

Construção de um Sistema Cartográfico Especialista: funções de classificação semi-automática de dados

Luciene Stamato Delazari ¹
Marcio Augusto Reolon Schmidt ¹
Luciana Fraron de Siqueira ¹

¹ Universidade Federal do Paraná – UFPR-DGEOM
Caixa Postal 19001 – 51531-990 - Curitiba - PR, Brasil
luciene@ufpr.br; marcio.schmidt@gmail.com; lufraron@yahoo.com.br

Abstract. This paper presents the elaboration of a cartographic expert system, for generation of thematic maps pursuing the implantation of the Organic Law of Social Assistance of the Paraná State. The cartographic expert system was created to assist users not specialists in cartography, to classify their information of economic and social indicators in the database that supports the Atlas. The system was built from Digital Interactive Atlas developed in a doctoral thesis and in its actual stage allows the semi-automatic classification of nominal and ordinal data. It was created a knowledge base that was used to classify the data accordingly the measure level. The system prevents the classification mistakes caused by the lack of knowledge of the criteria related to measurement level and graphic variables. Moreover, one of the advantages of the system is the exemption of training by users, that are not specialists in cartography.

Palavras-chave: thematic maps, data classification, mapas temáticos, classificação de dados.

1. Introdução

Este trabalho apresenta a elaboração de um sistema cartográfico especialista para geração de mapas temáticos. O objeto de estudo consiste de informações sobre Conselhos e Fundos Municipais no Estado do Paraná, bem como alguns indicadores estatísticos. O sistema, que está em constante desenvolvimento, consiste de um aprimoramento de um sistema computacional – denominado Atlas Digital Interativo – cujo objetivo é permitir análises espaciais de temas relativos à Assistência Social no Estado do Paraná.

Como esses usuários não têm conhecimento de cartografia, é fundamental que haja um sistema cartográfico que permita a manipulação dos dados e suas representações considerando os critérios e regras pré-estabelecidas, de modo que tais representações propiciem a aquisição da informação e do conhecimento. Além disso, a informação deve ser adequadamente processada no banco de dados sem intervenção do usuário. Dessa forma o treinamento do usuário em áreas não necessárias à exploração dos dados apresentados pelo Sistema Cartográfico pode ser dispensado, e se evitam erros de classificação por desconhecimento dos critérios inerentes à atividade do cartógrafo.

2. O Atlas interativo - evolução

Na primeira versão do Atlas, as informações coletadas em 1998 foram agrupadas em 3 grandes grupos de modo a produzir mapas que permitam identificar o tema geral que está sendo representado. As informações foram divididas em informações relacionadas aos Conselhos Municipais de Assistência Social, aos Fundos Municipais de Assistência Social e informações sobre a Assistência Social.

Nesta versão, Delazari (2004) desenvolveu o projeto cartográfico de modo a determinar as classificações dos dados e aparência visual de cada mapa individualmente, mas também de modo a manter a coerência entre os 26 mapas coropléticos. O objetivo principal do protótipo era possibilitar a representação espacial de informações relativas à Assistência Social no Estado do Paraná, de modo a subsidiar o entendimento da variação espacial de fatores associados à mesma. A partir deste objetivo outros foram propostos como análise e

comparação de padrões espaciais, análise espacial e de tendências, métodos de representação e a combinação de diferentes tipos de informações.

O protótipo foi desenvolvido em linguagem de programação Visual Basic 6.0, juntamente com o MapObjects 2.0a, desenvolvido pela empresa ESRI (*Environmental Systems Research Institute*). O MapObjects 2.0a é uma biblioteca de classes que fornece ferramentas para criação de aplicativos que possam ter funcionalidades de visualização cartográfica, permite trabalhar com projeções cartográficas, acessar bancos de dados, entre outras operações. Este programa compreende um controle ActiveX chamado Map e um conjunto de 46 objetos de automação ActiveX, que podem ser utilizados em ambientes de programação tais como Visual Basic, Visual C++, Delphi e Visual Basic for Applications (VBA).

Como resultado, obteve-se um sistema com as seguintes funcionalidades:

- O usuário pode escolher o tema, a informação a ser representada e esquema de cores;
- O sistema gera automaticamente um gráfico com a distribuição das informações;
- A legenda apresenta interatividade com o mapa;
- O usuário é capaz de uma análise sobre o mapa, digitando um texto que é armazenado e pode ser recuperado numa nova consulta.

