

Desenvolvimento de uma Ferramenta Web para Migração e Visualização de Dados Abertos do DATASUS

Marcelo Henrique Casali¹
Cristiano Bertolini²
Guilherme Bernardino da Cunha³
Sidnei Renato Silveira⁴

Resumo. Dados da área da saúde são complexos e precisam de sistemas de visualização eficientes. Neste contexto, o principal objetivo deste trabalho foi o de analisar, implementar e migrar as bases de dados governamentais abertas na área da saúde de acordo com os dados do DATASUS. Os dados extraídos são referentes ao Sistema Único de Saúde (SUS). A maior contribuição é o sistema de visualização e migração das informações, representadas em mapas e gráficos por consultas em um banco de dados MySQL, trazendo todas as informações disponíveis de cada município e estado referente à morbidade hospitalar do SUS por causas externas.

Palavras-Chave. Sistemas de Visualização, Dados Abertos, DATASUS.

Abstract. Health data is complex and needs efficient visualization systems. The main objective of this work is to analyze, implement and migrate open government databases in the health area according to DATASUS data. The data extracted will refer to the single health system (SUS). The main contribution is the information visualization and migration system, represented in maps and graphs by queries in a MySQL database, bringing all available information from each municipality and state regarding SUS hospital morbidity due to external causes.

Keywords. Visualization Systems, Open DataBases, DATASUS.

1 INTRODUÇÃO

A lei de acesso à informação (Lei 12.527/2011) regulamenta o acesso a dados e informações mantidas pelo governo e constitui um marco para a democratização da informação nacional pública (BRASIL, 2011). O maior benefício da reutilização dos dados abertos é a facilidade de combinar dados nos formatos entregues (JSON – *Java Script Object Notation*, XML – *eXtensible Markup Language*, CSV – *Comma-Separated Values*, RDF – *Resource Description Framework*) com múltiplas fontes de dados, interligando-se com outras iniciativas de dados abertos na *web* (BRASIL, 2011).

A principal fonte de informação sobre dados abertos no Brasil é o Portal Brasileiro de Dados Abertos (PERREIRA, 2012). O portal está sendo construído e

¹ Curso de Bacharelado em Sistemas de Informação - Universidade Federal de Santa Maria - UFSM - Campus de Frederico Westphalen

² Departamento de Tecnologia da Informação - UFSM/Frederico Westphalen

³ Departamento de Tecnologia da Informação - UFSM/Frederico Westphalen

⁴ Departamento de Tecnologia da Informação - UFSM/Frederico Westphalen

estruturado de forma colaborativa entre diferentes instituições - públicas, privadas e sociedade civil (DADOS.GOV.BR, 2018).

Nota-se, também, que o DATASUS, com a base de dados sobre as atividades do Sistema Único de Saúde (SUS), vinculado ao Ministério da Saúde, também tornou-se uma importante ferramenta para a divulgação e publicação de dados abertos. Assim, além dos dados com base na área da saúde, também é possível obter informações financeiras e contábeis sobre o SUS (DATASUS, 2018).

Dados abertos governamentais são produzidos e colocados à disposição das pessoas para tornar possível não apenas sua leitura e acompanhamento mas, também, sua reutilização em novos projetos, *sites* e aplicativos. Uma dificuldade para se trabalhar com os dados abertos disponibilizados pelos portais brasileiros é o trabalho na modelagem dos dados, que realizam a limpeza, transformação e reorganização de todo o conteúdo que será utilizado. O principal desafio é o de equilibrar as necessidades da aplicação com os dados e formatos disponibilizados.

A partir do ano de 2009 o governo brasileiro sancionou a lei de acesso à informação, que determina que os órgãos públicos disponibilizem para a sociedade informações a respeito da gestão pública e, a partir de então, os dados abertos passaram a ser disponibilizados. “São citados ainda pelo W3C3 cerca de três benefícios que atingem diretamente ao cidadão a partir da disponibilização dos dados abertos: inclusão, transparência e responsabilidade” (OPENDATAMANUAL.ORG, 2018) (W3.ORG, 2018).

Não somente no Brasil, mas no mundo inteiro, a criação de projetos que utilizam dados abertos é utilizada para melhorar a vida da população como um todo. Diante deste cenário, o presente trabalho se propôs a definir um sistema único, com os dados disponibilizados pelo DATASUS referente à morbidade hospitalar dos pacientes atendidos pelo SUS por causas externas. Segundo as normas do SIH/SUS, as internações provocadas por causas externas devem ser classificadas no diagnóstico principal como: acidentes que englobam as quedas; envenenamento; afogamento; queimaduras; acidentes de trânsito.

Neste contexto, este trabalho está organizado da seguinte forma: a Seção 2 apresenta um breve referencial teórico composto por sistema de visualização, dados abertos e banco de dados; a Seção 3 apresenta os trabalhos relacionados e a Seção 4 descreve a solução implementada - o desenvolvimento de uma aplicação *web* denominada *WebSUS*. Por fim, são apresentadas as conclusões e as referências empregadas.

2 REFERENCIAL TEÓRICO

Esta seção apresenta um breve referencial teórico sobre as áreas de dados abertos, banco de dados e bibliotecas de visualização e análise de dados. Dividiu-se em três subseções: sistema de visualização, dados abertos e banco de dados.

2.1 Sistema de visualização

Sistemas de visualização consistem em comunicar a informação de maneira clara e efetiva utilizando meios gráficos. Não significa que a visualização de dados necessita ter um visual muito sofisticado. Para transmitir ideias efetivamente, tanto a forma estética quanto as necessidades funcionais precisam estar equilibradas, promovendo a compreensão de um complexo conjunto de dados, comunicando seus principais aspectos de uma forma mais intuitiva (FREITAS, 2001).

A ferramenta *GoogleCharts*, desenvolvida pela *Google*, fornece uma maneira adequada de visualizar dados incorporados em um *site*. O gráficos e mapas são expostos como classes em *JavaScript*, gráficos de linhas simples a mapas de árvore hierárquicos complexos (DEVELOPERS.GOOGLE.COM, 2018).

A ferramenta *GeoCharts* é um mapa geográfico, podendo assim trabalhar dados com países, continentes ou regiões, com cores e valores atribuídos a partes específicas. Os valores são exibidos como uma escala de cores, podendo personalizar um gráfico ou mapa para se adequar à aparência de um *site*. O mapa é processado no navegador usando um *plugin* incorporado. Entretanto, este mapa não é rolável ou móvel, mas pode ser configurado para permitir o *zoom* em alguns estados, porém o Brasil não está habilitado para esta opção no momento (DEVELOPERS.GOOGLE.COM, 2018).

2.2 Dados Abertos

O cenário de manipulação e criação de aplicativos com dados abertos ocorre entre o setor público e privado, permitindo que o cidadão saia do papel de mero receptor de informações e passe a ser o atuante que interpreta os dados abertos em seu benefício e de sua comunidade (OPENDATAMANUAL.ORG, 2018).

