

UMA ARQUITETURA PARA SISTEMA DE INFORMAÇÃO SOBRE SOLOS VOLTADA PARA O ZONEAMENTO AGRÍCOLA

RENATO FILETO¹
MARIA LEONOR RIBEIRO CASIMIRO LOPES ASSAD²
JOÃO DOS SANTOS VILLA DA SILVA¹
AMARINDO FAUSTO SOARES¹
LAURIMAR GONÇALVES VENDRUSCULO¹

RESUMO

A disponibilidade de informação sobre solos é fundamental para a tomada de decisão no agronegócio. Para cobrir tal demanda, os dados devem ser coletados de diversas fontes, integrados e processados por ferramentas de análise de informação e extração de conhecimento, de modo que os resultados possam ser facilmente assimilados. Este artigo descreve a arquitetura geral do Agrissolos, um sistema de informação sobre solos voltado para as necessidades agrícolas. Esta arquitetura facilita o acesso à informação e a evolução contínua do conhecimento sobre os solos brasileiros, segundo o ciclo: coleta de dados, análise da informação, extração de conhecimento e avaliação dos resultados.

PALAVRAS CHAVES: Solos do Brasil, banco de dados, ontologias, mapas de solos.

A SOILS INFORMATION SYSTEM ARCHITECTURE FOR AGRICULTURAL ZONING

ABSTRACT

The availability of information about the soils is central for decision making in the agribusiness. In order to address this demand, data must be collected from several sources, integrated and processed by tools for information analysis and knowledge extraction, so that the results can be easily assimilated. This paper describes the general architecture of Agrissolos, an information system intended to the agricultural needs. This architecture facilitates information access and continuous evolution of the knowledge about Brazilian soils, according to the cycle: data collecting, information analysis, knowledge extraction and results evaluation.

KEYWORDS: Brazilian soils, databases, ontologies, maps of soils.

¹ Doutor em Ciência da Computação, Doutor em Engenharia Agrícola, Mestre em Agronomia e Mestre em Ciência da Computação, respectivamente - Embrapa Informática Agropecuária, Av. Dr. André Tosello, 209, Cidade Universitária (UNICAMP), C. P. 6041, CEP 13083-886, Campinas-SP.

² Doutora em Ciência do Solo - Departamento de Recursos Naturais e Proteção Ambiental (DRNPA), Universidade Federal de São Carlos (UFSCar), Rodovia Anhanguera, Km 174, CEP 13600-970, Araras-SP.

1. INTRODUÇÃO

O zoneamento agrícola é uma atividade de tomada de decisão que requer informação proveniente de várias esferas (e.g., clima, solos, relevo) para determinar as melhores opções de uso produtivo e sustentável das terras. Atualmente, uma das carências mais marcantes do zoneamento agrícola brasileiro é um sistema de informações sobre solos.

Autores distintos propõem diferentes arquiteturas para sistemas de suporte a decisão e variações dessas arquiteturas, de acordo com os objetivos do sistema, modo de funcionamento e área de aplicação. Todavia, a despeito dessas variações e discrepâncias de nomenclatura, todos reconhecem a necessidade de dois componentes básicos, além da interface com o usuário: o banco de dados e a base de conhecimento (POWER 2002, TURBAN 1995).

No zoneamento agrícola e particularmente em pedologia, assim como em outras áreas de aplicação onde a localização geográfica é atributo essencial, um outro componente fundamental é o sistema de informação geográfica, para gerenciar e manipular informação cartográfica. Assim, uma arquitetura de sistema de informação para tais aplicações deve contemplar também este componente.

Este artigo apresenta a arquitetura geral do Agrissolos, um sistema de informação que visa suprir o zoneamento agrícola e os processos de tomada de decisão na agricultura com informação consolidada e conhecimento útil sobre os solos brasileiros. A arquitetura do Agrissolos integra três componentes principais: o banco de dados, a base de conhecimento e o sistema de informação geográfica. As descrições desses componentes e os problemas envolvidos no processamento de dados de solos, contidos nesse artigo, visam contribuir para o estabelecimento de uma referência para canalizar esforços de coleta, gerenciamento, análise e utilização de informação sobre solos, visando subsidiar a tomada de decisão na agricultura.