A figura 1 apresenta a interface desenvolvida para o protótipo.

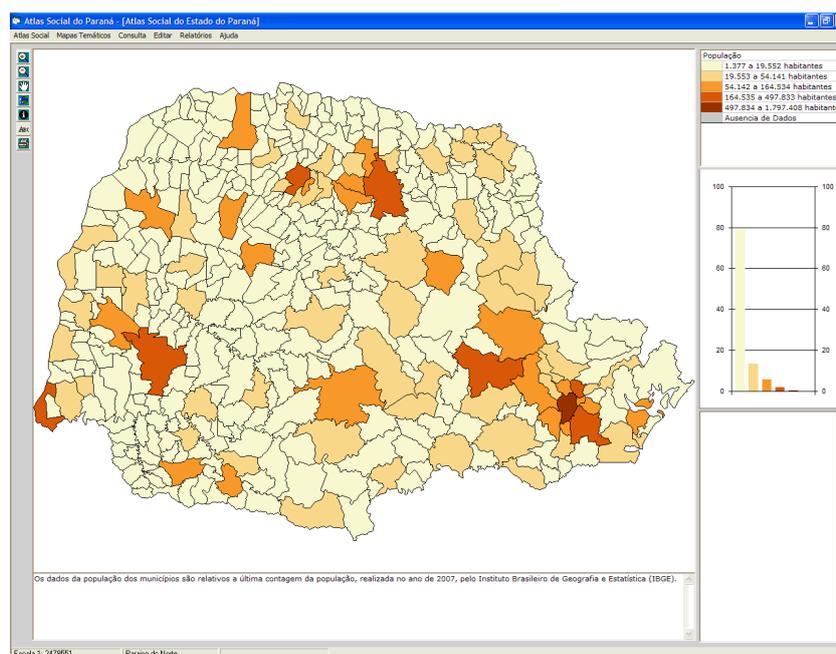


Figura 1 – Interface do Atlas Interativo

Na implementação original, o usuário acessava a seleção dos temas e os esquemas de cores na barra de menus na parte superior da interface. A relação do esquema de cores com o tema e a classificação temática era fixa no corpo do programa. Com isso, se evitava equívocos na seleção das cores para os mapas. Entretanto, qualquer nova informação inserida no banco de dados não era automaticamente classificada e representada. A inclusão de um sistema especialista (ver próxima sessão) e alterações na seleção dos esquemas de cores permitem que qualquer nova informação no banco de dados, após classificada seja passada para o mapa segundo as decisões de projeto cartográfico original. A nova implementação permite que o usuário escolha o método de classificação, o esquema de cores e consiga observar uma prévia de como o mapa ficará antes de sua criação.

Fachim (2008) realizou a inclusão de dados censitários atualizados dos 399 municípios do Estado do Paraná, tais como população, densidade demográfica, taxa de crescimento anual

populacional, IDH, PIB, PIB per capita, mortalidade infantil, taxa de pobreza, grau de urbanização, assistência médica, participação política e censo agropecuário. Esta autora também realizou a implementação de um sistema de segurança, com o objetivo de limitar o acesso de algumas funções de edição da base de conhecimento e do banco de dados para qualquer tipo de usuário. Somente o administrador tem acesso às funções de edição de informações, enquanto que a consulta aos mapas e informações, bem como geração de relatórios apresentam-se disponíveis para todos os usuários. A restrição de acesso a algumas funções faz com que o sistema seja mais confiável e estável.

2.1 Sistema especialista

Os usuários de SIG nem sempre tem o conhecimento do conceito de nível de medida e não sabe como escolher os símbolos cartográficos. A simbolização ocorre em termos da tríade formada pelo nível de medida, primitiva gráfica e variáveis visuais (Schmidt, 2009).

Existem várias possibilidades para tentar assegurar que o usuário não especialista faça a escolha correta, desde treinamentos de longa duração até outras opções, como inteligência artificial. Neste contexto os sistemas especialistas aparecem como uma possibilidade. Desta forma, Schmidt e Delazari (2009) criaram um sistema especialista para classificação de dados que compara informações textuais dos nomes das classes com um banco de palavras.