O termo Dados Abertos refere-se a dados que podem ser livremente usados, reusados e distribuídos por qualquer pessoa (OKFN.ORG, 2018). Os principais formatos utilizados para disponibilizar dados abertos são:

- JSON (*JavaScript Object Notation*), que é uma formatação de troca de dados, pois usa convenções que são as linguagens C e familiares, incluindo C++, *Java*, *JavaScript*, *Perl*, *Python*, entre outras (OKFN.ORG, 2018);
- XML (*eXtensible Markup Language*) recomendada pela W3C para a criação de documentos com dados organizados hierarquicamente, tais como textos, banco de dados ou desenhos vetoriais (LIBREOFFICE.ORG, 2018);
- CSV (*Comma Separated Values*): é um formato de arquivo de texto que pode ser usado para trocar dados de uma planilha entre aplicativos (SYMPHONY, 2018);
- RDF (*Resource Description Framework*): é uma sintaxe padrão para representar um grafo dirigido em XML desenvolvida pelo W3C - *World Wide Web Consortium* (W3C, 2001). É um modelo padrão para o intercâmbio de dados na *web*. Uma de suas características é a facilidade de fusão de dados (W3.ORG, 2018).

Visando promover, de forma mais organizada e produtiva, o desenvolvimento de aplicativos que utilizem dados abertos, as organizações públicas e privadas têm organizado movimentos onde desenvolvedores, analistas e profissionais de diversas áreas trabalham juntos na análise e tratamento dos dados para posterior desenvolvimento dos aplicativos.

Esses eventos são chamados de *hackathon* (OPENDATAMANUAL.ORG, 2018). O *hackathon* é um tipo de evento que conta com a parceria da *Open Knowledge Foundation* (OPENDATADAY.ORG, 2018) e da W3C Brasil (W3C.BR, 2018). Durante estes eventos a instituição realiza, organiza e disponibiliza a massa de dados aos participantes para análise e as equipes participantes têm o prazo geral de 48 horas para a análise e modelagem dos dados para o desenvolvimento, teste e entrega do projeto. No final deste período, os projetos são expostos e passam por uma votação onde os melhores projetos são premiados e disponibilizados ao público (OPENDATADAY.ORG, 2018).

A criação de projetos que utilizam dados abertos visa melhorar a vida da população global, auxiliando as pessoas a programarem melhor suas atividades e, até mesmo, para definir a localização de um futuro relacionando às suas condições de vida.

Diante deste cenário, o manual dos dados abertos governamentais cita algumas características a respeito: transparência e controle democrático; participação popular; melhores ou novos produtos e serviços privados; inovação; melhora na eficiência e efetividade dos serviços governamentais; medição do impacto das políticas; conhecimento novo a partir da combinação de fontes de dados e padrões (W3C.BR, 2018).

2.3 Banco de Dados

Em um mundo onde a globalização está cada vez mais presente na vida do cidadão e os processos estão cada vez mais automatizados, a necessidade de armazenamento de dados e informações se torna o primeiro passo para iniciar um projeto com dados abertos.

O *MySQL* é um servidor e um Sistema Gerenciador de Banco de Dados (SGBD) relacional, de licença dupla (sendo uma delas baseada em *software* livre), projetado inicialmente para trabalhar com aplicações de pequeno e médio portes, principalmente com aplicações para interface *web*, possuindo todas as características que um banco de dados de grande porte precisa, sendo reconhecido por algumas entidades como banco de dados *OpenSource* (PACHEVSASHA, 2007). O *MySQL*, além de ser extremamente rápido pelo fato de armazenar os dados em tabelas no modo ISAM (*Indexed Sequential Access Method*) (código de baixo nível), é altamente confiável. Por possuir essas características, a ferramenta é indicada para o uso em aplicações em todas as áreas, independentemente do tamanho da aplicação (MYSQL.COM, 2018).

3. TRABALHOS RELACIONADOS

A ação do Governo Federal em propiciar dados relacionados a múltiplos serviços, envolvendo dados da saúde, abre muitas possibilidades para visualização de dados e planejamento. Nesta seção são apresentados alguns trabalhos relacionados ao trabalho desenvolvido.

Segundo Ribeiro (2011), várias ações estão sendo realizadas para incentivar o trabalho com dados abertos, não somente no Brasil mas em nível mundial. Estas ações estão facilitando o intercâmbio de dados e a descoberta de informações relevantes que levam à melhoria dos serviços como, por exemplo, os relacionados à saúde pública.

Ribeiro (2011) cita que existe um projeto em desenvolvimento, chamado DadosGov desenvolvido pelo Comitê de Organização de Informações da Presidência da República (COI-PR). O projeto permite a criação de um catálogo de informações abertas, aprimorando a gestão pública e facilitando o acompanhamento por parte dos cidadãos.

Segundo Traina (2001), os seres humanos possuem uma capacidade inata para interpretar informações apresentadas de forma gráfica, conseguindo reconhecer mais facilmente o conhecimento associado às informações. Traina (2001) apresenta uma ferramenta para a visualização de dados de forma gráfica, onde os dados armazenados ficam em uma base relacional, a qual é independente de qualquer processo de sumarização dos mesmos. Entretanto, a maioria das ferramentas permite somente a visualização das informações estáticas sobre os dados e poucas permitem a visualização de forma dinâmica. Traina (2001) frisa o futuro desenvolvimento de uma técnica para a visualização de dados a partir de base de dados composta por qualquer tipo de informação, assim implementando uma ferramenta prática para a reutilização da base de dados.

Canban (2015), destaca que existem novas interfaces interativas na área da saúde em desenvolvimento, pois o volume de dados clínicos é grande. Assim, existe a necessidade de apoiar novos projetos para a saúde e, desta forma, médicos podem encontrar maneiras mais eficazes para tratamentos clínicos e futuras tomadas de decisões. O trabalho apresenta uma análise visual da saúde em forma de oportunidades no mercado tecnológico e desafios para pesquisas, gerando dados com pacientes em forma de relatórios, por meio dos quais os médicos podem comparar casos de pacientes semelhantes, facilitando assim o tratamento. Canban (2015) cita um sistema chamado HARVEST, que tem uma interface interativa para registros de cada paciente, com isso fornecendo um amplo conteúdo para ser analisado. Assim, os clínicos e pesquisadores podem gerar hipóteses e obter uma compreensão mais profunda de cada caso. Em particular percebeu que algumas áreas como a saúde, devem ter mais atenção quando o assunto é tecnologia, devido à falta de métodos padronizados de avaliação, validação e eficácia das ferramentas de análise visual.

