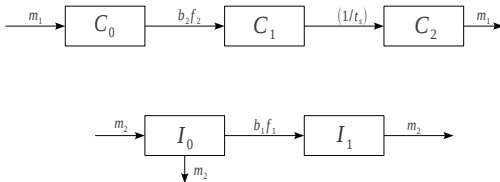
---

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

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

<sup>3</sup> Embrapa Informática Agropecuária, [vilamiu@gmail.com](mailto:vilamiu@gmail.com)

e  $I_1$  o compartimento de insetos infectados com HLB, que são descritos matematicamente pelo sistema de equações diferenciais em (1), já normalizadas.



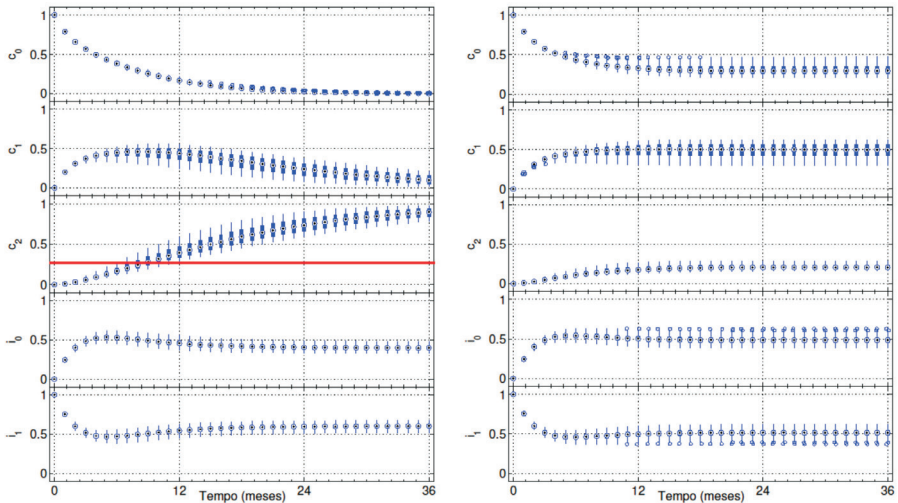
$$(1) \begin{cases} \frac{dc_1}{dt} = N_2 b_2 f_2 i_1 (1 - c_1 - c_2) - \frac{1}{t_s} c_1, \\ \frac{dc_2}{dt} = \frac{1}{t_s} c_1 - m_1 c_2, \\ \frac{di_1}{dt} = N_1 b_1 f_1 (1 - i_1)(c_1 + c_2) - m_2 i_1, \end{cases}$$

**Figura 1.** Compartimentos considerados no modelo matemático.

onde:  $N_1$  e  $N_2$  são as densidades de populações de plantas de citros e insetos por unidade de produção (UP), respectivamente (ambos assumidos constantes);  $c_1$  é a população de plantas em estado de incubação da doença;  $c_2$  é a população de plantas sintomáticas;  $i_1$  é a população de insetos infectados com HLB;  $b_1$  é a frequência de visitas que um psilídeo realiza por planta;  $b_2$  é a frequência de visitas que uma planta recebe por inseto;  $f_1$  é a probabilidade de aquisição de HLB pelo inseto;  $f_2$  é a probabilidade de transmissão de HLB do inseto para a planta;  $m_1$  é a taxa de remoção de plantas infectadas;  $m_2$  é a taxa de mortalidade do inseto;  $t_s$  é o tempo de incubação na planta.

Através de valores encontrados na literatura (BRAGA et al., 2011), foram realizadas 1000 simulações por LHS (Latin Hypercube Sampling), de forma a obter o comportamento geral da dinâmica para os valores dos parâmetros variáveis ( $f_1$ ,  $m_2$  e  $t_s$ ). O resultado pode ser verificado na Figura 2:

A análise da Figura 2 deixa clara a importância de se remover as plantas infectadas. Se a remoção não ocorre ( $m_1 = 0,0$ ), em menos de 2 anos mais de 95% da plantação estará infectada. Analisando o cenário onde  $m_1 = 0,2$ , verifica-se uma mudança drástica na dinâmica geral, uma vez que o ponto de equilíbrio médio será (0,295;0,495;0,210), ou seja, em média o equilíbrio será alcançado com 29,5% da população de plantas saudáveis, 49,5%



**Figura 2.** Resultado das dinâmicas geradas por LHS usando os parâmetros definidos e  $N_2 = 1000$  (Densidade populacional do inseto) e  $m_1 = 0,0$  (a) e  $m_1 = 0,2$  (b) (Taxa de remoção de plantas).