Foi construído um sistema especialista para a classificação de informações temáticas que pode ser distribuído para atender diferentes departamentos da Secretaria de Assistência Social do Estado do Paraná. Este sistema é colocado como uma camada entre o banco de dados propriamente e o dispositivo que realize a representação dos dados. Este sistema trabalha como um gerenciador do banco de dados através de regras do tipo SE-ENTÃO. Fixas no corpo do programa, estas regras identificam qual o tipo de informação é requisitada para representação, definindo-a em termos de nível de medida e a classificação armazenada na base de conhecimento. Diferentemente de outros sistemas especialistas, no caso deste sistema, as regras são fixas sendo que os parâmetros de avaliação usados na regra é que são continuamente atualizados. Assim quanto melhor definida e trabalhada as informações da base, maior será a eficiência da análise da regra.

Quando todas as classes são encontradas na base, as regras indicam dentre as paletas de cores disponíveis as mais adequadas para o tipo de dado; no momento estas paletas foram definidas pelo projeto cartográfico original de Delazari (2004). Para inserir novas informações, o sistema especialista testa o tipo de dado. Dados numéricos são classificados diretamente como nível de medida numérico. Dados nominais ou ordinais têm suas classificações e ordem de classes definidas na base de conhecimento. Quando se trata de dados não numéricos, ou seja, classificação semântica, a eleição de um ou outro nível de medida merece atenção, pois a correta classificação passa a ser subjetiva. Dados indicando temporalidade como primeiro e segundo semestres, são dados ordinais e tem de ser avaliados em detalhe.

Conforme indicado na figura 2, o sistema especialista procura na base de conhecimento por uma possível ordem das classes. Caso não seja possível, ou seja, as classes não são conhecidas, o nome das classes é decomposto em palavras. Esta lista de palavras de cada classe é então comparada novamente com a base de conhecimento que traz associada um dicionário com algumas palavras conhecidas elencadas na ordem correta. Ocorrendo sucesso, a base de conhecimento se retro alimenta com a nova classificação. Ocorrendo insucesso ou sucesso parcial, o usuário pode ser inquirido a intervir para indicar a ordem correta. Então o sistema armazena a classificação indicada, alimenta o dicionário e prossegue para a representação. À medida que os usuários utilizam a base e a alimentam com novas classificações o sistema passa a ser cada vez menos dependente da intervenção humana e com vocabulário mais complexo pode tratar de classificações mais complexas.

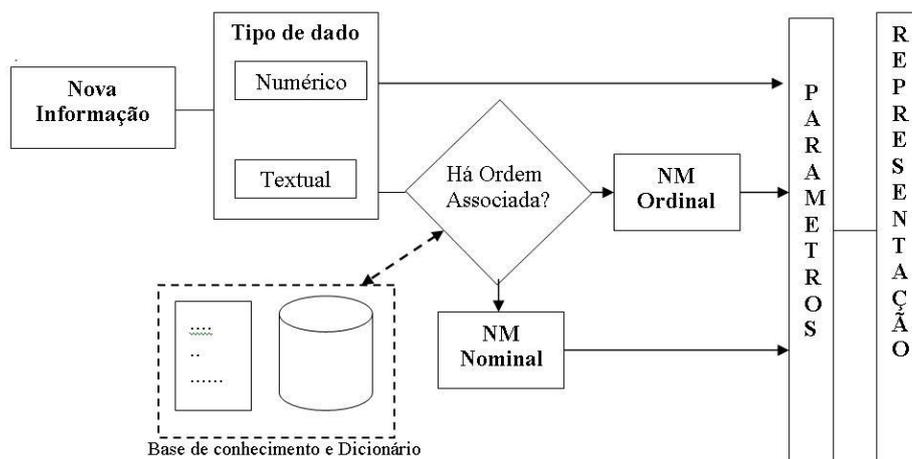


Figura 2 – Fluxograma de funcionamento do Sistema Especialista

A figura 3 apresenta o esquema de inserção de uma nova informação na base de dados. O usuário deve selecionar entre inserir um novo tema, à esquerda, ou simplesmente selecionar um dos temas existentes. Logo abaixo existem três campos para preenchimento. O campo Informação (1) é o nome que a coluna na base de conhecimento irá receber e pelo qual o usuário fará a pesquisa. O campo Classes (2) é onde o usuário faz a descrição de cada classe.