Germano (2008) apresenta um estudo realizado como parte de uma pesquisa em aglomerados de municípios brasileiros com precariedade das informações de óbito, visando estimar a mortalidade infantil nessas áreas no ano de 2000. O trabalho discute as principais fontes de informação do óbito infantil nesses municípios, assim como alguns dos problemas relacionados à operacionalização do Sistema de Informações sobre Mortalidade (SIM). A metodologia adotada foi a de busca ativa de óbitos em fontes oficiais e não oficiais de informações previamente cadastradas. A busca ativa identificou 344 óbitos não processados no SIM (66% do total de óbitos). Foi observado um baixo percentual de óbitos processados no SIM oriundos das fontes oficiais (cartórios e estabelecimentos de saúde), o que assinalou para problemas tanto na notificação do óbito e emissão da declaração de óbito quanto no fluxo de informações. A elevada contribuição do Programa Agentes Comunitários de Saúde/Programa Saúde da Família na identificação de óbitos infantis aponta para a necessidade de busca ativa de informações nesta fonte, como rotina nos municípios com precariedade dos dados.

Germano (2008) frisa que a insuficiente capacitação dos profissionais e a alta rotatividade de técnicos foram questões-chave para a deficiente operacionalização do sistema.

Segundo Zacarias (ZACARIAS et al. 2015), devido à falta de dados disponíveis e de alguns dados serem incompletos, fica difícil realizar comparações em bases de dados. Os autores perceberam que os dados referentes a acidentes de trabalho não têm a devida importância pois, além de poucos dados, eles não são atualizados periodicamente. Com uma maior quantidade de dados, acredita-se que a aplicação desenvolvida (Análise Comparativa dos Acidentes de Trabalho no Brasil a partir de Dados Abertos) poderia ser estendida para o uso de técnicas estatísticas, assim prevendo determinadas situações e, até mesmo, para servir de apoio a decisões do governo.

O trabalho aqui apresentando estende o proposto em Zacarias (ZACARIAS et al. 2015) e utiliza, além de um volume de dados maior, uma maior precisão na visualização dos mesmos.

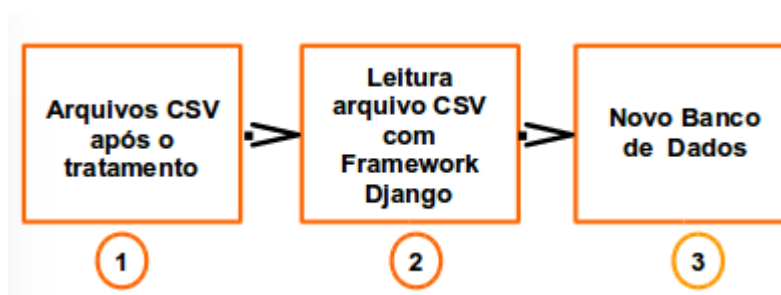
4. SOLUÇÃO IMPLEMENTADA

Nesta seção será apresentado o sistema de visualização de dados para o DATASUS denominado *WebSUS*. Tem-se, como solução implementada, a criação de uma aplicação que proporcionará ao usuário a visualização dos dados abertos disponibilizados no DATASUS, por meio de mapas e gráficos, sobre a morbidade hospitalar dos pacientes atendidos no SUS com causas externas.

4.1 Base de Dados

Para a construção da base de dados, foram utilizados os arquivos disponíveis no DATASUS, no formato CSV. Desenvolveu-se um algoritmo utilizando a linguagem de programação *Python* com o *Framework Django*, que faz a importação dos arquivos CSV para o SGBD *MySQL* permitindo, assim, a inserção dos dados. O processo de tratamento dos arquivos CSV e inserção dos dados para o SGBD *MySQL* foi desenvolvido em três etapas diferentes conforme mostra a Figura 1.

Figura 1. Etapas do processo de criação do Banco de Dados



Fonte: dos autores

Na primeira etapa do processo foi realizado o tratamento do arquivo CSV oriundo do DATASUS, onde é possível visualizar o arquivo original e o arquivo após o tratamento dos dados, para a leitura e a sua inserção manualmente no banco de dados *MySQL*.

A Figura 2 apresenta o arquivo CSV sem tratamento dos dados. Cada linha corresponde a uma cidade do Brasil. As colunas apresentam informações das cidades,

Fonte: dos autores

Na segunda etapa, desenvolveu-se um algoritmo em *Python* com o *Framework Django* para tratar especificamente dos arquivos CSV, possibilitando a leitura do arquivo CSV, até a inserção dos dados diretamente na tabela específica do banco de dados *MySQL*. O algoritmo é apresentado na Figura 4.

Figura 4. Algoritmo desenvolvido

Algorithm 1 Algoritmo

```
1: procedure def LERCSV(request):
2:   coluna = 1
3:   mes = 1
4:   anoInicial = 2008
5:   anoFinal = 2017
6:   while (anoInicial < anoFinal) : do
7:     with open('arquivo.csv','rb') as csvfile :
8:       spamreader = csv.reader(csvfile, delimiter = ',' , quotechar = '|')
9:       for row in spamreader: do
10:        if row[0] is not None: then
11:          t = Internacao()
12:          t.idmunicipio_id = row[0]
13:          t.quantidade = row[coluna]
14:          t.data = str(ano) + '-' + str(mes) + '-01
15:          t.save(force_insert = True)
16:          print t.idmunicipio_id, t.data, t.quantidade
17:        mes = mes + 1
18:        coluna = coluna + 1
19:        print mes
20:        if coluna % 12 == 1 : then
21:          print anoInicial
22:          mes = 1
23:          anoInicial = anoInicial + 1
```

Fonte: dos autores

O algoritmo apresentado na Figura 4, verifica o conteúdo no arquivo CSV e, após a verificação, insere todos os registros encontrados diretamente na tabela específica no banco de dados *MySQL*.

Na terceira etapa, pode-se visualizar a estrutura do banco de dados e o número de registros. Após a inserção de todos os dados, tem-se o novo banco de dados com dados inseridos do DATASUS, morbidade hospitalar dos pacientes atendidos no SUS com causas externas entre 2008 e 2015, como mostra a Figura 5.

