

DASHCARE: Um Dashboard para Análise dos Acidentes nas Rodovias Federais

Francisco Lucas L. Nunes¹, Samara Martins Nascimento¹, André Luiz Sena da Rocha¹

¹Universidade Federal do Semi-Árido (UFERSA)
Pau dos Ferros – RN – Brasil

{lusca.lima@outlook.com, {samara.nascimento, andre.rocha}@ufersa.edu.br}

Abstract. *The large number of accidents, which are observed on the federal highways of Brazil, brings a warning about the need changes in the infrastructure of the roads. For this, it is necessary to carry out a decision-making process, which can directly interfere with the type and time of displacement of the population. To assist this process, it is possible to build strategic tools that facilitate the visualization of the main problems of the roads and indicate, punctually, where new changes should be made. This work proposes the construction of a Dashboard, which shows the accidents of federal highways in a clear and simple way, which were obtained from data provided by the Federal Highway Police, using the strategy of Big Open Data.*

Resumo. *O grande número de acidentes, que são observados nas rodovias federais do Brasil, traz um alerta sobre a necessidade de mudanças na infraestrutura das estradas. Para isso, é necessário realizar um processo de tomada de decisões, que pode interferir diretamente no tipo e tempo de deslocamento da população. Para auxiliar este processo, é possível construir ferramentas estratégicas que facilitem a visualização dos principais problemas das rodovias e indiquem, pontualmente, onde devem ser realizadas novas mudanças. Este trabalho propõe a construção de um dashboard, que mostra os acidentes das rodovias federais de forma clara e simples, os quais foram obtidos a partir de dados disponibilizados pela Polícia Rodoviária Federal, utilizando a estratégia de Big Open Data.*

1. Introdução

O meio de transporte rodoviário configura-se como o maior e mais importante meio de locomoção, que serve desde fins de transporte pessoal até fins econômicos [Schmidt 2011]. Frente a isso, é válida a ótica de que se deve reforçar a ideia sobre o aprimoramento e a manutenção da segurança no trânsito. Nesse cenário, as Rodovias Federais (RFs) atuam como uma das principais vias para deslocamento dos automóveis em todo o país. É de responsabilidade da Polícia Rodoviária Federal (PRF) o levantamento das causas e dados acerca dos acidentes que ocorrem em uma RF [Brasil 2019].

Grande parte dos registros sobre os dados coletados a partir dos levantamentos realizados pela PRF são disponibilizados em uma base de Dados Abertos (DAs) [PRF 2019]. Os DAs são dados institucionais sem restrições ou licença, disponibilizados publicamente,

que fomentam inovação e crescimento econômico, quando aplicadas as devidas Tecnologias da Informação (TI) [Zuiderwijk 2014]. Dentro desse contexto, é válido citar o *Big Data*, que corresponde a uma estratégia computacional que lida, em suma, com a variabilidade dos dados e seu crescimento escalável, aliado à velocidade na mudança das informações. Como os DAs mudam continuamente e seu volume cresce na mesma velocidade, é possível indicar a necessidade de manipular informações, cedidas pela PRF, a partir do *Big Open Data*.

O presente trabalho tem como objetivo propor o desenvolvimento de um *dashboard*, chamado *DASHCARE: Um Dashboard para Análise dos Acidentes nas Rodovias Federais*, que foi construído a partir de tecnologias Web e conta com a persistência dos dados usando um banco de dados NoSQL. O DASHCARE monitora, analisa e exibe informações gerais sobre os acidentes, que ocorrem nas rodovias do Brasil. Essa informação tende a mudar continuamente, dadas atualizações que podem ser realizadas e cada mudança deve indicar a ocorrência de novos acidentes.

2. Contextualização

Dados divulgados pelo G1 (2018), revelam que, apesar do número total de acidentes ter diminuído ao longo do intervalo entre os anos de 2007 a 2017, o índice de ocorrências continua alto, e que ainda nesse intervalo, houve um pico de acidentes em 2011, com cerca de 192 mil casos. No ano de 2014, foi estimado que cada acidente em uma RF custou ao país a média de R\$ 75 mil, valor que aumenta significativamente no que se refere a casos de acidente acompanhado de óbito [Andrade 2019]. Mais adiante, no ano de 2017, foi contabilizado aproximadamente 89 mil ocorrências em RFs, dentre as quais, cerca de 6 mil resultaram em fatalidade [G1 2018]. Esses dados demonstram que, apesar da considerável queda do número de acidentes nas RFs, ainda existe um alto número de acidentes que transtorna o meio social, como também onera sobre os gastos com saúde pública no país.