das plantas estarão infectadas em estado latente e 21% serão sintomáticas. Pela análise das simulações percebe-se que a dinâmica do patossistema está intrinsicamente influenciada pela taxa de remoção das plantas infectadas. Comparando com a Instrução Normativa nº 53 (BRASIL, 2008) que determina a eliminação das plantas da unidade de produção quando o resultado da análise das amostras laboratoriais for positivo e a UP apresentar mais de 28% de plantas sintomáticas, o modelo mostra que, quando  $c_2(t) = 0,28$ , a população de plantas sadias será de apenas 30%. Tal resultado pode ser inferido pela Figura 2a, onde a linha vermelha determina  $c_2(t) = 0,28$ .

## Referências

BACAËR, N. A short history of mathematical population dynamics. Bondy, France: Springer, 2010.

BELASQUE JUNIOR, J.; YAMAMOTO, P. T.; MIRANDA, M. P. D.; BASSANEZI, R. B.; AYRES, A. J.; BOVÉ, J. M. Controle do huanglongbing no estado de São Paulo, Brasil . Citrus Research & Technology, Cordeirópolis, v. 31, p. 53 - 63. 2010.

BRAGA, G. A.; TERNES, S.; VILAMIU, R. G. d'A.; CASTRO, A.; SILVA, M. V. da; LARANJEIRA, F. F. Modelagem matemática da dinâmica temporal do HLB. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 8., 2011, Bento Gonçalves. Anais... Florianópolis: UFSC; Pelotas: UFPel, 2011. Não paginado. SBIAgro 2011.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Instrução Normativa Nº 53, de 16 de outubro de 2008**. Disponível em: <<http://www.legisweb.com.br/legislacao/?legislacao=569432>>. Acesso em: 18 ago.2011.

FUNDECITRUS. **Controle da *Diaphorina citri***. Guia técnico, 2011.

VANDERMEER, J.; POWER, A. An epidemiological model of the corn stunt system in central america. **Ecological Modelling**, v. 52, p. 235 – 248,1990.

# Análise de custos de tecnologias da Embrapa Informática Agropecuária

Fernanda da Rocha Fernandes<sup>1</sup>  
Aryeverton Fortes de Oliveira<sup>2</sup>

A avaliação de impactos dos projetos de pesquisa da Unidade é baseada em um amplo conjunto de critérios. Dentre esses, foi sinalizado pela Secretaria de Gestão e Estratégia (SGE) que a análise de custos em 2010 deveria ser reforçada. A Empresa Brasileira de Pesquisa Agropecuária (Embrapa) define uma metodologia para levantamento das informações e cálculo dos custos, tendo sido aplicada com sucesso para estimar os custos das tecnologias Agrolivre, Diagnose Virtual, Agência e, em andamento, para o Ainfo. As informações podem ser também indicadores úteis para a gestão, definindo custos em novas iniciativas de pesquisa e transferência de tecnologia. O sistema de custos da Embrapa ainda é incapaz de recuperar informações históricas por projeto, e para superar essa limitação, em colaboração com outros empregados<sup>3</sup>, trabalhou-se na recuperação e sistematização das informações.

A forma de cálculo dos custos foi evoluída mantendo-se em acordo com a metodologia definida no manual da avaliação de impactos (AVILA et al., 2008), cujas principais orientações para uma estimativa dos custos de uma dada tecnologia são:

- a) custos de pessoal: referem-se à remuneração anual bruta mais encargos sociais do pessoal envolvido na geração e na transferência da tecnologia.

---

<sup>1</sup> Universidade Estadual de Campinas, fernandarf@cnptia.embrapa.br

<sup>2</sup> Embrapa Informática Agropecuária, ary@cnptia.embrapa.br

<sup>3</sup> Com destaque para os seguintes colaboradores: Adriana Farah Gonzalez; Aparecido Teodoro de Souza Filho; Helano Póvoas de Lima; Luciano Benedito Caldeira de Lima; Luiz Carlos do Nascimento; Marcos Dias Soares; Marcos Fernandes; Ronaldo de Carvalho Coutinho; Sílvia Maria Baú Idalgo; Sílvia Maria Fonseca Silveira Massruhá; Sílvio Roberto de Medeiros Evangelista; Suzilei Almeida Carneiro

- b) custeio da pesquisa: refere-se aos gastos anuais com a geração da tecnologia (exceto pessoal), estimados com base no orçamento dos projetos.
- c) depreciação de capital: corresponde à depreciação anual de todos os bens do centro de pesquisa, distribuída segundo a participação no esforço de pesquisa do centro.
- d) custos de administração: referem-se a uma parcela dos custos fixos atribuídos à tecnologia. Estão inclusos: o custeio com pessoal ligado à administração, os custos dos setores de campos experimentais e máquinas agrícolas e o custeio geral do centro (vigilância, limpeza, telefone, energia, xerografia, combustíveis, correio, etc.).
- e) custos de transferência tecnológica: consideram-se a elaboração de circulares, folderes, cursos, palestras, dias de campo, seminários, visitas, unidades de observação ou demonstrativas, etc.