Figura 5. Banco de Dados Implementado

Tabela	Ação	Linhas	Tipo	Colaço	Tamanho	Sobrecarga
aih	Visualizar Estrutura Procurar Inserir Limpar Eliminar	341,190	InnoDB	utf8_general_ci	15 MB	-
causas	Visualizar Estrutura Procurar Inserir Limpar Eliminar	110,385	InnoDB	utf8_general_ci	20.1 MB	-
django_migrations	Visualizar Estrutura Procurar Inserir Limpar Eliminar	0	InnoDB	utf8_general_ci	16 KB	-
estados	Visualizar Estrutura Procurar Inserir Limpar Eliminar	27	InnoDB	utf8_general_ci	16 KB	-
internacao	Visualizar Estrutura Procurar Inserir Limpar Eliminar	341,190	InnoDB	utf8_general_ci	22 MB	-
municipios	Visualizar Estrutura Procurar Inserir Limpar Eliminar	5,565	InnoDB	utf8_general_ci	560 KB	-
obitos	Visualizar Estrutura Procurar Inserir Limpar Eliminar	149,634	InnoDB	utf8_general_ci	8 MB	-
permanencia	Visualizar Estrutura Procurar Inserir Limpar Eliminar	341,190	InnoDB	utf8_general_ci	17 MB	-
populacao	Visualizar Estrutura Procurar Inserir Limpar Eliminar	44,520	InnoDB	utf8_general_ci	3.7 MB	-
taxa	Visualizar Estrutura Procurar Inserir Limpar Eliminar	149,634	InnoDB	utf8_general_ci	8 MB	-
valor	Visualizar Estrutura Procurar Inserir Limpar Eliminar	341,190	InnoDB	utf8_general_ci	17 MB	-
11 tabelas	Soma	1,824,525	InnoDB	utf8_general_ci	111.6 MB	0 Bytes

Fonte: dos autores

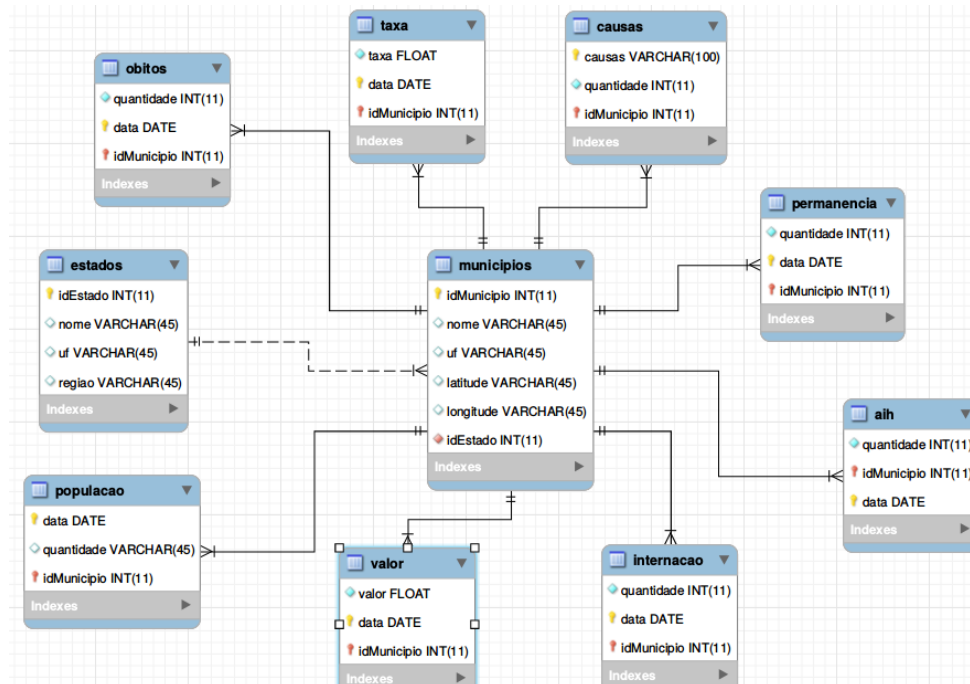
4.2 Consultas à Base de Dados

As consultas são baseadas na morbidade hospitalar do SUS com causas externas entre janeiro de 2008 a dezembro de 2015, podendo então interagir e relacionar os dados para futuras análises:

- Número de óbitos: designa a morte da pessoa que estava internada;
- Número de internações: qualquer internação motivada por uma causa externa, que por sua gravidade ou complexidade, exija um tratamento clínico ou cirúrgico;
- Valor total gasto: valores gastos com médicos, hospitais, internações, medicamentos e serviços em gerais;
- Taxa de mortalidade: a taxa de mortalidade é um índice demográfico obtido pela relação entre o número de mortos de uma população e um determinado espaço de tempo, neste caso, um ano;
- Autorizações de internações hospitalares: arquivo do SUS conhecido como "AIH". A AIH é originada quando um hospital ou uma unidade de saúde gera uma solicitação de internação hospitalar e tem o objetivo de pré-validar os dados de internação;
- Dias de permanência: relação entre o total de pacientes/dia e o total de internados que tiveram saída do hospital em determinado período, incluindo os óbitos. Representa o tempo médio em dias que os pacientes ficaram no hospital;
- População: número de habitantes referente ao ano informado. Com base neste dado é possível verificar a média de óbitos pela quantidade de habitantes.

A Figura 6 apresenta o modelo Entidade Relacionamento (ER) do banco de dados, onde é possível visualizar o modelo dos dados implementados.

Figura 6. Modelo Entidade-Relacionamento



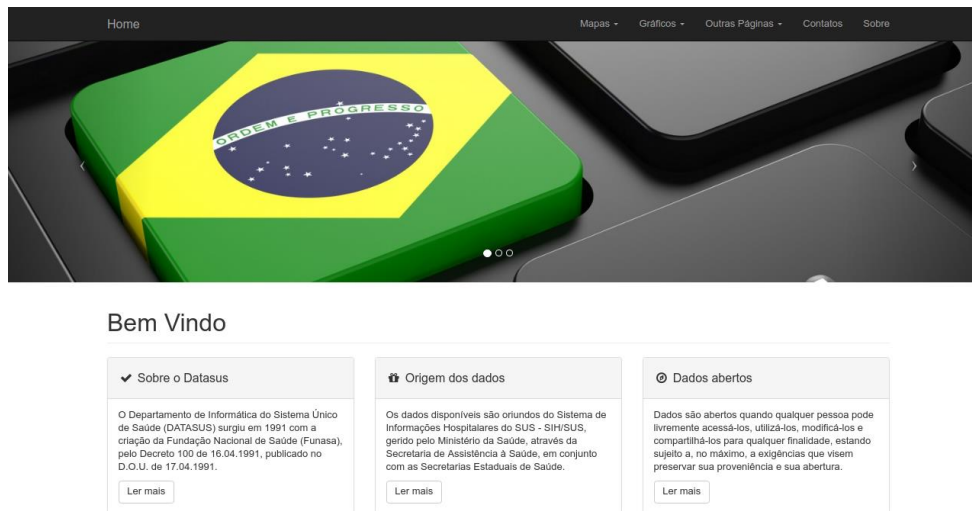
Fonte: dos autores

4.3 Sistema de Visualização

Os dados utilizados no Sistema de Visualização desenvolvido são fornecidos em um arquivo CSV, o que dificulta a interpretação dos mesmos, sem um tratamento para este projeto. Traina (2001) enfatiza que os seres humanos não são eficientes para “interpretar” grande volume de dados em forma numérica ou textual, especialmente em espaços de altas dimensões, mas têm uma percepção muito boa quando esses dados são apresentados em forma de gráficos.