Ao fazer a adição de cada uma das classes essas são remetidas ao campo da direita (3), onde elas podem ser mantidas na ordem de inserção, ou ter sua ordem alterada manualmente ou excluídas através de botões. Essa ordenação junto com a descrição de cada classe será remetida para a base de conhecimento para armazenagem.

O sistema especialista procura no banco de dados e no dicionário os parâmetros para classificação da nova informação. Se não for possível classificar essa informação, o usuário é solicitado a informar se a ordem definida (4) deve ser seguida. O sistema traz para todas as informações uma pequena descrição ou histórico para auxiliar o usuário em sua pesquisa. Essa descrição pode ser inserida no campo Descrição da Informação (5).

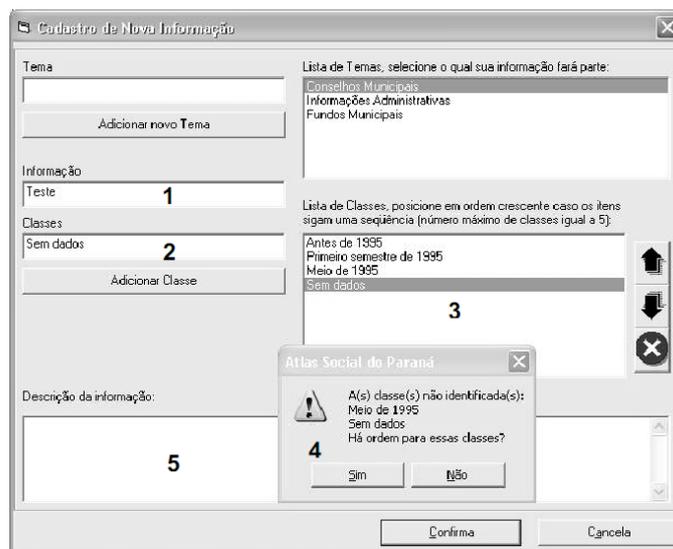


Figura 3 – Interface para o Sistema especialista

Por definição do projeto cartográfico, o programa limita o número de classes em 5. Caso o sistema especialista não consiga encontrar a classificação adequada e o usuário responda afirmativamente, uma mensagem é apresentada ao usuário. Em seguida há uma confirmação

da inclusão dos dados no banco de dados e na base de conhecimento. Então o programa abre uma janela para a associação das classes aos municípios.

2.2 Métodos de classificação de dados

Os algoritmos desenvolvidos até o momento realizam a classificação pelos métodos de intervalos constantes, quebras máximas, quantis e média e desvio-padrão. Futuramente, o algoritmo de classificação de Fisher-Jenks será incorporado ao rol de classificadores completando a etapa de classificação numérica do sistema.

A edição de dados nominal e ordinal é gerenciada pelo sistema especialista. Também está em fase de conclusão um módulo de pesquisas SQL que realizará a conexão entre o banco de dados e o programa e se conectará ao sistema especialista. Dentre as muitas vantagens, este módulo permitirá a inserção de consultas em SQL pelos usuários e facilitará o trabalho de manutenção e atualização das informações pelo administrador com a ajuda do sistema especialista.

A figura 4 abaixo mostra a opção para a classificação dos dados.