Foi realizado um amplo levantamento de informações externas e junto aos setores de pesquisa e suporte, que envolveu: o Núcleo de Comunicação Organizacional (NCO), no fornecimento de informações de palestras e cursos, com indicação de horas de treinamento e técnicos envolvidos, folderes e materiais produzidos e adquiridos; o Setor de Gestão de Infraestrutura e Logística (SIL) e o site<sup>4</sup>, informaram valores de folderes e materiais produzidos para transferência de tecnologia, informações sobre compras dos projetos e patrimônio da Unidade; Área de Gestão de Pessoas (AGP), que disponibilizou dados históricos da folha de pagamentos e do quadro de funcionários, por consulta ao SIRH; a SGE, que repassou as informações sobre a depreciação de capital da unidade; o Setor de Gestão Orçamentária e Financeira (SOF), que repassou informações do Sistema Integrado de Administração Financeira (Siafi).

Foram entrevistados pesquisadores, analistas e assistentes que se envolveram nos projetos, a fim de definir seus níveis de dedicação por ano de desenvolvimento. Os valores de gastos e do quadro de funcionários nos agrupamentos de pesquisadores, analistas e assistentes foram levantados no SIRH, por ano, impressos, digitados em planilha (Figura 1.a), e serviram de referência para os custos com pessoal no desenvolvimento da tecnologia. Após definida a dedicação do grupo à tecno-

---

<sup>4</sup> Disponível em: < [www.comprasnet.gov.br](http://www.comprasnet.gov.br) >.

logia, foi calculado o total da força de trabalho da Embrapa Informática Agropecuária alocado ao projeto, empregado para rateio da depreciação do capital e do custeio geral da Unidade, além do cálculo de custos da hora de trabalho, para posterior cômputo dos custos com transferência de tecnologia.

Vale esclarecer que os custos atualmente disponíveis são agregados (a todo o centro, desde a sua criação), não existindo os custos desagregados por projeto, atividade ou tecnologia. Somente a partir da implantação do sistema de custos em andamento na Embrapa é que os custos desagregados estarão disponíveis.

Os custos gerais da Unidade foram calculados a partir de levantamentos de despesas registradas no Siafi, dos anos de 1996 a 2010, cujos relatórios anuais, por limitação do sistema tiveram que ser impressos, digitalizados e analisados em planilha. As contas identificadas com seis níveis, ou seja, mais desagregadas possíveis, foram classificadas segundo a pertinência para rateio. Adotou-se como critério que contas com baixa correspondência com os gastos realizados por projetos de pesquisa da unidade seriam rateadas integralmente. Contas com alguma sobreposição com despesas executadas por projetos foram rateadas em 50% de seu valor. Contas associadas com despesas dos projetos foram desconsideradas no rateio, em função de já estarem presentes no custeio geral dos projetos. Os valores associados com despesas gerais da Unidade foram somados e rateados para o projeto em função do tempo de dedicação da equipe no desenvolvimento da tecnologia em comparação com a força de trabalho total da Unidade.

Houve também o esforço em resgatar informações sobre a dedicação dos técnicos para os projetos de pesquisa e em quantificar com maior exatidão os custos que os envolve. Atualmente há uma planilha que sumariza as informações e a equipe estuda a utilização de um sistema de registro e análise de dados para armazenar e disponibilizar informações, via SAS. Os relatórios de impactos estão disponibilizados no endereço interno<sup>5</sup>.