A visualização gráfica nos permite inferências sobre os dados e possibilita a comparação entre fontes diferentes. Sendo assim, este trabalho apresenta o desenvolvimento de uma aplicação *web* que utiliza os dados do DATASUS e os apresenta de forma gráfica. Para o desenvolvimento do projeto foram utilizadas as tecnologias: linguagens de programação PHP e JavaScript, HTML (*HyperText Markup Language*) interagindo com a biblioteca *Google Charts* e o SGBD *MySQL*. A Figura 7 permite a visualização da interface *web* da aplicação desenvolvida.

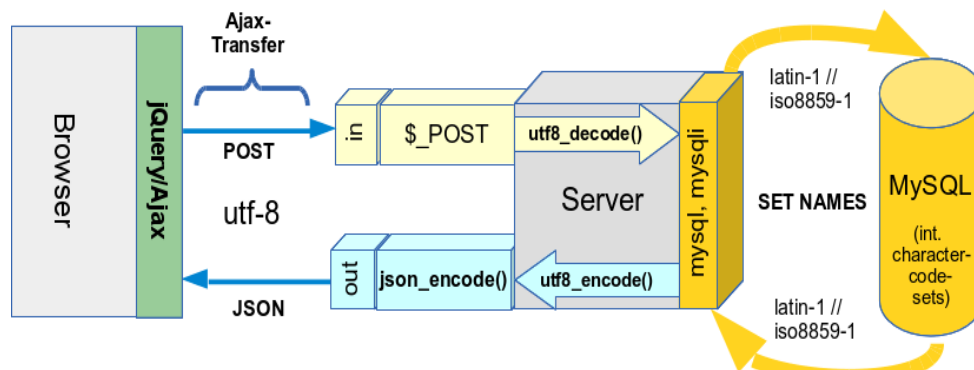
Figura 7. Interface Web da Aplicação Desenvolvida



Fonte: dos autores

Um dos objetivos deste trabalho foi o tratamento dos dados para a criação de um banco de dados relacional. Desta forma foi possível utilizar estes dados na biblioteca *GoogleCharts* para indicar a localização e, conseqüentemente, apresentar os dados em forma de gráficos e mapas. Esses dados são obtidos por meio de uma requisição em JSON a partir de consultas no banco de dados. A Figura 8 mostra a comunicação entre as tecnologias que foram utilizadas no projeto *WebSUS* para realizar a comunicação dos dados com a ferramenta *GoogleCharts*.

Figura 8. Comunicação entre as Tecnologias adotadas no projeto



Fonte: dos autores

De acordo com a Figura 8 tem-se:

- O *browser* envia dados de um formulário pelo método \$POST (envio de dados);
- O arquivo desenvolvido em PHP recebe os dados enviados pelo método \$POST e trabalha juntamente com o SGBD *MySQL*, realizando a consulta diretamente no servidor;
- O `JSON encode()` envia para o *browser* os dados obtidos na consulta por meio de um *array*;
- O *Google Charts* trabalha juntamente com o *browser* utilizando o JSON recebido.

A Figura 9 apresenta um exemplo de arquivo no formato JSON, onde AIH representa o número de autorizações de internações hospitalares, o nome que representa a cidade, e o valor representa o valor total gasto.

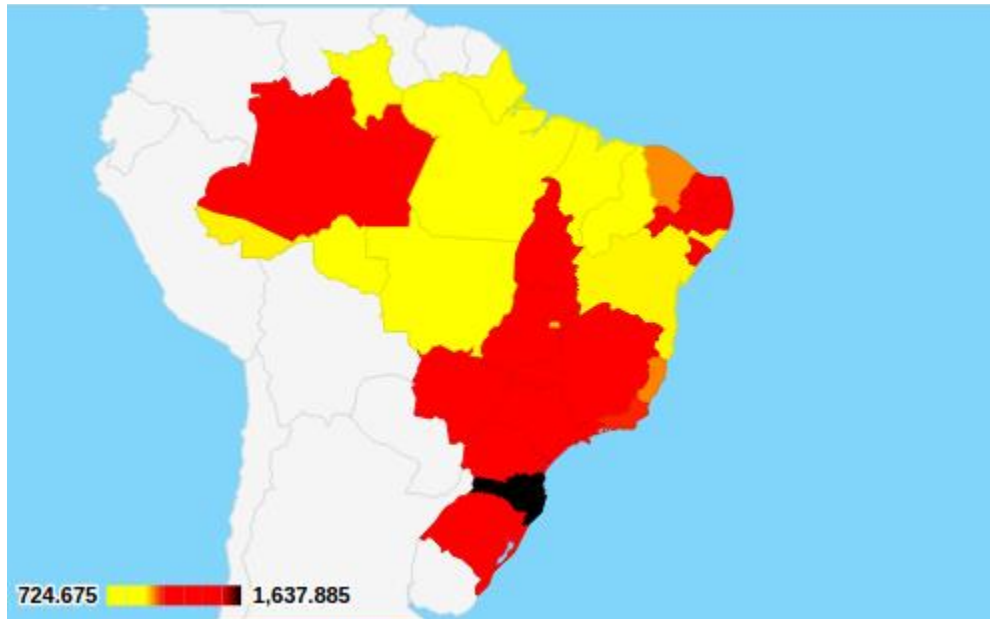
Figura 9. Exemplo de um Arquivo JSON

```
localhost/valor/valor.2014.php
[{"aih":"4678","nome":"Acre","valor":"4126272.4800720215"}, {"aih":"13593","nome":"Alagoas","valor":"13742680.124191284"}, {"aih":"1994","nome":"Amapá","valor":"1329802.1221733093"}, {"aih":"10020","nome":"Amazonas","valor":"9781477.04921341"}, {"aih":"68265","nome":"Bahia","valor":"68122334.15128326"}, {"aih":"51493","nome":"Ceará","valor":"51212976.524513245"}, {"aih":"17495","nome":"Distrito Federal","valor":"17228890"}, {"aih":"24343","nome":"Espírito Santo","valor":"24819432.56194687"}, {"aih":"41918","nome":"Goiás","valor":"48724016.550457"}, {"aih":"35515","nome":"Maranhão","valor":"25961064.15840912"}, {"aih":"25924","nome":"Mato Grosso","valor":"21572378.423171997"}, {"aih":"19944","nome":"Mato Grosso do Sul","valor":"22309506.458946228"}, {"aih":"126446","nome":"Minas Gerais","valor":"164306427.9190445"}, {"aih":"51256","nome":"Pará","valor":"40725003.95681763"}, {"aih":"15225","nome":"Paraíba","valor":"18562851.41984558"}, {"aih":"80017","nome":"Paraná","valor":"110249288.33564758"}, {"aih":"55191","nome":"Pernambuco","valor":"71236711.13385773"}, {"aih":"21013","nome":"Piauí","valor":"19917831.895137787"}, {"aih":"65780","nome":"Rio de Janeiro","valor":"70042158.13788223"}, {"aih":"12102","nome":"Rio Grande do Norte","valor":"13828167.586532593"}, {"aih":"62484","nome":"Rio Grande do Sul","valor":"79554368.58679962"}, {"aih":"14642","nome":"Rondônia","valor":"11403203.085609436"}, {"aih":"2918","nome":"Roraima","valor":"2371435.660003662"}, {"aih":"44767","nome":"Santa Catarina","valor":"68237952.10068512"}, {"aih":"235557","nome":"São Paulo","valor":"296687549.46819305"}, {"aih":"7841","nome":"Sergipe","valor":"8401461.533565521"}, {"aih":"13554","nome":"Tocantins","valor":"14781669.150894165"}]
```