O objetivo da publicação desta arquitetura é introduzir o Agrissolos na nossa comunidade e fomentar discussões com vistas a solucionar algumas questões relacionadas à construção e à operacionalização do sistema:

- verificar os dados de solos disponíveis atualmente em diferentes instituições;
- estabelecer parcerias para desenvolvimento e alimentação do sistema;
- definir estratégias tecnológicas para integração e análise de dados;
- dar prosseguimento aos trabalhos de implementação.

***V Congresso Brasileiro de Agroinformática, SBI-AGRO
Londrina, 28 a 30 de setembro de 2005***

A meta é obter um sistema de informação sobre solos que congregue os esforços de diversas instituições, disponibilizando publicamente via Web informações acuradas e consolidadas para subsidiar processos decisórios, em benefício de todos.

As seções subseqüentes deste artigo descrevem cada um dos componentes do Agrissolos individualmente e a arquitetura geral do sistema, discutindo alguns problemas envolvidos no processamento dos dados e delineando soluções sendo pesquisadas.

2. O BANCO DE DADOS DE PERFIS DE SOLOS

No Brasil, grande parte dos dados disponíveis sobre os solos foram obtidos por meio de levantamentos realizados por diversas instituições e projetos. Além desses mapeamentos, um grande número de informações de solos encontra-se disponível em trabalhos científicos publicados ao longo dos mais de cem anos de história da pesquisa agrícola no país.

Os dados de perfis e de horizontes formam o banco de dados de solos que contém descrições morfológicas de perfis de solos de diversos pontos do território nacional e medidas oriundas de análises laboratoriais de amostras de solos colhidas nesses perfis, que constituem a fonte de informação primária para aferir as características dos solos.

A Figura 1 apresenta o esquema (estrutura) do banco de dados. Um levantamento de solos gera uma série de perfis, cada qual com vários horizontes. Para cada horizonte são armazenados atributos morfológicos, químicos, físicos e mineralógicos. Esse conjunto de informações permite definir unidades taxonômicas (ou classes) e contribuem para o mapeamento de solos. Em geral, uma unidade taxonômica (classe de solo) não corresponde a apenas uma unidade cartográfica (ou unidade de mapeamento), pois esta depende de fatores de escala. Uma classe de solo pode participar de uma ou mais unidades de mapeamento.

O primeiro protótipo do banco de dados de perfis está sendo desenvolvido sobre o Microsoft Access, a fim de permitir o cadastramento dos dados localmente em microcomputadores de diferentes instituições, segundo o esquema apresentado na Figura 1, com manutenção de consistência e integridade. Paralelamente a este cadastro local, será desenvolvido um sistema mais potente e robusto sobre o PostgreSQL (POSTGRESQL 2005), para centralizar o gerenciamento dos dados de perfis de solos ao mesmo tempo em que se permite o acesso remoto, com os devidos recursos de segurança.