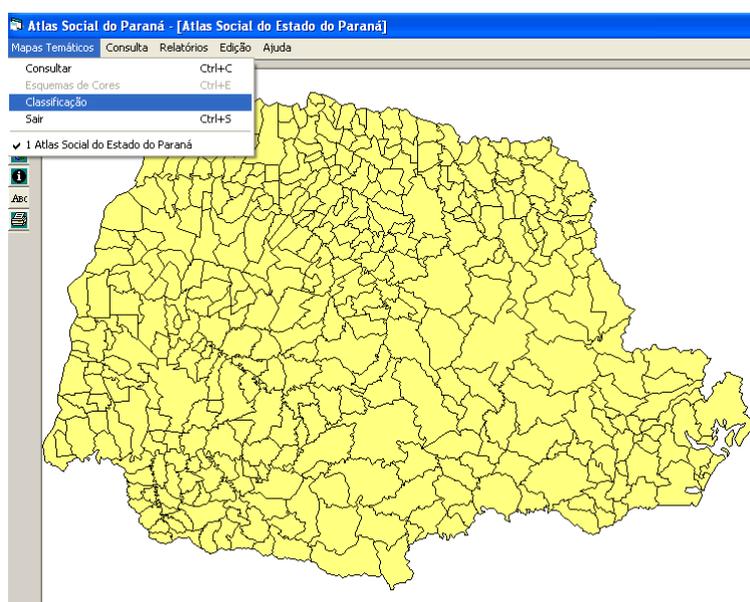


Figura 4 – Menu de edição para a classificação numérica dos dados

Ao optar pela classificação dos dados numéricos o usuário seleciona a informação que deseja representar (Figura 5), o método de classificação e o número de classes (Figura 6), sendo que após a seleção dos parâmetros o sistema apresenta o mapa temático.

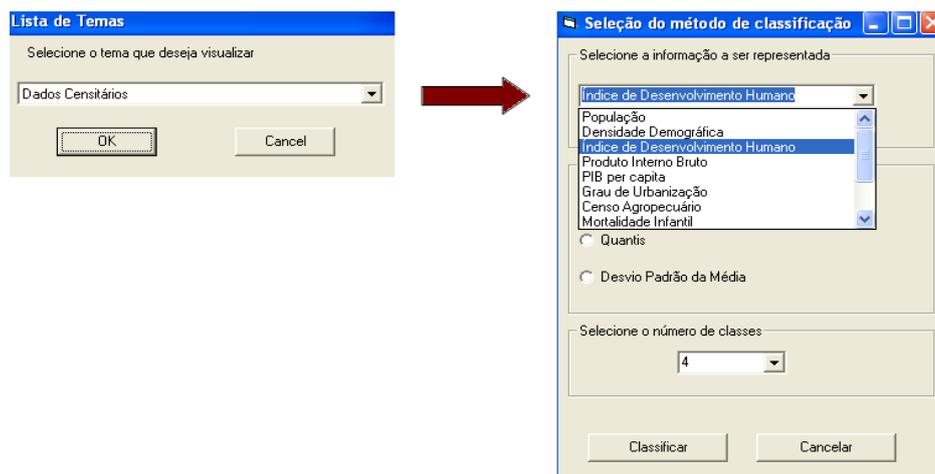


Figura 5 – Lista de temas seguida da escolha dos parâmetros para a classificação numérica.

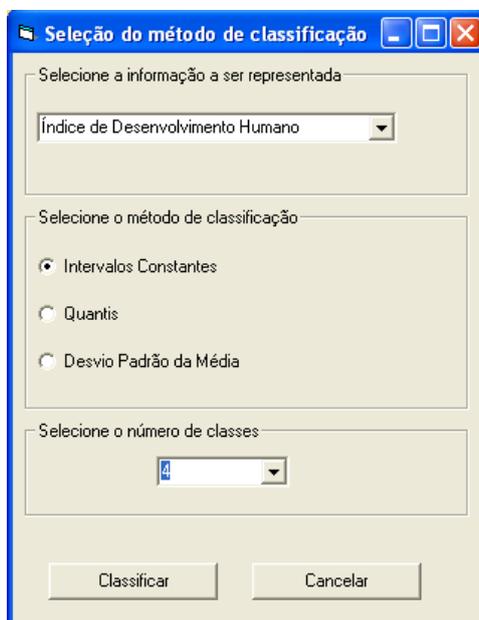


Figura 6- Seleção do método de classificação e do número de classes.

2.3 Outras funcionalidades do sistema

2.3.1 Arquivo de ajuda e relatórios

O Atlas Interativo oferece diversas funções, o que torna necessário a criação de um arquivo de ajuda cuja finalidade é auxiliar, direcionar e esclarecer dúvidas do usuário. O arquivo de ajuda foi desenvolvido em HTML integrado às aplicações Visual Basic, seguindo o padrão Windows. A ferramenta utilizada para trabalhar com os arquivos e gerar o produto final foi o HTML Help Workshop, da Microsoft.