Com a aplicação da Tecnologia da Informação (TI), há indícios da construção de sistemas, que garantem uma essencial aplicabilidade de soluções informatizadas de análise, organização e apresentação dos dados em tempo real [Kopetz 2011], os quais computam *streams*¹. Sendo assim, é possível, com a solução adequada, gerenciar e validar as informações dos dados levantados pela PRF das RFs, servindo como estratégia norteadora para o processo de tomadas de decisões, buscando o aumento na segurança das estradas nacionais.

A quantidade dos dados acumulados ao longo do tempo, por órgãos como a PRF, é ligeiramente grande. Por exemplo, se fossem convertidos em uma planilha, essa quantidade chegaria a conter centenas de milhares de linhas. Segundo Marquesone (2016), a grande limitação para se construir sistemas com processamento em tempo real é o tempo de processamento, quando se considera grandes volumes de dados. Assim, existe uma limitação quanto ao processamento dos DA disponibilizados pela PRF. Uma possível solução seria a de construir sistemas escaláveis, que podem mudar continuamente e ainda garantir que o desempenho do mesmo não seja perdido, dado o recebimento de novas *streams*.

¹Segundo [Aggarwal 2007], *streams* de dados é a emissão e coleta de dados de modo contínuo, podendo ser gerados a partir de várias fontes distintas.

Para garantir a construção de sistemas que computam *streams* de dados continuamente, é comum a otimização do processamento por meio de programação paralela, que divide o processamento em várias máquinas utilizando uma abordagem distribuída [Chen 2014]. Essa estratégia garante o processamento com baixo tempo de resposta e alta disponibilidade [Marquesone 2016].

3. Estado da Arte

Na literatura, é possível encontrar diferentes propostas de *dashboards* que permitem a visualização sumarizada de um conjunto de dados. A proposta de Joazeiro (2015) foi o desenvolvimento de um *dashboard* para análise de *Business Intelligence* (BI), dentro de um contexto governamental. Foi realizado um estudo de caso sobre a Secretaria de Estado do Mato Grosso, onde aplicou técnicas de BI para a sistematização de estratégias que compreendem os valores, impacto social e processos internos da secretaria [Joazeiro 2015]. Entretanto, sua proposta não faz menção quanto à escalabilidade, dado o aumento do volume de dados.

Santos (2017) propôs um painel de controle que visa monitorar a qualidade dos *softwares* para órgãos públicos brasileiros. A solução tem o foco na simplicidade e objetividade das informações, tendo em vista que grande parte dos órgãos públicos terceirizam o serviço de *software* [Santos 2017]. Já o estudo de Silva (2018) visa a criação de um *dashboard* que emprega técnicas de BI buscando o levantamento de aspectos importantes, que sirvam de indicadores para o segmento do agronegócio. Esta aplicação trata de uma plataforma Web, que além das tecnologias base, suporta estratégias de manipulação e agrupamento de dados [Silva 2018].

Os trabalhos desenvolvidos por Joazeiro (2015), Santos (2017) e Silva (2018) estão relacionados com este, pois todos visam a construção de um *dashboard*, que oferece suporte ao processo de tomadas de decisões, as quais são baseadas em dados. No entanto, os estudos mostrados não focam na criação de ferramentas para trabalho com grandes volumes de dados ou DAs, como o *DASHCARE*, cuja solução será mostrada a seguir.

4. DASHCARE: Um Dashboard para Análise dos Acidentes nas Rodovias Federais

Em uma visão macro, a aplicação proposta é uma interface de visualização da informação extraída a partir dos DAs dos acidentes das RFs. Na Figura 1, é possível identificar a existência de duas áreas: (i) Módulos preexistentes; e (ii) *Dashboard* proposto, na qual foi concentrado o esforço deste trabalho.

A área de “Módulos preexistentes” mostra como seria o possível plano de fundo para qualquer serviço, que queira consumir informações a partir da proposta deste trabalho. Nessa área, é possível encontrar quatro módulos: (i) geradores de dados, que são as fontes responsáveis por produzir os dados, os quais devem ser analisados; (ii) *stream collector*, que é responsável pela coleta de dados gerados por fontes específicas em tempo real, a exemplo do GPS (*Global Positioning System*); (iii) dados acidentes, que corresponde ao conjunto de dados relacionados, obtidos a partir do módulo anterior; e (iv) o módulo ETL (*Extract, Transform, Load*), que integra os dados originários de diversas fontes de acordo com as regras de negócio da entidade utilitária e realiza o processo de limpeza dos mesmos.