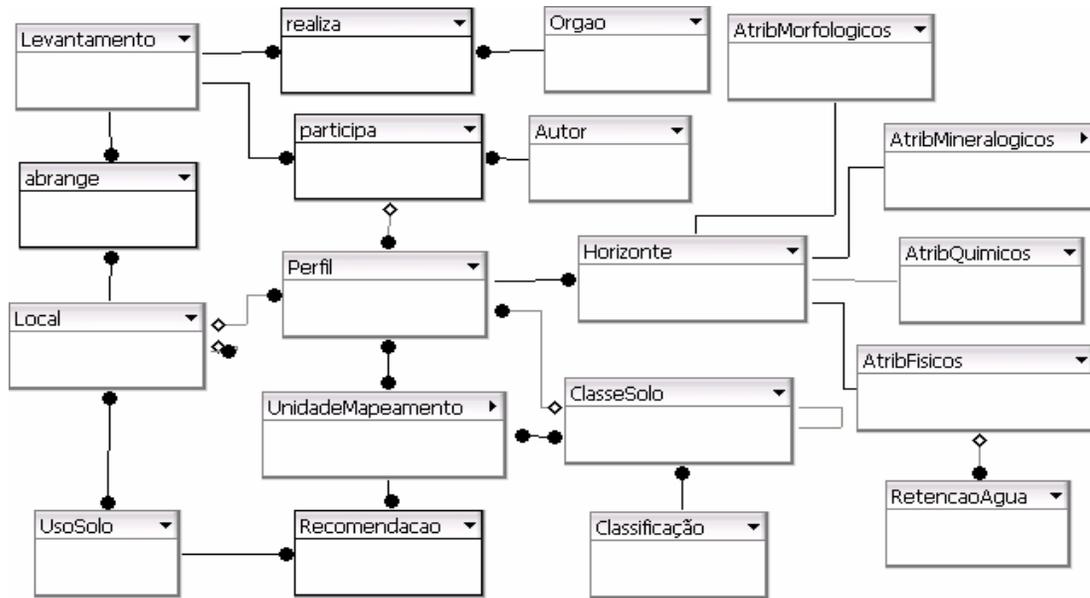


Figura 1: O esquema do banco de dados de perfis de solos

3. A BASE DE CONHECIMENTO

O Agrissolos utilizará técnicas computacionais de extração (HAN & KAMBER 2000), representação (BERTINO *et al.* 2001, USCHOLD & GRUNINGER 1996) e uso de conhecimento (BERNERS-LEE *et al.* 2002), para apoio à decisão na agricultura. A base de conhecimento é, portanto, um módulo central na arquitetura do Agrissolos, pois permite armazenar o conhecimento resultante do processamento do banco de dados de perfis, juntamente com o conhecimento inserido no sistema por especialistas dos domínios de pedologia e agricultura, mantendo a consistência, refinando gradativamente a base e promovendo o uso do conhecimento em processos decisórios.

Atualmente, dois esforços estão sendo empreendidos para o gerenciamento de conhecimento no Agrissolos:

1. Formalização de ontologias (USCHOLD & GRUNINGER 1996), sobre o Protégé (NOY *et al.* 2002), visando solucionar incompatibilidades entre classificações de solos de diferentes fontes de dados.
2. Utilização de técnicas de mineração de dados (HAN e KAMBER 2000), sobre o Weka (WITTEN & FRANK 2000), a fim de determinar classificações específicas para a agricultura, a partir dos dados do banco de perfis de solos e das classificações de solos conhecidas.

A elaboração de recomendações específicas para a agricultura considera aspectos como profundidade, textura e capacidade de retenção de água dos solos, entre outros.

A Figura 2 mostra uma tela de edição da ontologia relativa ao Sistema Brasileiro de Classificação de Solos (EMBRAPA, 1999) no Protégé (NOY *et al.* 2002). No canto superior esquerdo aparecem os níveis e fases deste sistema de classificação. No quadro superior central aparecem as ordens pedogenéticas, com a descrição da ordem *Neossolos* à direita e do grupo *Neossolos Quartzarênicos* na janela localizada no canto inferior esquerdo da Figura 2.