Fonte: dos autores

Para a construção do mapa, foi utilizado o pacote *GeoCharts* do *GoogleCharts*. O serviço permite incorporar um mapa em uma página *web* usando a tecnologia *JavaScript*. A Figura 10 apresenta um mapa, referente ao valor gasto e um valor médio de cada internação por causas externas do SUS no ano de 2015. Pode-se observar que o estado de Santa Catarina apresentou um maior custo por internação em relação à quantidade de internações, sendo um custo médio por internação de R\$ 1.637,88 e gasto total de R\$ 74.394,35. Os estados com menos gastos por internações são: Amapá, com custo médio por internação de R\$ 763,34 e gasto total de R\$ 1.664,83 e Maranhão, com custo médio por internação de R\$ 724,67 e gasto total de R\$ 26.376,71.

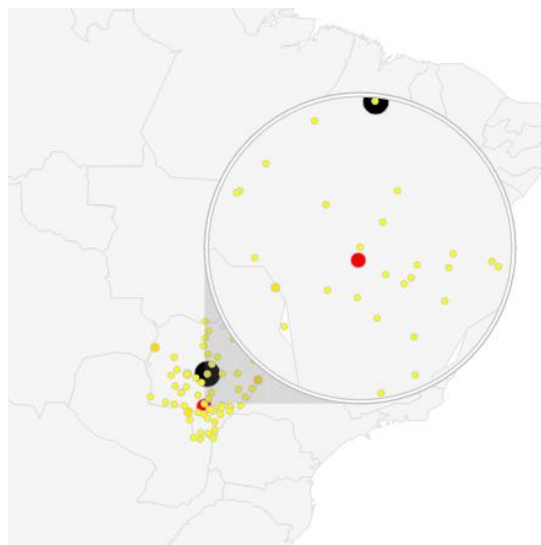
Figura 10. Mapa Indicador de Estados



Fonte: dos autores

A Figura 11 apresenta os dados analisados no Estado do Mato Grosso do Sul e seus municípios, referentes à quantidade de internações por causas externas do SUS. Pode-se notar que os municípios com mais internações foram Campo Grande-MS (na cor preta) com 10.202 internações no ano de 2015 e, em seguida, Dourados-MS (na cor vermelha) com 2.114 internações no ano de 2015. Neste mapa o usuário poderá escolher o Estado, data inicial e final e o conteúdo desejado.

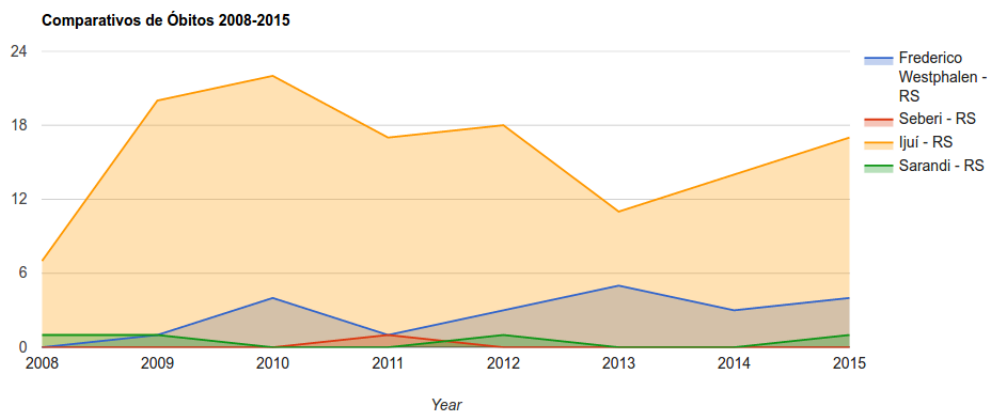
Figura 11. Mapa Indicador de Municípios



Fonte: dos autores

A Figura 12 apresenta um gráfico comparativo de óbitos entre quatro municípios, considerando causas externas do SUS, entre 2008 e 2015, indicando a quantidade de óbitos. Na vertical, mostra-se a quantidade de óbitos dos municípios e, no eixo da horizontal, o período entre 2008 a 2015. De acordo com o gráfico, o município de Ijuí apresentou, durante o período amostrado, uma quantidade de óbitos maior que as outras três cidades apresentadas.

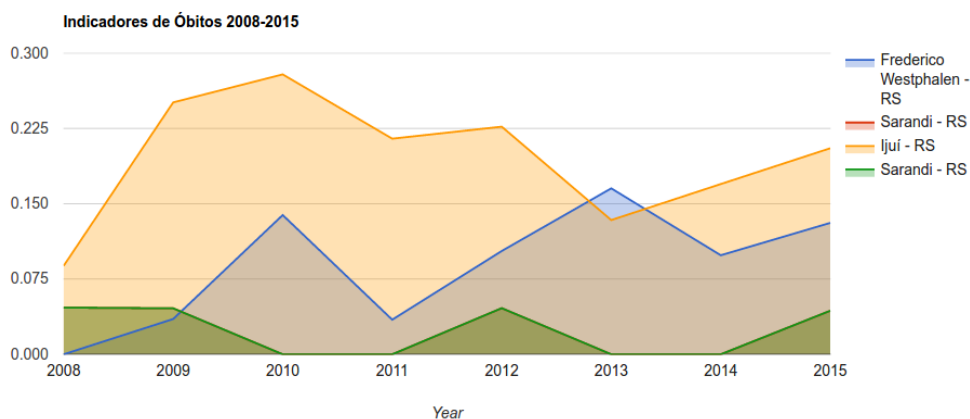
Figura 12. Gráfico Comparativo de Óbitos



Fonte: dos autores

A Figura 13 apresenta um gráfico indicador de óbitos com quatro municípios por causas externas do SUS, entre 2008 e 2015, indicando a quantidade de óbitos por mil habitantes. Na vertical mostra-se o índice de óbitos dos municípios por mil habitantes e, no eixo da horizontal, o período entre 2008 a 2015.

