

7 Conclusão e Trabalhos Futuros

O objetivo deste capítulo é resumir a pesquisa apresentada neste trabalho, enfatizando as principais contribuições e os possíveis trabalhos futuros.

7.1. Conclusões

Esta dissertação propôs uma sistemática baseada em data warehousing para apoio ao Governo Eletrônico. A motivação para o seu desenvolvimento foi um trabalho desenvolvido pelo Departamento de Informática da PUC-Rio para uma Organização Não-Governamental (ONG). Neste trabalho, foi possível perceber que o uso das estatísticas públicas vai muito além de quem a produz e o difícil trabalho para encontrar, acessar e usar as estatísticas públicas. Apesar do governo disponibilizar os dados brutos na Web, apenas os usuários especialistas conseguem tornar as informações estatísticas usáveis de forma analítica.

Foi então elaborada a hipótese que as informações estatísticas disponibilizadas em formato adequado para análise poderiam beneficiar tanto o governo quanto o cidadão. Além disso, seria um estímulo à pesquisa, pois muitos são os interessados nas estatísticas públicas produzidas pelos governos. Dessa forma, apoiar os governos para oferecer facilidade no acesso e no uso das estatísticas tornaria a exploração dos dados em um processo autônomo sem a dependência de um especialista. O objetivo de apoiar o governo para disponibilizar as estatísticas públicas em formato analítico na Web é estimular o uso da informação estatística pública, já que os governos estão reformulando suas formas de interação com os cidadãos, tornando-os mais participativos do processo de decisão pública, colaborando assim, com o processo de transparência.

Para colocar em prática esse apoio ao governo, foi desenvolvida uma sistemática baseada em data warehousing. Sua aplicação, entretanto, não se

restringe às estatísticas públicas. No meio corporativo, essa sistemática pode ser utilizada para apoiar a exploração de dados das diversas pesquisas estatísticas realizadas através de questionários fechados, como por exemplo, na área de marketing. Essa sistemática também pode ser usada na área acadêmica, em projetos de pesquisas quantitativas.

Para oferecer uma aplicação como prova-de-conceito seguindo as normas da W3C foram utilizados dados públicos reais e tecnologias recomendadas XML e Web service em um problema real do governo.

O protótipo apresentado mostra a simplicidade do acesso aos dados visto que basta indicar o endereço do cubo na Web. Quanto a usabilidade, o protótipo por ser uma aplicação OLAP oferece as características necessárias do grau de usabilidade conforme Cappelli(2009).

Dessa forma, observa-se que a construção de data warehouses e a aplicação de ferramentas OLAP oferecem um nível de usabilidade e transparência de um determinado processo que, embora não seja ideal para usuários sem conhecimentos aprofundados, atuam definitivamente fornecendo uma maior produtividade e eficiência na análise exploratória dos dados provenientes das estatísticas públicas.

Com isso, tanto o governo quanto a sociedade ganham. Para o governo fornecendo o serviço de acesso às informações, conseqüentemente mais transparência. Para o cidadão a independência de um terceiro, ou seja, autonomia de pesquisa. A usabilidade da informação estatística fortalece a interação entre governo e sociedade, pois oferece um ambiente analítico que permite que ambos possam conduzir os debates sobre políticas públicas fundamentados em informações comuns.

Esta subseção apresentou a conclusão da pesquisa apresentada nessa dissertação. A próxima apontará sugestões de trabalhos futuros relacionados à sistemática proposta nesse trabalho.

7.2. Trabalhos futuros

A sistemática proposta não se limita aos passos explorados nas etapas apresentadas. Dependendo da tecnologia e da pesquisa estatística em questão, novas idéias podem ser implementadas e exploradas. Nas próximas seções serão apresentados tópicos de interesse para pesquisas futuras guiadas pela sistemática proposta nesse trabalho.

7.2.1. Intervalo de Confiança em Pesquisas Amostrais

Nesta dissertação aplicamos a sistemática proposta para apresentar as informações da PNAD usando a tecnologia OLAP. A PNAD é uma pesquisa aplicada a uma parte representativa da população e depois são realizados cálculos estatísticos para estimar valores para toda a população. Nesse processo existe um conceito chamado de intervalo de confiança que indica a confiabilidade de uma estimativa. Um problema comum em dados multidimensionais é que se pode chegar facilmente ao nível de detalhe de uma célula. Isso causa problemas nos resultados das análises, porque quanto maior o número de cortes da análise maior se torna o intervalo de confiança para serem úteis (Li et. al, 2008).

A sistemática buscou oferecer todos os dados quantitativos das estatísticas e seus contextos para livre exploração, mas como a confiança dos dados é uma questão fundamental em exploração de dados, essa questão poderia ser mais explorada. Para trabalhos futuros, este protótipo ou outra aplicação da sistemática poderia sofrer melhorias para apresentar o intervalo de confiança nos cubos para o usuário final. Dessa forma, o acesso e a usabilidade dos dados através da tecnologia OLAP seriam conduzidos de forma análoga à proposta desse trabalho.

7.2.2. Tecnologia “Open Source”

Na arquitetura proposta, o uso do XMLA garante a interoperabilidade de sistemas permitindo que o acesso aos dados no servidor OLAP seja através de ferramentas gratuitas ou não. O modelo ideal para apoio ao e-Governo é o encontro do “open data”, “open souce”.

Dessa forma, as atuais ferramentas “open source” poderiam ser utilizadas para o processo de DWing, desde o acesso aos microdados e a disponibilização do cubos nos servidores OLAP, assim como, o uso de aplicações OLAP na Web que permitam o acesso aos cubos sem instalar nenhum software específico na máquina do usuário final.

7.2.3. Entrega de dados geográficos

Uma ferramenta SOLAP permite a navegação rápida e fácil nos bancos de dados espaciais e oferece vários níveis de granularidade de informação, muitos temas, diversos períodos e modos de exibição que podem estar sincronizados ou não: mapas, tabelas e diagramas. Esta tecnologia surgiu com a fusão da tecnologia OLAP e os componentes GIS (geographic information systems). Ferramentas SOLAP oferecem na forma tabular um guia interativo, com gráficos e visualizações cartográficas de cubos de dados espaciais, que são conjuntos de dados multidimensional que inclui dados geográficos, como a geometria dos países. Isso permite aos analistas explorarem plenamente o componente espacial.

Entretanto, a linguagem XMLA, utilizada nessa sistemática, para garantir a interoperabilidade de dados multidimensionais não possui suporte para troca de dados geográficos. Dessa forma, futuras aplicações dessa sistemática podem, em particular, inserir os dados geográficos nos cubos para consultas posteriores. Estudo de linguagens relacionada à troca de dados multidimensionais e geográficos com a implementação técnica de serviços Web devem ser realizados a fim de possibilitar a entrega de dados SOLAP aos clientes na Web.