Portanto, o trabalho de recuperação e sistematização de informações referentes aos custos contribuiu para aprimorar as atividades de avaliação de impactos das tecnologias geradas na Empresa (Figura 1.b). Destaca-

---

<sup>5</sup> Disponível em: <<http://intranet.cnptia.embrapa.br/tt/spat/relatorios>>.

a)

The screenshot shows a detailed financial planilha (spreadsheet) for the period 1997-2010. It is organized into columns for different agencies (e.g., Agência, Agência de Assistência Técnica, Agência de Assistência Social) and rows for each year. Each cell contains numerical data representing financial values. The spreadsheet includes various formulas and data validation rules, as indicated by the status bar at the bottom.

b)

**8. CUSTOS DA TECNOLOGIA**

**8.1 - Estimativa dos Custos**

**Tabela 8.1.1. - Estimativa dos custos**

Ano	Custos de Pessoa	Outros Custos	Depreciação de Capitais	Custos de Administração	Custos de Transfereência Tecnológica	Total
1997	54.499,21	7.201,74	6.999,02	18.858,83	0	87.358,80
1998	110.009,50	9.222,03	22.203,66	45.829,80	0	187.264,78
1999	104.197,28	8.414,61	23.107,45	37.696,56	0	173.415,90
2000	51.080,98	6.182,03	6.544,69	18.786,18	0	82.573,87
2001	24.138,71	5.775,17	6.379,48	18.483,85	0	54.777,19
2002	24.122,50	5.335,96	2.235,66	6.802,47	0	38.546,59
2003	21.319,76	8.666,74	6.316,37	14.416,45	0	48.719,32
2004	31.244,28	7.493,17	6.274,38	17.814,44	0	62.826,27
2005	68.173,28	11.835,54	34.995,17	37.707,13	0	152.711,12
2006	62.025,62	6.982,71	5.117,56	15.886,56	0	90.012,45
2007	156.502,89	14.295,17	14.370,60	44.702,93	160,34	230.031,93
2008	136.447,18	47.702,30	11.018,62	37.875,26	3153,71	235.995,26
2009	53.943,99	45.986,33	3.164,07	11.964,98	402,41	115.461,78
2010	56.781,44	43.556,92	2.991,46	11.408,02	400,00	115.137,83

**8.2 - Análise dos Custos**

Dado que a menor nota que a avaliação de impactos recebeu foi no quesito custos, em 2009 e 2010 o esforço foi concentrado nesta fragilidade. O sistema de custos da Embrapa ainda é incapaz de recuperar informações históricas. Para superar esta limitação, o trabalho de recuperação e sistematização de informações foi desenvolvido em colaboração com um grande número de funcionários, com destaque para os seguintes colaboradores: Adriana Farah-Gonzalez; Aparecido Teodoro de Souza Filho; Helano Póvoas de Lima; Luciano Benedito Caldeira de Lima; Luiz Carlos do Nascimento; Marcos Dias Soares; Marcos Fernandes; Ronaldo de Carvalho Coutinho; Silvia Maria Baú Idalgo; Silvia Maria Fonseca Silveira-Massurah; Silvio Roberto de Medeiros Evangelista; Suzilei Almeida Carneiro.

Os custos das tecnologias foram melhorados e calculados conforme metodologia definida no manual da avaliação de impactos. Foi realizado um amplo levantamento

Figura 1. a) Planilha para armazenamento e b) cálculo dos custos e compilação de resultados no relatório de impactos.



se ainda que tais avaliações tornaram-se um importante documento orientador, principalmente em nível institucional, sendo útil também para demonstrar, à sociedade, os impactos dos investimentos realizados na Empresa.

## Referências

AVILA, A. F. D. (Ed.); RODRIGUES, G. S. (Ed.); VEDOVOTO, G. L. (Ed.). **Avaliação dos impactos de tecnologias geradas pela Embrapa: metodologia de referência.** Brasília, DF : Embrapa Informação Tecnológica , 2008. 189 p.

# Relação entre a radiação fotossinteticamente ativa e a radiação global para diferentes classificações de céu

Fernanda Ferraz Brocardi<sup>1</sup>

Giuliana Amalfi Pinto<sup>1</sup>

Gustavo Costa Rodrigues<sup>2</sup>

## Introdução

Medições da Radiação Fotossinteticamente Ativa (RFA) são importantes para diferentes áreas relacionadas à fisiologia vegetal e às ciências agrárias, pois estão diretamente associadas à produção de biomassa, à morfologia e ao crescimento das plantas (CARTER; KLINKA, 1992; ACKERLY E BAZZAZ, 1995), sendo também essenciais em estudos ecológicos (BROCK, 1981). Esta fração da Radiação Global (Rg), compreendida na faixa espectral do visível (0,40 a 0,70  $\mu\text{m}$ ), é responsável pela excitação das moléculas de clorofila, fornecendo energia para o processo fotossintético. Entretanto, dados da RFA não são disponíveis em larga escala, sendo, em geral, estimados a partir da radiação solar global.