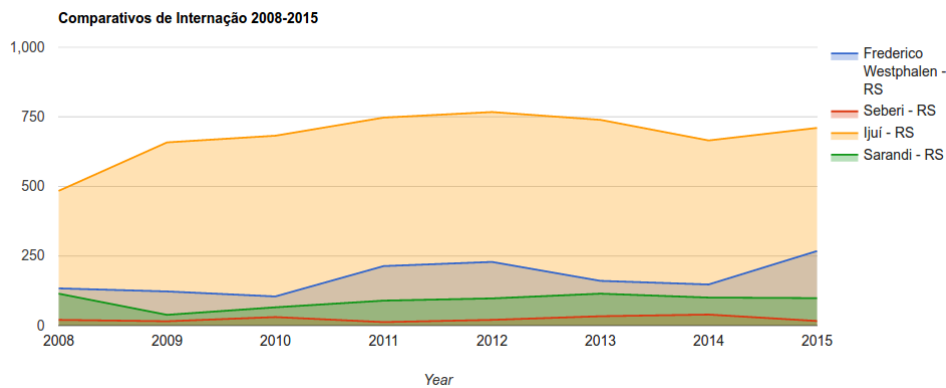
Figura 13. Gráfico Indicador de Óbitos



Fonte: dos autores

A Figura 14 apresenta um gráfico comparativo de óbitos com quatro municípios por causas externas do SUS, entre 2008 e 2015, indicando a quantidade de internações. Na vertical, mostra-se a quantidade de internações dos municípios e, no eixo da horizontal, o período entre 2008 a 2015.

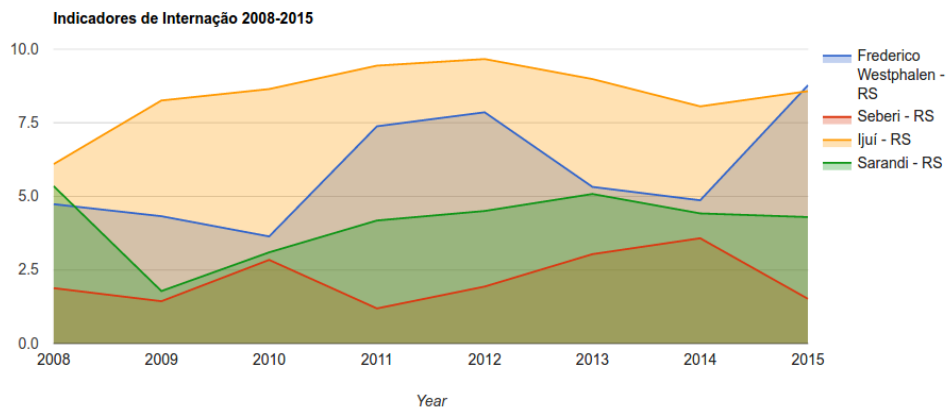
Figura 14. Gráfico Comparativo de Internações



Fonte: dos autores

A Figura 15 apresenta um gráfico indicador de internações com quatro municípios por causas externas do SUS, entre 2008 e 2015, indicando a quantidade de internações por mil habitantes. Na vertical, mostra-se o índice de internações dos municípios por mil habitantes e, no eixo da horizontal, o período entre 2008 a 2015.

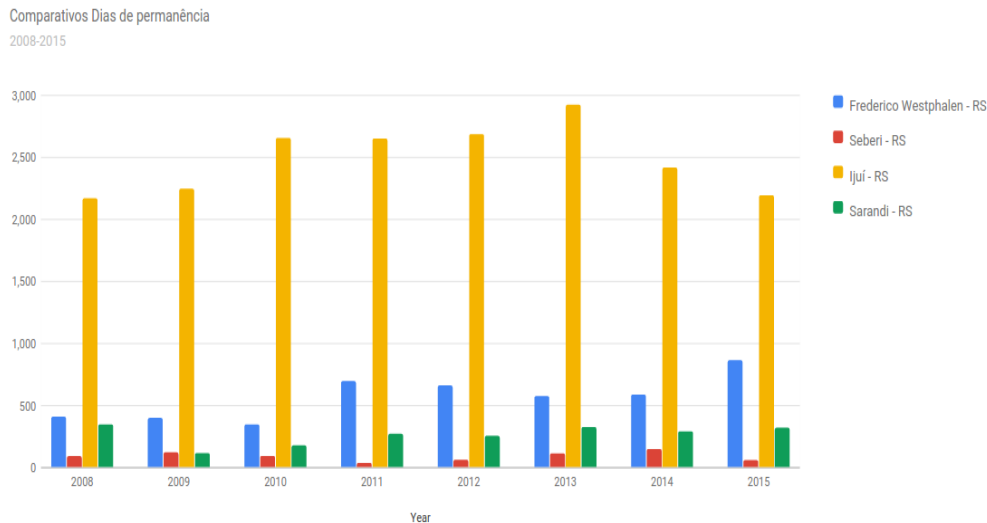
Figura 15. Gráfico Indicador de Internações



Fonte: dos autores

A Figura 16 apresenta um gráfico comparativo dos dias de permanência hospitalar com quatro municípios por causas externas do SUS, entre 2008 e 2015. Na vertical, mostra-se a quantidade de dias permanecidos no hospital dos municípios e, no eixo da horizontal, o período entre 2008 a 2015.

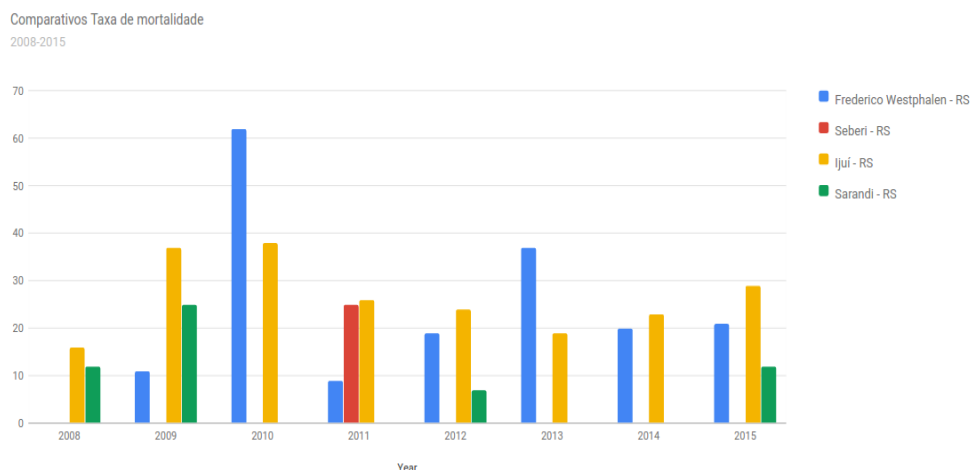
Figura 16. Gráfico Comparativo – Dias de Permanência Hospitalar



Fonte: dos autores

A Figura 17 apresenta um gráfico comparativo da taxa de mortalidade com quatro municípios por causas externas do SUS, entre 2008 e 2015, indicando a quantidade de óbitos e o número de AIHs aprovadas, computadas como internações no período, multiplicado por cem. Na vertical, mostra-se a taxa de mortalidade e, no eixo da horizontal, o período entre 2008 a 2015.

