

AGROSOFT 95

Feira e Congresso de Informática Aplicada à Agropecuária e Agroindústria

Proposta de uma Cooperativa de Banco de Dados sobre Recursos Naturais na Embrapa

Contato

Sergio A. B. Cruz

CNPTIA - EMBRAPA

Rodovia SP 65 Km 143,6 - Bairro dos Amarais

Caixa Postal 5.010

13.031-970 - Campinas - SP

e-mail: coopdb@cnptia.embrapa.br

Autoria

Sérgio A. B. Cruz - CNPTIA - EMBRAPA

Laurimar Gonçalves - CNPTIA - EMBRAPA

Suzana Druck Fuks - CNPS - EMBRAPA

Andrea Silva Hemerly - IBM Brasil - Centro Científico Rio

RESUMO

Visando atender a demanda por informações geradas por seus centros de pesquisa, a EMBRAPA (Empresa Brasileira de Pesquisa Agropecuária) iniciou esforços para implantação de sua cooperativa de banco de dados. Uma cooperativa de banco de dados é formada por um conjunto de centros de dados interligados por um sistema de comunicação. Os centros de dados são elementos que permitem a coleta, armazenamento, processamento, acesso e distribuição de dados. Alguns centros de pesquisa da EMBRAPA atuarão na cooperativa administrando os centros de dados. A consulta aos dados armazenados na cooperativa será realizada pelo usuário através dos serviços hipermedia do WWW. Dados georeferenciados tais como mapas temáticos e levantamentos de solo poderão ser acessados pelos usuários da cooperativa.

ABSTRACT

In order to make available the information generated by its own research centers, EMBRAPA has started an implementation for its cooperative database system. A cooperative database system is formed by some data centers interconnected by a communication system. The data centers are elements that allow data gathering, storage, processing, access and distribution. Some EMBRAPA centers will manage the data centers. The data stored on the cooperative database system will be searched by users through the WWW hypermedia services. Geographically referenced data such as thematic maps and soil research will be accessed by the cooperative data system users.

1. Introdução

A quantidade crescente de informações necessárias ao gerenciamento de grandes organizações têm exigido um esforço cada vez maior na padronização e descentralização no tratamento da informação. Este esforço visa facilitar a manipulação da informação e evitar a duplicação de esforços e inconsistências através da distribuição organizada das atividades necessárias ao processamento da informação. Esta estrutura onde a informação esta distribuída e integrada através de um sistema de comunicação define uma cooperativa de banco de dados. Uma cooperativa de banco de dados é formada por um conjunto de centros de dados interligados por um sistema de comunicação. Os centros de dados seriam responsáveis pela coleta, armazenamento, processamento e distribuição de dados.

O surgimento de novas tecnologias tem possibilitado a agilização na implementação destes esforços. Alguns exemplos de sistemas que estão sendo desenvolvidos [4][5], adotam a tecnologia dos SGBD (Sistema Gerenciador de Banco de Dados) em ambiente cliente-servidor para gerenciamento dos dados e a rede Internet [2] como meio de comunicação.

Diferentemente de um sistema de banco de dados distribuído, o acesso remoto via Internet numa cooperativa de dados visa basicamente a consulta. É responsabilidade do centro de dados "popular" o banco de dados com as informações a serem consultadas e distribuídas. Quanto mais organizada for a distribuição dos temas a serem tratados pelos diversos centros de dados menores serão as inconsistências e as redundâncias nas informações da cooperativa.

Uma visão integrada dos dados de uma cooperativa através da definição de um esquema global geral é dificultada pela heterogeneidade dos bancos de dados e modelos de dados adotados na criação destes diversos bancos. Este projeto adotará esquemas de integração de dados fracamente acoplados, onde a transparência de localização é realizada através do uso de metadados (descrição padronizada dos dados), visando a preservação das estruturas de armazenamento já existentes.

Uma estrutura geral para uma cooperativa usando a Internet, SGBD e serviços da WWW (World Wide Web) será apresentada. No final é descrito o projeto piloto de uma cooperativa de dados servindo dados sobre recursos naturais fornecidos pelo CNPS (Centro Nacional de Pesquisa de Solos) e CPAC (Centro de Pesquisa Agropecuária dos Cerrados), centros de pesquisa da EMBRAPA.

2. Estrutura de uma Cooperativa de Banco de Dados

Inúmeras iniciativas tem sido feitas na tentativa de fornecer acesso a dados geoespaciais [4] [5] [7] [9]. Uma estrutura geral no modo como os serviços WWW e os SGBD são usados nestas iniciativas é mostrado na seguinte figura.

