

O projeto WebMAPS: desafios e resultados

Carla Geovana do N. Macário^{1 2}, Claudia Bauzer Medeiros¹,
Rodrigo Dias Arruda Senra¹

¹ Instituto de Computação
Universidade Estadual de Campinas
Caixa Postal 6176 – 13084-971 – Campinas/SP – Brasil

²Embrapa Informática Agropecuária
Caixa Postal 6041 – 13083-886 – Campinas/SP – Brasil

{carlamac,cmbm}@ic.unicamp.br, rsenra@acm.org

Abstract. *This paper describes challenges and results of WebMAPS, a multi-disciplinary project under development at UNICAMP. Its goal is to develop a platform based on Web Services for agro-environmental planning. It requires state of the art research in specification and implementation of software that relies on several kinds of distributed information - satellite images, data from sensors and from agricultural production and geographic data.*

Resumo. *Este trabalho descreve desafios e resultados do projeto WebMAPS, um esforço multidisciplinar envolvendo ciências agrárias e de computação, em desenvolvimento na UNICAMP. Seu objetivo é desenvolver uma plataforma baseada em serviços Web para o planejamento agro-ambiental. Requer pesquisa de ponta voltada à especificação e à implementação de software com acesso a vários tipos de informação distribuída - imagens de satélite, dados provenientes de sensores, dados de produção agrícola e dados geográficos.*

1. Introdução

A agricultura é uma atividade de destaque na economia brasileira, contribuindo significativamente para o PIB brasileiro. Em 2005, o PIB atingiu R\$ 1.929 bilhões, sendo R\$ 537 bilhões provenientes de atividades agrícolas, o que corresponde a quase 30% do montante brasileiro. Diversos fatores contribuem para estes números. Um deles é a disponibilidade cada vez maior de sistemas que auxiliam o planejamento e o gerenciamento da produção.

Pesquisas em Ciências Agrárias geram informação essencial para a agricultura brasileira. No entanto, decisões sobre o que plantar (e quando, onde e como) exigem acesso confiável a dados e informação atualizada. Além disso, há necessidades de desenvolvimento de modelos sofisticados, requerendo cooperação entre especialistas do domínio e cientistas da Computação. Portanto, soluções que ajudem a aquisição, o processamento e a disseminação de dados em tempo hábil são cada vez mais necessárias. O WebMAPS (Sistema baseado na WEB Semântica para Monitoramento Agrícola e Previsão de Safras) visa contribuir nesse sentido, dando apoio ao processamento de dados científicos heterogêneos provenientes de diversas fontes. Seu objetivo final é o estabelecimento de uma plataforma base para a formulação, implementação e avaliação

de políticas integradas de planejamento agrícola. O projeto, que iniciou em 2003 apoiado por um edital Universal do CNPq, envolve pesquisadores de três unidades da UNICAMP - do Instituto de Computação, da Faculdade de Engenharia Agrícola e do CEPAGRI [Medeiros et al. 2006]. Este trabalho descreve alguns dos desafios e resultados do projeto obtidos até o momento, estando organizado da seguinte forma: a seção 2 descreve os principais tópicos de pesquisa envolvidos no projeto. A seção 3 aborda a metodologia usada no desenvolvimento do sistema, focando na implementação do protótipo disponível na web. Algumas das iniciativas existentes nas mesmas áreas de pesquisa do projeto encontram-se na seção 4. Por fim, a seção 5 apresenta conclusões e atividades em desenvolvimento.

2. Alguns Aspectos de Pesquisa Envolvidos

A especificação e o desenvolvimento do WebMAPS envolve pesquisa multidisciplinar em Ciências da Computação e Ciências Agrárias. Do lado computacional, estão sendo estudados pontos em aberto nas áreas de grandes bancos de dados e algoritmos para manipulá-los, engenharia de software de grandes sistemas, processamento digital de imagens e interfaces multimodais. Do lado de Ciências Agrárias, os estudos vêm envolvendo aspectos de sensoriamento remoto, desenvolvimento de novas metodologias de previsão de safra, especificação de novos métodos e ferramentas de apoio a decisão no domínio agrícola, dentre outros. A combinação dos resultados destas duas áreas é o principal diferencial do projeto.