O Projeto Sistema de Organização Nacional de Dados Ambientais (Sonda) para o setor de energia foi desenvolvido pelo Centro de Previsão do Tempo e Estudos Climáticos (CPTEC) do Instituto Nacional de Pesquisas Espaciais (Inpe) e envolve uma rede de coleta de dados destinada a levantar e melhorar a base de dados sobre os recursos de energia solar e eólica no Brasil. Algumas dessas estações dispõem de sensores para a coleta de dados de componentes da radiação solar (difusa, direta e global) e da RFA.

---

<sup>1</sup> UNICAMP, [fernanda.brocardi@gmail.com](mailto:fernanda.brocardi@gmail.com), [giu.amalfi@gmail.com](mailto:giu.amalfi@gmail.com)

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

No presente trabalho, dados de vários anos da estação do projeto Sonda, localizada em Planaltina, DF, foram utilizados com o objetivo de se obter a relação entre a  $R_g$  e a RFA para diferentes tipos de céu, caracterizados por índices de claridade e brilho.

## Material e métodos

Foram analisadas as médias horárias de 2005 a 2010 da estação Brasília do projeto Sonda, localizada em Planaltina, DF (15° 36' 03" S, 47° 42' 47" O, 1023m acima do nível do mar).

Para a classificação dos dias usou-se as relações:

$$(1) \quad \epsilon = \frac{R_d + R_b}{R_d} \quad (2) \quad \Delta = \frac{R_d}{R_{0n} \cdot \text{sen} h}$$

sendo  $\epsilon$  o índice de nitidez do céu,  $R_d$  a radiações difusa,  $R_b$  a radiação direta,  $\Delta$  o índice de claridade do céu,  $R_{0n}$  radiação extraterrestre (1360W/m<sub>2</sub>) e  $h$  o ângulo de elevação solar. Foram considerados apenas valores de  $h$  superiores a 5°.

O ângulo de elevação solar foi obtido do Measurement and Instrumentation Data Center (MIDC..., 2011).

As horas de radiação no período avaliado foram agrupadas de acordo com os índices  $\epsilon$  e  $\Delta$ , sendo:

- Nublado:  $\epsilon < 1,2$ ;  $\Delta < 0,1$
- Parcialmente nublado:  $\epsilon < 1,2$ ;  $0,1 < \Delta < 0,3$
- Intermediário:  $1,2 < \epsilon < 5,2$ ;  $0,2 < \Delta < 0,3$
- Claro:  $\epsilon > 5,2$ ;  $\Delta < 0,1$
- Muito claro:  $\epsilon > 6,0$

Para cada condição foi estimada a relação entre a RFA e a radiação global ( $R_g$ ) utilizando-se um modelo linear.

## Resultados e discussão

Na tabela 1 é mostrada a relação entre RFA e  $R_g$  para as diferentes condições de céu. Observa-se tendência da fração da RFA ser maior em dias nublados, sendo assim um componente importante da radiação difusa.

**Tabela 1.** Relação entre a radiação fotossinteticamente ativa (RFA) e  $R_g$  para Planaltina, DF, sob diferentes condições do céu.

Condição do Céu	RFA/ $R_g$
Nublado	2,40
Parcialmente nublado	2,21
Intermediário	2,11
Claro	2,07
Muito Claro	2,07

## Conclusão

A média para todos os tipos de céu observada no presente trabalho (2,08) é compatível com os resultados obtidos em trabalhos realizados em diferentes partes do mundo e pode ser usada na estimativa da RFA incidente em superfícies vegetais quando somente os dados de  $R_g$  estiverem disponíveis.

## Referências

ACKERLY, D.D.; BAZZAZ, F.A. Seedling crown orientation and interception of diffuse radiation in tropical forest gaps: ecology: New York, v. 76, p. 1134-1146, 1995.

BROCK, T.D. Calculating solar radiation for ecological studies. **Ecological Modelling**, Amsterdam, v. 14, p. 1-19, 1981.

CARTER, R.E.; KLINKA, K. Variation in shade tolerance of Douglas-fir, western hemlock, and western red cedar in coastal British Columbia. **Forest Ecology and Management**, Amsterdam, v. 55, p. 87-105, 1992.

INPE - **Projeto Sonda**: Rede Nacional de estações para coleta de dados meteorológicos aplicados ao setor de energia, 2007.

MDIC: SPA calculator. Disponível em: < [www.nrel.gov/midc/solpos/spa.html](http://www.nrel.gov/midc/solpos/spa.html) >. Acesso em: 1 dez. 2011.



---

*Informática Agropecuária*

Ministério da  
Agricultura, Pecuária  
e Abastecimento



CGPE 9797