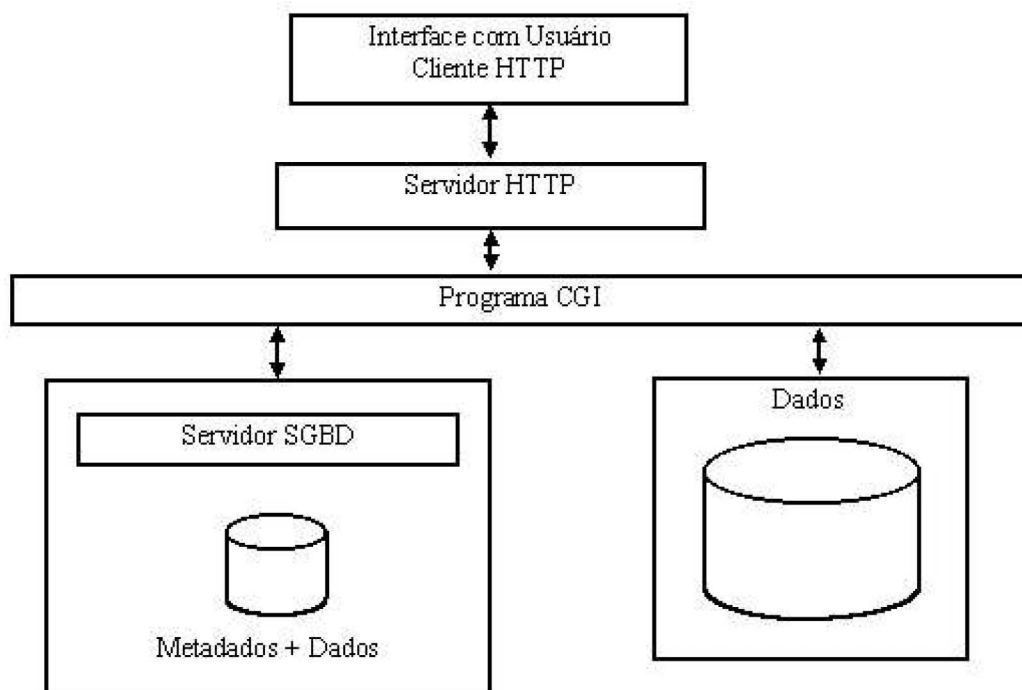


Figura 1 - Sistema de disponibilização de informação

A estrutura dos sistemas de disponibilização de informações atuais é suportada por dois elementos básicos permitindo os serviços da WWW: o cliente e o servidor http (hypertext transfer protocol)[2]. Estes dois elementos permitem a navegação pela Internet através interfaces hipermídia descritas usando a linguagem HTML (Hypertext Markup Language)[2].

O usuário interage com o Sistema da Cooperativa através destas interfaces gráficas. Uma "query" é especificada através do preenchimento dos formulários e/ou seleção via mouse de regiões nos mapas. Desta maneira são especificados parâmetros básicos de busca, tais como, limite geográfico da região de interesse (determinados pela seleção de regiões em mapas ou pelo nome da região), temas de interesse, limites temporais, escalas desejadas dos mapas, qualidade dos dados, etc. Programas clientes http, tais como o Mosaic ou Netscape [2] podem ser usados para visualizar a interface.

Este conjunto de parâmetros usados na descrição dos dados são designados metadados e servem como índice ou catálogo dos dados. Os dados podem estar armazenados juntos aos metadados e sendo gerenciados por um mesmo SGBD ou podem estar armazenados em algum repositório de dados separado.

A falta de uma padronização de formato dos metadados e dados dificulta o processo de busca e compartilhamento dos dados. Surgiram várias tentativas de padronização, como é o caso do FGDC (Federal Geographic Data Committee), entidade governamental americana que definiu um padrão de metadados [6], o qual todas as entidades federais americanas que produzem ou manipulam informações geoespaciais concordaram em adotar.

Os parâmetros de busca são passados para o servidor http através da Internet. Qualquer usuário tendo acesso a Internet, independente da sua localização poderá se conectar ao servidor http. Alguns serviços fornecidos pela WWW definem algum controle de acesso, através de identificação do usuário e senhas, impedindo desta forma o acesso indiscriminado à informação. Os parâmetros de busca passados para o servidor http são adequadamente tratados por um programa CGI (Common Gateway Interface) [8] e traduzidos para um formato adequado, geralmente o SQL, para definição de uma consulta ao banco de dados contendo os metadados.

Uma descrição sucinta dos resultados desta consulta são passados para usuário. Esta descrição pode conter um resumo do metadado associado aos dados selecionados pela consulta além de uma amostra dos dados. No caso dos dados serem imagens, as amostras correspondem a versões reduzidas da imagem original, para efeito de otimização do tráfego de informação na WWW. Este resultado parcial pode ser refinado pelo usuário através de consultas sucessivas até que se obtenha um conjunto desejado de dados. A cada apresentação de resultado o usuário poderá também selecionar os dados que deseja requisitar. Ao final desta sessão de consultas o usuário pode requisitar o envio dos dados escolhidos definindo os formatos e mídias em que deseja obtê-los. Alguns dados que são fornecidos "on-line" em forma de arquivos digitais podem ser requisitados durante a sessão de consultas.