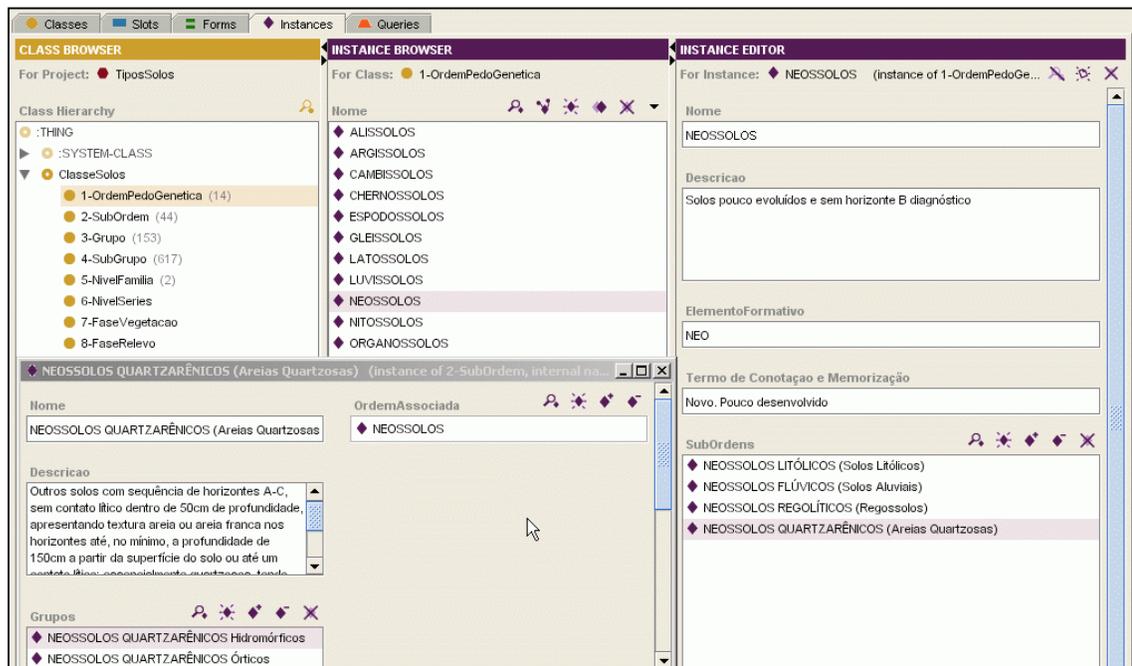


Figura 2: Uma ontologia de categorias de solos sendo editada no Protegé

4. O SISTEMA DE INFORMAÇÃO GEOGRÁFICA

O sistema de informação geográfica é o módulo que efetua o tratamento e a apresentação de dados cartográficos. Atualmente, as cartas do RadamBrasil estão sendo digitalizadas sobre o SPRING (INPE 2005), para permitir a manipulação eletrônica e a publicação dos mapas na Web. Uma das dificuldades encontradas no tratamento dos dados do RadamBrasil é a incompatibilidade de legendas de cartas publicadas em diferentes épocas. O sistema de classificação de solos foi evoluindo a longo do período em que essas cartas foram produzidas, sendo que cada carta segue o sistema de classificação vigente na época de sua publicação. Além disso, um novo Sistema Brasileiro de Classificação de Solos (SiBCS) foi

publicado posteriormente (EMBRAPA 1999). A situação atual da digitalização dos mapas no Agrissolos é relatada por SILVA *et al.* (2004). O problema de incompatibilidade de legendas e a abordagem atualmente adotada para tentar solucioná-lo são descritos por SOARES & SILVA (2005). A Figura 3 mostra um extrato de uma carta do RadamBrasil (à esquerda) e o resultado da sua digitalização no SPRING (à direita).

