

Desenvolvimento do módulo de consulta CSV para o Sistema de Informação de Solos Brasileiros

Victor André Cetara Berti¹

Luciano Vieira Koenigkan²

Stanley Robson de Medeiros Oliveira²

O Sistema de Informação de Solos Brasileiros (OLIVEIRA et al., 2009) foi criado a partir de uma parceria entre a Embrapa Informática Agropecuária e a Embrapa Solos, e objetiva ser um repositório de dados com informações de perfis de solos coletados e analisados nas mais diversas regiões do Brasil. Esse sistema contempla informações organizadas em três grupos de dados: perfis de solos, informações sobre fertilidade e unidades de mapeamento, constituindo, dessa forma, uma importante fonte de informações para pesquisas, principalmente em agricultura.

As consultas originalmente disponíveis no sistema não ofereciam um método para que os dados fossem exportados, para sua utilização por outros sistemas de software. Tal fato permitiria a seus usuários, entre eles estudantes e pesquisadores, uma maior flexibilidade na utilização dos dados provenientes do banco de solos, ampliando sua possibilidade de utilização em projetos acadêmicos e científicos.

De forma a contribuir com essa flexibilização na utilização dos dados, este trabalho tem como objetivo implementar um módulo que permita a exportação dos dados para o formato de texto separado por vírgulas, *Comma Separated Value* (CSV), por ser este um formato que possibilita a fácil interoperabilidade com outros sistemas de software.

O Sistema de Informação de Solos Brasileiros (BD Solos), foi desenvolvido em linguagem PHP (PHP..., 2012) e a arquitetura é organizada de

¹ Faculdade de Tecnologia de Americana, victor.berti@colaborador.embrapa.br

² Embrapa Informática Agropecuária, {luciano.vieira, stanley.oliveira}@embrapa.br

forma que as consultas sejam realizadas em três etapas: “Seleção de Atributos”, “Seleção de Filtros” e “Seleção e exibição de Resultados”. Para tanto, o sistema é dotado de funcionalidades que permitem ao usuário realizar as consultas seguindo as etapas pré-estabelecidas.

Na primeira etapa os identificadores dos atributos selecionados para exibição são armazenados e encaminhados para a etapa seguinte. A segunda etapa consiste na seleção dos filtros que representam os critérios para a seleção dos dados desejados.

Ao prosseguir para a próxima etapa, o usuário realiza a seleção de filtros que são interpretados e traduzidos para os campos equivalentes do modelo de dados. Por sua vez esses campos são armazenados em diferentes sessões de acordo com os grupos de atributos, sendo em seguida gerada e processada uma consulta SQL baseada nas informações resultantes das primeira e segunda etapas.

Ainda na terceira etapa é exibida uma síntese dos resultados retornados para que o usuário possa selecionar quais dados serão exibidos por completo, bem como o formato de exibição dos dados. Os formatos disponíveis são: Páginas HTML, arquivo ZIP e arquivo CSV.

Para o desenvolvimento da consulta em formato CSV foram estendidas as funcionalidades do sistema de forma a conservar sua arquitetura original. Também se buscou realizar as operações de forma otimizada com intuito de reduzir a carga de processamento de cada consulta realizada. Por essa razão ao selecionar o formato CSV para exibição, os parâmetros de seleção dos dados (atributos e filtros), obtidos na primeira e na segunda etapa, são utilizados na geração de uma única operação de “consulta” que retorna à totalidade dos dados.

Os resultados retornados pela consulta são utilizados pela função principal de processamento do arquivo CSV, responsável pela formatação dos dados. Ao final do processamento o usuário tem a possibilidade de salvar o arquivo CSV com os dados de sua consulta.

Ao final do desenvolvimento foram realizados testes com intuito de medir o tempo de processamento das consultas em formato CSV, com o objetivo de avaliar o desempenho da implementação. Os testes consistiram na execução de consultas a partir da interface do sistema para diferentes conjuntos de dados com tamanhos distintos. Os testes foram executados em um equipamento com 4 GB de memória RAM, processador “AMD

Phenom II X4 B93”, com 5.586,02 bogomips (VAN DORST, 1996), sistema operacional Linux e navegador Mozilla Firefox versão 16.0.2.

Nos testes realizados observou-se que, para consultas com: 3.000, 15.0000 e 30.0000 registros, foram obtidos respectivamente os seguintes tempos de resposta aproximados: 70 segundos, 420 segundos e 800 segundos.

O trabalho de desenvolvimento realizado conseguiu atingir o objetivo proposto de contribuir para a flexibilização das consultas do sistema BD Solos por meio da implementação do formato de consulta CSV; as funcionalidades foram estendidas seguindo a arquitetura original do sistema, o que contribui para facilitar a realização de futuras manutenções.

A implementação da nova consulta utilizando a estratégia de consultar a totalidade dos dados por meio de uma única requisição ao banco de dados (POSTGRESQL, 2012) se mostrou eficiente, uma vez que os testes realizados apresentaram tempos de resposta satisfatórios, mesmo para conjuntos com grandes quantidades de dados com cerca de 30.000 registros, apresentando um tempo de resposta de pouco mais de 10 minutos.

O trabalho desenvolvido terá continuidade com a participação dos usuários do sistema que iniciarão a validação das novas funcionalidades, para que sejam realizados o ajustes necessários.

Referências

OLIVEIRA, S. R. de M.; ZURMELY, H. R.; LIMA JÚNIOR, F. A. de; MEIRELLES, M. S. P.; SANTOS, H. G. dos; YAMASHITA, H. H. Um sistema para organização de informação de solos do Brasil na internet. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 7., 2009, Viçosa, MG. **Anais...** Viçosa, MG: UFV, 2009. Não paginado. SBIAgro 2009.

PHP Hipertext Preprocessor. Disponível em: <<http://www.php.net>>. Acesso em: 1 nov. 2012.

POSTGRESQL. Disponível em: <<http://www.postgresql.org/>>. Acesso em: 1 nov. 2012.

VAN DORST, WIM. “The Quintessential Linux Benchmark”. **Linux Journal**. Jan. 1996. Disponível em: <<http://www.linuxjournal.com/article/1120>>. Acesso em: 1 nov. 2012.