O Atlas ainda cria relatórios a partir do Menu relatórios, sendo possível gerar relatórios específicos ou completos da maneira que o usuário desejar organizar, tanto por classificação das linhas como disposição das colunas

2.3.2 Manutenção do banco de dados

O usuário pode incluir novos dados e editar os antigos de modo a manter atualizado o sistema. Para isso, foram elaboradas novas telas e foi criado um novo item no menu Edição. Um item importante a ser preenchido quando o usuário insere uma nova informação e suas respectivas classes, é se aquelas classes seguem uma ordem seqüencial. Com isso o sistema,

no item de geração dos mapas temáticos, classificará como “nominal” ou “ordinal” aqueles dados. Na edição de ações é apresentada uma “combo-box” com cada ação disponível para aquela determinada informação, o que facilita e agiliza o preenchimento da tabela.

2.3.3 Interatividade entre a legenda e o mapa

Um dos objetivos do desenvolvimento desta pesquisa consiste na inclusão de funções de interatividade. No sistema em questão, foi implementada uma função que permite a interação entre a legenda e o mapa. Com a seleção de algum item da legenda, destaca-se no mapa a classe correspondente, com a sua respectiva cor na legenda (Figura 7), e ao selecionar o título da legenda, o mapa completo é mostrado novamente.

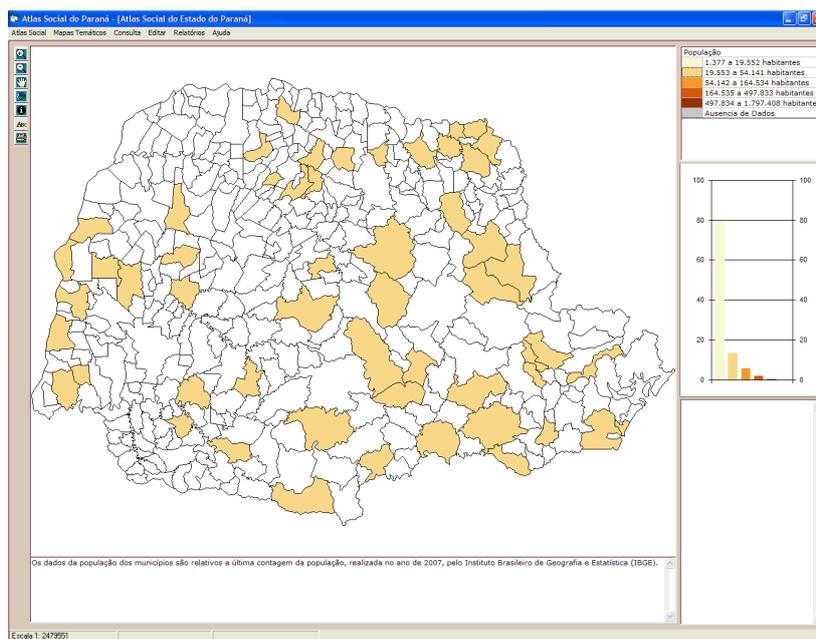


Figura 7 – Interatividade entre a legenda e o mapa

3 Trabalhos futuros

O sistema está em constante atualização e ampliação. O Atlas Social está sendo convertido para GeoServer (http://www.cartografia.ufpr.br/atlas_social/), um servidor de mapas baseado em Java. Nesta nova implementação, as informações do banco de dados original e o sistema especialista são armazenados no servidor da UFPR. A base de conhecimento definida pelo usuário fica armazenada em seus computadores pessoais. O sistema especialista é necessário para gerir adequadamente o fluxo de dados, identificando quando solicitar os dados que serão representados. Para conectar as informações dos dois bancos de dados, uma mensagem será enviada ao servidor de mapas web através de um script java. O script carrega as definições de classificação e simbologia do projeto cartográfico do Atlas Social e envia uma nova mensagem para os computadores dos usuários para proceder o carregamento dos novos dados. A comparação da base de conhecimento realizará a classificação dos novos dados ou atualização do banco de dados. Depois disso, o script gera um arquivo SLD, como um arquivo XML que define como os dados serão exibidos aos usuários do computador. A figura 8, apresenta a interface web do Atlas Social. Todas as ações serão gravadas em arquivos de registros (*log*) para controle e acompanhamento. As novas implementações são previstas para funcionar no segundo semestre de 2011.