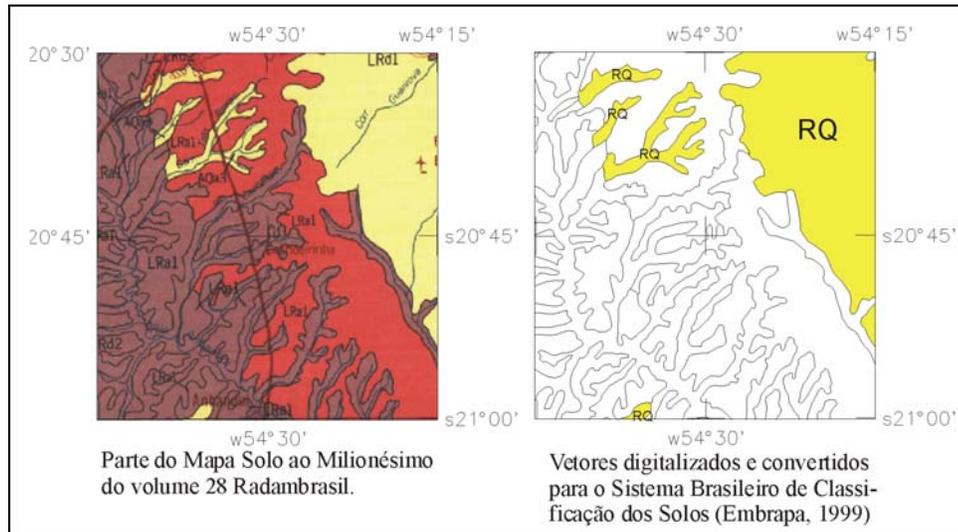


Figura 3: Mapas de solos do interior do Mato Grosso do Sul

5. A ARQUITETURA GERAL DO AGRISSOLOS

A arquitetura geral do Agrissolos contempla portanto três módulos básicos:

- (i) o banco de dados de perfis de solos, cobrindo todo o território brasileiro;
- (ii) a base de conhecimento, com descrições de solos para fins de uso agrícola, na forma de ontologias (representando, em formato computacionalmente processável, as classificações de solos) e árvores de decisão (representando os critérios adotados nas classificações e em tomadas de decisão);
- (iii) um sistema de informação geográfica para manipulação e apresentação de informação cartográfica.

A Figura 4 ilustra a arquitetura geral do Agrissolos. Esta arquitetura envolve, além dos módulos de armazenamento de dados descritos acima, alguns processos de carga e manipulação de dados, uma síntese da base de dados públicos e um portal para acesso via Web. O processo *Carga de Perfis* alimenta o banco de perfis de solos com dados de *Parceiros e Sistemas Legados*. O processo *Importação Cartográfica* carrega mapas no

V Congresso Brasileiro de Agroinformática, SBI-AGRO
Londrina, 28 a 30 de setembro de 2005

sistema de informação geográfica (SIG) via digitalização de *cartas* ou importação de informação de *outros SIGs*. O processo *Extração de Conhecimento*, por sua vez, consiste de uma coleção de ferramentas para extração e inserção de conhecimento. Periodicamente, a síntese do conteúdo dos repositórios é atualizada na base de dados públicos, com mecanismos de segurança e controle de acesso para publicação no portal Web. O fluxo de trabalho constituído por entrada de perfis e mapas, atualização da base de conhecimento, publicação e avaliação dos resultados é contínuo, de modo a possibilitar a evolução gradual do conteúdo do sistema.