Na área de Ciência da Computação vários pontos vêm sendo estudados. Um dos tópicos centrais são sistemas de bancos de dados espaço-temporais, para integrar informação extraída de imagens de satélite a dados climatológicos de modo a permitir fazer associações *ad-hoc* entre eles, para regiões e períodos arbitrários. Outros dois desafios são a especificação e implementação de algoritmos que possibilitem estabelecer correlações entre séries temporais e sua evolução, a partir do uso de novos resultados na área de casamento de padrões, e a especificação e desenvolvimento de algoritmos para processamento das imagens de satélite, a partir de características de conteúdo, visando busca por padrões temporais. Ainda outro tópico, o gerenciamento da rastreabilidade de produtos agrícolas, envolvendo dentre várias, técnicas de *workflow* e interoperabilidade de serviços Web. Finalmente, estão sendo levados em conta aspectos de *design* e implementação de interfaces que permitam diferentes modos de visualização e interação com os dados armazenados, visando facilitar consultas interativas sobre evolução espaço-temporal dos dados.

Em Ciências Agrárias, a ênfase tem sido dada a aspectos de sensoriamento remoto, envolvendo séries de imagens de satélite [Lunetta et al. 2003]. Essas imagens podem ser apresentadas em bandas individuais ou em índices de vegetação, como o NDVI (*Normalized Difference Vegetation Index*), que é calculado pela diferença de valores de reflectância das bandas Infravermelho (IV) e Vermelho (V) e normalizado pela soma dos valores de reflectância destas duas bandas. O índice permite avaliar as condições da biomassa de uma cultura, podendo ser usado para monitoramento e previsão de safra. Uma curva de valores NDVI para uma região em um período pode ser comparada com curvas de outros períodos (para determinar o comportamento da biomassa) e também com outras curvas do mesmo período mas de outras regiões. Estes dados podem também ser comparados com séries históricas de temperatura, de chuva ou de outros dados meteorológicos. Estão

previstas famílias de ferramentas para: (a) manipulação de dados de cadastro e produção; (b) manipulação de dados de sensores; (c) manipulação de imagens de satélite; (d) uso de *workflows* para modelagem de processos. A idéia é disponibilizar produtos que sejam gerados por tais ferramentas, mas também permitir a execução *on-line*, combinada, de um ou mais módulos.

3. Aspectos de Implementação

O WEBMaps vem sendo desenvolvido combinando prototipação rápida com a disponibilização de ferramentas e produtos aos especialistas da área, para obter *feedback* quanto ao uso dos módulos desenvolvidos. Tal *feedback* vem possibilitando aperfeiçoar o levantamento de requisitos e a especificação de módulos adicionais. A análise de requisitos foi guiada pela Semiótica Organizacional [Liu et al. 2007, Schimiguel et al. 2005, Prado et al. 2000] e o desenvolvimento do projeto envolve metodologias de teste de desempenho [Torres-Zenteno et al. 2006]. A maior parte da implementação vem utilizando a tecnologia J2EE e o servidor de aplicações Zope, adotando um modelo de arquitetura multi-camadas. Dados são armazenados no PostGreSQL, prevendo-se sua migração para PostGIS. A variação espaço-temporal dos dados ainda está em discussão, podendo, por exemplo, vir a usar a proposta de [Faria 1998]. Para geração dos perfis de NDVI estão sendo utilizadas imagens do sensor MODIS (*Moderate Resolution Imaging Spectroradiometer*) com resoluções: espacial de 250 metros, temporal diária, radiométrica 16 bits e espectral de 36 bandas.