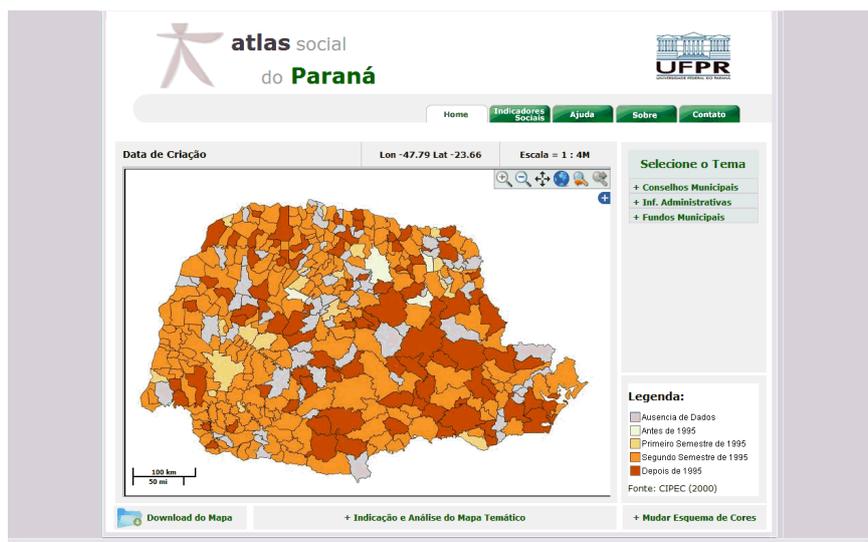


Figura 8 – Atlas Social na Internet

4 Conclusões

A pesquisa orientada à visualização cartográfica visa o desenvolvimento de novos métodos de representação e apresentação das informações geográficas. A Associação Cartográfica Internacional (ICA), em suas recomendações, destaca a importância da discussão e inclusão de novos conceitos em projetos para Atlas Digitais, tais como dos recursos de interatividade e de visualização cartográfica disponíveis atualmente.

A rápida disseminação das tecnologias SIG sem os conceitos cartográficos implica na necessidade de que esses sistemas orientem o usuário através do controle de funções e ferramentas de análises em SIG. O desenvolvimento de um sistema cartográfico especialista oferece ao usuário a possibilidade de construir seus próprios mapas, e de que estes estejam corretamente construídos no que diz respeito aos princípios de projeto cartográfico.

Até o momento foram desenvolvidas funções para inserção e classificação de dados nominais e ordinais. Além disso, possui diversas funções de interatividade e de auxílio à exploração e análise dos dados. As pesquisas atuais envolvem a classificação automática de dados numéricos, bem como a transposição deste sistema para web.

5. Referências

Delazari, L. S. 2004. Modelagem e implementação de um Atlas Eletrônico Interativo utilizando métodos de visualização cartográfica. São Paulo. 155p. Tese de doutorado na Escola Politécnica da Universidade de São Paulo – Departamento de Engenharia de Transportes.

Fachim, M. A. Schmidt, M.A.R., Delazari, L.S. Modelagem e implementação de um Atlas Eletrônico Interativo utilizando conceitos de visualização cartográfica. In: 16 EVINCI. *Anais...* Curitiba. 2008.

Schmidt, M. A. R.; Delazari, L. S. Expert system to classify semantic information to improve map design. In: 24th International Cartographic Conference, 2009, Santiago - Chile. Conference Proceedings: The World's Geo-Spatial Solutions. Santiago - Chile: ICA/ICC2009, 2009. v. 1. *Anais...* Disponível em http://www.icaci.org/documents/ICC_proceedings/ICC2009/html/refer/9_4.pdf

Schmidt, M.A.R. 2008. Automação de decisões para classificação de informações em atlas eletrônico utilizando sistema especialista baseado em regras. Curitiba. 124p. Dissertação de mestrado no Programa de Pós Graduação em Ciências Geodésicas da Universidade Federal do Paraná.

Agradecimentos

Ao CNPq pela bolsa Produtividade em Pesquisa, processo número 308892/2008-9.