O módulo mostrado na figura 1 apresenta as funções de disponibilização de dados e distribuição de produtos inerentes ao centros de dados. Existem outras funções de responsabilidade dos centros de dados, tais como: aquisição (coleta, pré-processamento, conversão e controle de qualidade) de dados, alimentação da base ou repositório de dados de maneira organizada, obedecendo aos padrões de metadado, gerenciamento do cadastro de usuários, produtos e fornecedores e controle de acesso aos dados. A cooperativa é formada pela interligação de diversos sistemas como o da figura 1.

3. Estratégia de Implantação da Cooperativa de Banco de Dados na EMBRAPA

A cooperativa de banco de dados será implantada na EMBRAPA em duas etapas. A primeira, corresponde a implantação de um projeto piloto, objetivando tornar disponíveis dados coletados pelos CNPS e CPAC sobre recursos naturais. Estas unidades da EMBRAPA atuarão como Centros de Dados nos temas de solo, chuva e clima. Os resultados desta etapa serão posteriormente aplicados à outros Centros da EMBRAPA e instituições afins, podendo também abranger outras áreas de conhecimento.

Os dados que serão fornecidos pelo CNPS correspondem aos diversos levantamentos de solos realizados pelo Centro. Além de dados técnicos aos pesquisadores, outras informações úteis aos produtores e cooperativas agropecuárias poderão ser fornecidas, tais como: mapa de solos, aptidão agrícola, necessidade de fertilização, impedimento à mecanização, avaliação pedoclimática por cultura, zoneamento agroecológico. Os dados tabulares associados a estes levantamentos estão atualmente armazenados no sistema SISOLOS [3], desenvolvido utilizando o ambiente NTIA, tecnologia gerada pelo CNPTIA/EMBRAPA. Os mapas devem ser digitalizados através do sistema SPRING (Sistema de Processamento de Informação Georeferenciada)[10].

Os dados que serão fornecidos pelo CPAC correspondem a estudos climáticos [1] baseados nos dados coletados em cerca de 67 estações climatológicas e 100 estações pluviométricas das regiões dos cerrados. Os mapas associados a estes levantamentos estão armazenados no sistema SGI/SITIM [10] e as tabelas em arquivos no formato Xbase. Resultados de processamentos estatísticos sobre os dados coletados serão fornecidos, tais como: normais climatológicas (precipitação pluviométrica, umidade relativa do ar, evapotranspiração,

insolação, déficit hídrico, pressão atmosférica) e análises pluviométricas (tabelas de distribuição por classes de dias sucessivos sem chuva e tabelas de análises frequenciais de precipitação pluviométrica).

SPRING e SGI/SITIM são SIG 's (Sistemas de Informação Geográfica) desenvolvidos pelo INPE (Instituto Nacional de Pesquisas Espaciais).

A implantação deste projeto piloto, inicialmente implementará as funções de disponibilização de dados, distribuição de produtos, gerenciamento de usuários, produtos e fornecedores. Atualmente o projeto encontra-se em fase de modelagem e desenvolveu-se alguns protótipos que auxiliaram no levantamento de requisitos.

4. Bibliografia

- [1] ASSAD.E.D., coord. Chuva no cerrados: análise e espacialização. Brasília: EMBRAPA/SPI, 1994., 422 pg.
- [2] KROL, E. The Whole Internet: User 's Guide & Catalog. Sebastopol, CA: O 'Reilly & Associates, 1992, CA , 376 pg.
- [3] EMBRAPA. Serviço Nacional de Levantamento e Conservação de Solos (Rio de Janeiro, RJ). SISOLOS - Manual de Uso, Rio de Janeiro, 1983. 245p. (EMBRAPA.SNLCS. Documentos, 4).
- [4] FREW, J.; SMITH, T.; ZHENG, Q.; SIMPSON, J.; Alexandria Digital Library (<http://alexandria.sdc.ucsb.edu>) . [S.I] 1995.
- [5] ENVIRONMENTAL RESOURCE CENTER. The ERC Earth Observation Images Database (<http://192.156.133.18/ERC/SSEOPhomepage.html>). [S.I.], 1995.
- [6] STITT, S. Content Standards for Digital Geospatial Metadata, (<http://www.its.nbs.gov/nbs/meta/meta.html>). [S.I]: Federal Geographic Data Committee, 1995.
- [7] PLEWEE, B. Geoweb Home Page (<http://wings.buffalo.edu/geoweb>)_[S.I.] State University of New York at Buffalo, 1995.
- [8] NATIONAL CENTER OF SUPERCOMPUTING APPLICATIONS. The Common Gateway Interface (<http://hoohoo.ncsa.uiuc.edu/cgi>). [S.I.], 1995.
- [9] UNITED STATE GEOLOGICAL SURVEY. USGS Node of National Geospatial Data Clearinghouse (<http://h2o.er.usgs.gov:80/nsdi/>) [S.I.], 1995.
- [10] CÂMARA, G. Anatomia de um SIG Fator GIS - A Revista de Geoprocessamento, Curitiba, v.1, n.4, p. 11-15, jan./mar. 1994.