Algumas destas ferramentas já estão disponíveis [Macário et al. 2007]. Permitem o cadastro de propriedades, talhões e culturas, além de algumas consultas sobre tais entidades (figura 1). A figura 2 ilustra o resultado de uma consulta *on-line* ao sistema, na Web. Ela mostra a evolução, ao longo do tempo, dos valores NDVI de uma região selecionada pelo usuário. Este exemplo, cujos parâmetros foram as coordenadas da região e um intervalo de tempo (jan. 2001 a jan. 2003), usou como base um conjunto de 100 imagens do sensor MODIS das regiões Sul e Sudeste do Brasil.

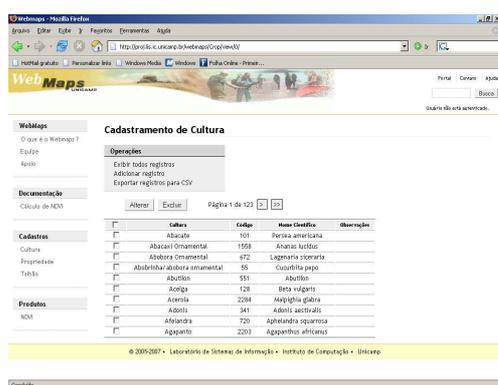


Figura 1. Cadastro de Culturas

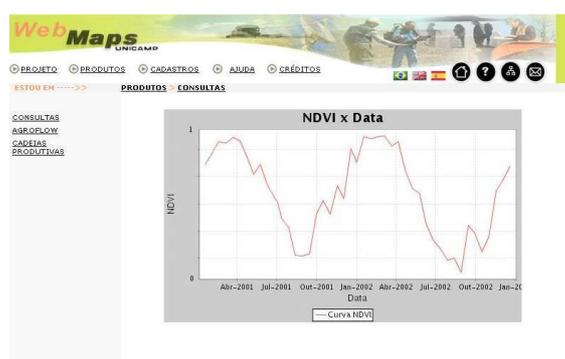


Figura 2. Resultado de Consulta

Para o usuário, a curva indica o comportamento da cultura (no caso, cana) naquela região ao longo daquele período. Este comportamento pode ser relacionado a coletas obtidas para fenômenos associados (chuva, temperatura) e à produtividade da cultura para a mesma região e período. De posse desse gráfico, o usuário pode comparar o comportamento de sua lavoura com anos anteriores que tiveram comportamento semelhante,

permitindo a definição de melhor época para realizar a colheita e a previsão da qualidade da cana produzida e de quanto será sua produção.

Do ponto de vista da Computação, a cópia de tela corresponde à materialização, via código, de vários aspectos de pesquisa de ponta, ressaltando-se: integração de pesquisa em processamento de imagens de satélite e bancos de dados geográficos; implementação de novos métodos de teste de software para a Web, ou projeto de interface de software geográfico para a Web.

Ainda outros resultados do trabalho incluem modelagem de transações em cadeias produtivas [Bacarin et al. 2004] e o desenvolvimento de novos descritores de curvas obtidas a partir de séries históricas de chuva e temperatura, que facilitam a busca por similaridade [Mariote et al. 2007]. Finalmente, o trabalho de [Kondo et al. 2007] especificou e desenvolveu um conjunto de serviços web para rastreabilidade de eventos e produtos em cadeias produtivas agrícolas, usando como base *workflows* da cadeia do leite. Usuários enviam requisições a um serviço que intermedia interações com outros serviços dedicados, responsáveis por encapsular o acesso a diferentes repositórios de registros de eventos. Este conjunto de serviços permite executar uma grande gama de consultas que combinem eventos de uma cadeia agrícola, facilitando gerenciar a qualidade de seus processos e produtos.

4. Trabalhos Correlatos

O acesso confiável a dados na Web é essencial para a geração de informação de apoio a políticas agrícolas. Apesar de iniciativas que estão surgindo nos EUA [Jakubauskas et al. 2001] e na Europa [Lemoine and Kidd 1998], há ainda uma enorme carência de ferramentas que agilizem a produção da informação estratégica. As iniciativas internacionais são centradas em pesquisa em Ciências Agrárias, sendo a Computação usada principalmente como fornecedora de infra-estrutura.