Figura 17. Gráfico Comparativo – Taxa de Mortalidade referente à quantidade de óbitos e o número de AIHs aprovadas



Fonte: dos autores

No sistema desenvolvido o usuário pode realizar diversas consultas aos dados do DATASUS de qualquer município ou Estado do Brasil. A análise de comparativos e

indicadores é visualizada em gráficos, sendo que todos os dados dos indicadores foram trabalhados com um índice de mil habitantes. Na área de mapas pode-se visualizar os Estados e municípios.

5. CONSIDERAÇÕES FINAIS

O presente artigo teve, como enfoque principal, o estudo de bases de dados abertas na área da saúde. do SUS, extraindo os dados do *site* DATASUS, convertendo-os em uma forma mais prática e atrativa para a sua interpretação, conseguindo assim demonstrar de forma gráfica como ocorreu a evolução da morbidade hospitalar do SUS por causas externas, permitindo identificar em que locais houve maior uso deste serviço. Os dados disponibilizados pelo DATASUS são mensalmente atualizados. Percebe-se que os dados, mesmo sendo atualizados, são fornecidos de uma forma que não favorece comparações e análises, já que os mesmos são disponibilizados no formato CSV.

A utilização de dados abertos com informações integradas a ferramentas para visualização facilita o dia-a-dia dos cidadãos e da sociedade. Acessando os mapas e gráficos, como os apresentados pela ferramenta desenvolvida, o usuário tem acesso às informações com mais clareza e visibilidade, sem que seja necessário o acesso a tabelas e arquivos em forma de texto. Isso permite ao usuário a visualização de informações com mais facilidade, possibilitando uma maior compreensão das mesmas.

É importante frisar que a principal contribuição deste trabalho foi a migração dos dados obtidos do DATASUS para um banco de dados no SGBD *MySQL*. Esta base de dados poderá ser reutilizada por outros projetos, podendo trazer ainda mais benefícios à população brasileira, permitindo que o cidadão economize tempo na busca de informações, interpretando os dados de uma forma mais atrativa e significativa. Os gráficos facilitam a análise e interpretação de um conjunto de dados, o que pôde ser comprovado neste trabalho. Relacionando os dados obtidos com o número de habitantes, pode-se verificar a real situação do SUS com relação à morbidade hospitalar por causas externas.

A aplicação *web* desenvolvida poderia ser estendida para usar mais formas estatísticas e implementar algum método para atualizar a base de dados mensalmente, podendo servir como apoio a decisões em âmbito governamental.

Como trabalhos futuros, serão realizados os refinamentos e ajustes na aplicação desenvolvida, visando aproveitar o máximo possível do banco de dados, além de abordar a implementação de uma abordagem estatística dos dados disponíveis.

REFERÊNCIAS

BRASIL (2011). Lei nº 12.527, de 18 de novembro de 2011. Disponível em: <http://www.presidencia.gov.br/ccivil_03/_Ato2011-2014/2011/Lei/L12527.html>. Acesso em maio, 2018.

- CANBAN, D. G. (2015). Visual Analytics in Healthcare: opportunities and research challenges. Oxford University - JAMIA. Disponível em: <http://www.skateboardingalice.com/papers/2015_Caban.pdf>. Acesso em novembro, 2016.
- DADOS.GOV.BR (2018). Portal Brasileiro de Dados Abertos. Disponível em: <<http://dados.gov.br>>. Acesso em abril, 2018.
- DATASUS (2018). Portal da saúde. Disponível em: <<http://www2.datasus.gov.br/DATASUS/index.php>>. Acesso em abril, 2018.
- DEVELOPERS.GOOGLE.COM (2018). Display Live Data on your Site. Disponível em: <<https://developers.google.com/chart/>>. Acesso em março, 2018.
- FREITAS, C. M. D. S. F. (2001). Introdução à Visualização de Informações. RITA, pages 2–16. Disponível em: <<https://goo.gl/uJ19kJ>>. Acesso em novembro, 2017.
- GERMANO, P. G. F. (2008). Sistema de informações sobre mortalidade: estudo de caso em municípios com precariedade dos dados. SCPRJ, pages 2–10. Disponível em: <<http://www.scielo.br/pdf/csp/v24n10/07.pdf>>. Acesso em novembro, 2017.
- LIBREOFFICE.ORG (2018). Formatos de Arquivos XML. Disponível em: <https://help.libreoffice.org/Common/XML_File_Formats/pt-BR>. Acesso em abril, 2018.
- MYSQL.COM (2018). MySQL. Disponível em: <<https://www.mysql.com/>>. Acesso em março, 2018.
- OKFN.ORG (2018). Open knowledge. Disponível em: <<https://okfn.org/>>. Acesso em abril, 2018.
- OPENDATADAY.ORG (2018). International Open Data Hackathon. Disponível em: <<http://opendataday.org/>>. Acesso em novembro, 2017.
- OPENDATAMANUAL.ORG (2018). Manual of Open Data. Disponível em: <http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf>. Acesso em novembro, 2017.
- PACHEVSASHA, S. P. (2007). Understanding MySQL Internals. O'Reilly. Disponível em: <<http://bootski.net/books/Technical.Ebooks.Pack.2/OReilly%20Understanding%20MySQL%20Internals.pdf>>. Acesso em novembro, 2017.
- PERREIRA, C. R. P. (2012). Elementos para Construção de um Modelo de Análise de Iniciativas de Publicações de Dados Governamentais Abertos. EACH. Disponível em: <<https://goo.gl/CF7DSY>>. Acesso em novembro, 2017.
- RIBEIRO, C. J. S. R. (2011). Dados Abertos Governamentais (open government data): Instrumento para exercício da cidadania pela sociedade. Enancib. Disponível em: <<https://core.ac.uk/download/pdf/11890148.pdf>>. Acesso em agosto, 2017.

- SYMPHONY, I. L. (2015). Abrindo e salvando arquivos CSV de texto. Disponível em: <http://infolib.lotus.com/resources/symphony/3.0.0/sym20abd014/pt_br/text/scalc/guide/csv_files.html>. Acesso em agosto, 2017.
- TRAINA, A. J. M. T. (2001). Dados Governamentais Abertos no Brasil. USP. Disponível em: <<http://www.lbd.dcc.ufmg.br/colecoes/sbbd/2001/007.pdf>>. Acesso em agosto, 2017.
- W3.ORG (2018). W3c RDF. Disponível em: <<https://www.w3.org/>>. Acesso em março, 2018.
- W3C.BR (2018). W3C Brasil. Disponível em: <<http://www.w3c.br/Home/WebHome>>. Acesso em março, 2018.
- ZACARIAS, I., VITALLI, R. A., PREDIGER, D., MOERSCHBACHER, J., BERTOLINI, C. (2015). Análise Comparativa dos Acidentes de Trabalho no Brasil a partir de Dados Abertos. Anais do Computer on the Beach. Disponível em: <<http://siaiap32.univali.br/seer/index.php/acotb/article/view/7038>>. Acesso em agosto, 2017.