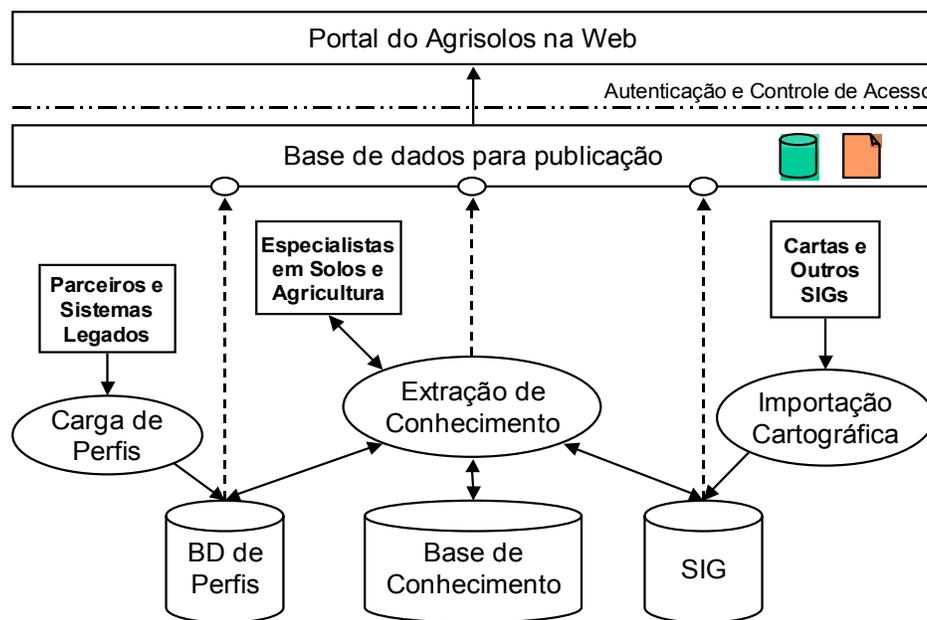


Figura 4: A arquitetura geral do Agrissolos

6. CONCLUSÕES

O Agrissolos é um sistema de informação sobre solos voltado especificamente para as necessidades da agricultura, particularmente do zoneamento agrícola. A arquitetura do Agrissolos, apresentada neste artigo, contempla três módulos principais além da interface Web: (i) um banco de dados de perfis de solos, cobrindo todo o território brasileiro; (ii) uma base de conhecimento, com descrições de solos para fins de uso agrícola, na forma de ontologias e árvores de decisão e (iii) um sistema de informação geográfica, para a manipulação de informação cartográfica e apresentação gráfica do conhecimento. Esses três módulos, integrados em um só sistema, vão possibilitar o fácil acesso à informação e a

evolução contínua do conhecimento sobre os solos brasileiros, segundo o ciclo: coleta de dados, análise da informação, extração de conhecimento e avaliação dos resultados. O desenvolvimento e a operacionalização do Agrissolos, segundo a arquitetura aqui apresentada, visa promover o uso e evolução do SiBCS e propiciar ao cidadão fácil acesso à informação e conhecimento públicos sobre solos, de maneira adequada para subsidiar suas decisões, promovendo assim a eficiência e a sustentabilidade das atividades agrícolas.

7. REFERÊNCIAS BIBLIOGRÁFICAS

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic Web. *Scientific American*, May 2001.

BERTINO, E.; ZARRI, G.; CATANIA, B. *Intelligent Database Systems*. Addison Wesley, 2001, 464p.

EMBRAPA. *Sistema Brasileiro de Classificação de Solos*, Empresa Brasileira de Pesquisa Agropecuária, Brasília, 1999.

HAN, J.; KAMBER, M. *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers, August 2000. 550p.

INPE. *SPRING: tutorial de geoprocessamento*. Disponível em <http://www.dpi.inpe.br/spring/portugues/tutorial/consulta.html>. Acesso em: 03.jan.2005.

NOY, N. F.; SINTEK, M.; DECKER, S.; CRUBEZY, M.; FERGERSON, R. W.; MUSEN, M. A. Creating semantic Web contents with Protégé-2000. *IEEE Intelligent Systems*, 16(2):60–71, 2002.

POSTGRESQL. <http://www.postgresql.org>. (as of May 2005).

POWER, D. J. *Decision Support Systems: Concepts and Resources for Managers*. Greenwood Publishing, 2002.

SILVA, J. S. V.; SOARES, A. F.; BEZERRA, H. S.; CALVE, L. Situação da base de solo brasileira na escala de 1:1.000.000. In: *Simposio Latinoamericano en Percepción Remota y Sistemas de Información Espacial en Santiago de Chile*, 11, Santiago, Chile, 2004.

SOARES, A. F.; SILVA, J. S. V. *Uniformização da legenda de solos do Brasil ao milionésimo*. Campinas: Embrapa Informática Agropecuária, 2005. 32p. (Documentos / Embrapa Informática Agropecuária; 49).

TURBAN, E. *Decision support and expert systems: management support systems*. Englewood Cliffs, N.J., Prentice Hall, 1995.

USCHOLD, M.; GRUNINGER, M. Ontologies: principles, methods, and applications. *Knowledge Engineering Review*, 11(2):93–155, 1996.

**V Congresso Brasileiro de Agroinformática, SBI-AGRO
Londrina, 28 a 30 de setembro de 2005**

WITTEN, H.; FRANK, E. Data Mining: *Practical machine learning tools with Java implementations*. Morgan Kaufmann, San Francisco, 2000.

**V Congresso Brasileiro de Agroinformática, SBI-AGRO
Londrina, 28 a 30 de setembro de 2005**