A especificação e o desenvolvimento de sistemas como o WebMAPS são alvo de pesquisa multidisciplinar em nível mundial. Do lado da Computação, há demanda por novos resultados em: bancos de dados e serviços Web [Lockemann et al. 1997] (gerenciamento de fontes de dados volumosos e heterogêneos contendo imagens de satélite, mapas digitais, dados climáticos, econômicos e outros); processamento de imagens [Beucher and Meyer 1993, Soile 1996] (algoritmos específicos para segmentação e recuperação por conteúdo); engenharia de software (especificação, testes e desenvolvimento do software), redes e sistemas distribuídos [Fenstermacher and Ginsburg 2002] (para integrar e processar os dados e ferramentas e disponibilizá-los na Web); e interfaces [Kosch 2002] (contemplando múltiplos tipos e propósitos de interação). Do lado da pesquisa em Ciências Agrárias (domínio alvo), há problemas na identificação dos dados, no procedimento de amostragem, no uso de sensoriamento remoto e no aperfeiçoamento de algoritmos para análise e visualização de informação. Outros problemas a serem abordados consideram, também, processamento de séries temporais [Muthukrishnan et al. 2004, Wu et al. 2005], rastreabilidade de processos e cadeias produtivas [Roder and Tibken 2006] e ontologias [Hochmair 2005].

5. Conclusões e Trabalhos em Andamento

O WebMAPS é motivado por problemas em aberto na geração de informação agrícola estratégica para a tomada de decisões tanto no âmbito governamental (apoio a políticas públicas) quanto no de cooperativas agrícolas e agronegócios. Há vários diferenciais deste projeto em relação a iniciativas internacionais de natureza semelhante: a ênfase em pesquisa multidisciplinar em Computação aplicada a Ciências Agrárias - na maioria das iniciativas, a pesquisa é centrada no último domínio e apenas utiliza recursos da Tecnologia da Informação; e a adequação ao contexto geográfico brasileiro, com forte ênfase em dados obtidos via sensores; utilização de novos resultados de pesquisa sobre a Web Semântica; exploração, em tempo real, do conteúdo de imagens, e não apenas de dados textuais; consideração de aspectos de interação humano-computador.

Há vários trabalhos em andamento. Um deles está voltado à disponibilização de ferramentas de projeto e execução de *workflows* para atividades em agricultura. Outras iniciativas envolvem desenvolvimento de consultas por conteúdo a imagens de satélite, modelos baseados em curvas NDVI para identificação de culturas e mecanismos para melhorar a interoperabilidade semântica de dados geoespaciais.

Referências

- Bacarin, E., Medeiros, C. B., and Madeira, E. (2004). A collaborative model for agricultural supply chains. In *OTM Confederated International Conferences, CoopIS, DOA, and ODBASE 2004*, number 3290 in LNCS, pages 319–336.
- Beucher, S. and Meyer, F. (1993). The morphological approach to segmentation: The watershed transformation. In *In Proc. of the International Symposium on Mathematical Morphology*, pages 433–481.
- Faria, G. (1998). Um banco de dados espaço-temporal para desenvolvimento de aplicações em sistemas de informação geográfica. Master's thesis, Instituto de Computação - Universidade Estadual de Campinas.
- Fenstermacher, K. and Ginsburg, M. (2002). A lightweight framework for cross-application user monitoring. *IEEE Computer*, 35(3):51–59.
- Hochmair, H. H. (2005). Ontology matching for spatial data retrieval from internet portals. In *First International Conference GeoSpatial Semantics, GeoS*, pages 166–182.
- Jakubauskas, M., Legates, D., and Kastens., J. (2001). Harmonic analysis of time-series avhrr ndvi data. *Photogrammetric Engineering and Remote Sensing*, 67(4):461–470.
- Kondo, A. A., Medeiros, C. M., Bacarin, E., and Madeira, E. R. M. (2007). Traceability in food for supply chains. In *3rd International Conference on Web Information Systems and Technologies (WEBIST)*, pages 121–127, Barcelona, Spain.
- Kosch, H. (2002). MPEG and multimedia database systems. *ACM SIGMOD Record*, 31(2):34–39.
- Lemoine, G. and Kidd, R. (1998). *Operational European Cereal Monitoring: Methodological Consideration*. Ispra, Italy. Disponível em <<http://conferences.esa.int/98c07/papers/P089.pdf>>. Acesso em jul, 2007.
- Liu, K., S.Y.Liao, and Chong, S. (2007). *Semiotics for Information Systems Engineering – re-use of high-level artefacts*. Disponível em:

<<http://www.scit.wlv.ac.uk/jphb/cp4040/rolandonotes/CSNDSP2002/Papers/A1/A1.1.pdf>>. Acesso em jul, 2007.

- Lockemann, P., Kolsch, U., Koschel, A., Kramer, R., Nicolai, R., Wallrath, M., and Walter, H. (1997). The network as a global database: Challenges of interoperability, proactivity, interactiveness, legacy. In *23rd VLDB Conference*, pages 239–250.
- Lunetta, R., Johnson, D., Lyon, J., and J., C. (2003). Impacts of imagery temporal frequency on landcover change detection monitoring. *Remote Sensing and Environment*, 89(4):444–454.
- Macário, C. G. N., Senra, R. D. A., Medeiros, C. B., Lamparelli, R. A. C., Júnior, J. Z., Rocha, J. V., Madeira, E. R. M., Martins, E., Baranauskas, M. C. C., Leite, N. J., and Torres, R. S. (2007). Monitoramento de safras via web: Um caso de sucesso em pesquisa multidisciplinar. In *60. Congresso Brasileiro de Agroinformática - SBIAgro 2007*. Aceito para publicação.
- Mariote, L. E., Medeiros, C. M. B., and Torres, R. S. (2007). Diagnosing similarity of oscillation trends in time series. In *International Workshop on spatial and spatio-temporal data mining - (SSTDM'07)*, Omaha, USA.
- Medeiros et al., C. M. B. (2006). *WebMAPS II - Sistema baseado na WEB Semântica para Monitoramento Agrícola e Previsão de Safras*. Projeto Universal CNPq. Iniciado em 2003 - renovado em 2006.
- Muthukrishnan, S., Shah, R., and Vitter., J. S. (2004). Mining deviants in time series data streams. In *6th International Conference on Scientific and Statistical Database Management*.
- Prado, A., Baranauskas, M., and Medeiros, C. (2000). Cartography and geographic information system as semiotic systems: A comparative analysis. In *8th ACM Symposium on Advances in Geographic Information System*, pages 161–166.
- Roder, A. and Tibken, B. (2006). A methodology for modeling inter-company supply chains and for evaluating a method of integrated product and process documentation. *European Journal of Operational Research*, 169(3):1010–1029.
- Schimiguel, J., Baranauskas, M. C. C., and Medeiros., C. B. (2005). Usabilidade de aplicações SIG Web na perspectiva do usuário: um estudo de caso. In *VI Simpósio Brasileiro de Geoinformática - GEOINFO*, Campos do Jordão, SP.
- Soile, P. (1996). Morphological partitioning of multispectral images. *Journal of Electronic Imaging*, 5(3):252–265.
- Torres-Zenteno, A., Martins, E., Torres, R. S., and Cuaresma, M. J. E. (2006). Teste de desempenho em aplicacoes sig web. In *IX Workshop Iberoamericano de Ingenieria de Requisitos y Ambientes de Software- IDEAS*, La Plata, Argentina.
- Wu, H., Salzberg, B., Sharp, G., Jiang, S. B., Shirato, H., and Kaeli., D. (2005). Subsequence matching on structured time series data. In *ACM SIGMOD Conference on Management of data.*, pages 682–693.