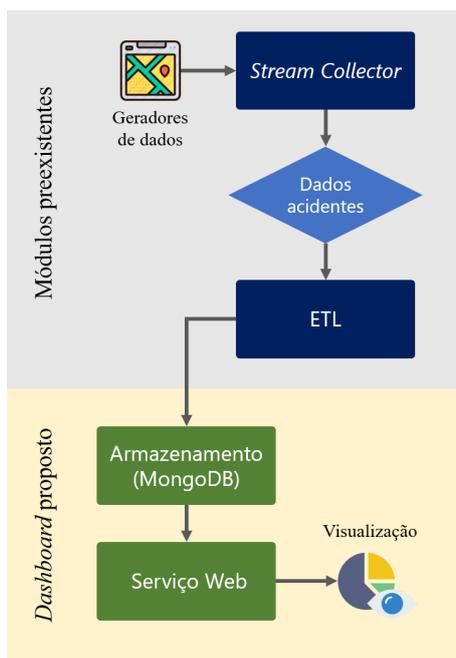


Figura 1. Arquitetura simplificada do sistema

A área de “*Dashboard* proposto” corresponde ao escopo deste trabalho. Nela é projetada a existência de três módulos: (i) Armazenamento, que disponibiliza uma API (*Application Programming Interface*) que se conecta com o módulo ETL e persiste os dados no Banco NoSQL MongoDB; (ii) Serviço Web, que culmina no *Dashboard* propriamente dito, onde ocorre o processamento, análise e seleção de relações entre os dados; e (iii) Visualização, onde o usuário pode interagir com a ferramenta, requisitando consultas pelos parâmetros desejados e recebendo as informações em elementos gráficos (e. g. tabelas e gráficos).

Os DAs fornecidos pela PRF são contidos em formato CSV (*Comma-Separated Values*) e não em uma base de dados única. Esses dados, considerados semi-estruturados, foram importados para o Banco NoSQL MongoDB, porque esse banco suporta o armazenamento de novos documentos, independente da variabilidade da informação obtida. A estrutura do servidor atua como intermediário entre o módulo de Visualização e Armazenamento. O objetivo do servidor é receber as requisições solicitadas e realizar uma consulta ao banco de dados, enviando uma resposta ao módulo de visualização. Este módulo deve, além de mostrar as informações contidas nos dados ao usuário, ser capaz de transformar comandos de maior abstração (maior granularidade dos dados), em comandos de mais baixo nível (menor granularidade dos dados, indicando uma maior riqueza de detalhe). A abordagem utilizada no *DASHCARE* remove do lado servidor o ônus de renderizar os elementos (deixando totalmente a cargo do navegador) e enviar uma nova página a cada requisição. Assim, o servidor ganha espaço de processamento, focando apenas na manipulação dos dados.

5. Resultados e discussões

O *DASHCARE* foi desenvolvido para monitorar e visualizar as informações contidas nos DAs da PRF acerca dos acidentes no Brasil. Assim, a primeira requisição ao ser exibida

a primeira tela no navegador é a de mostrar as informações dos acidentes em um escopo nacional. Esse estágio inicial do *DASHCARE* é ilustrado pela Figura 2. A figura mostra, na parte superior da tela inicial, três componentes principais, os quais são listados a seguir:

1. Barra de navegação: nela é possível ter o nome da região em foco, que no caso dessa tela é o “Brasil” (indicando o mais alto nível de granularidade dos dados). Além disso, são encontradas duas caixas de seleção, a primeira define o foco para algum estado e a segunda seleciona o mês em que os dados devem ser buscados.
2. Mapa do Brasil: que é um componente responsável por indicar o mapa político definido pelos estados. Nesse caso, cada estado tem uma área que permite a seleção, indicando a capacidade da ferramenta diminuir o nível de granularidade do dado (mostrando os acidentes por estado);
3. *Ranking* BRs: este componente mostra uma tabela que ordena, de forma decrescente, um *ranking* de nove BRs, com os maiores números de vítimas em absoluto, por gravidade dos ferimentos e óbitos.



Figura 2. Tela inicial do *DASHCARE* (superior)

A Figura 3 revela a parte inferior da tela inicial no *DASHCARE*. Ela contém nove componentes que plotam gráficos de acordo com o campo especificado, todos eles são explicados a seguir:

1. Acidentes por causa: nesse gráfico é possível ter acesso aos dados sobre os acidentes de acordo com a causa (ações relacionadas ao condutores dos veículos);
2. Acidentes por região: esse gráfico mostra o número de acidentes por região, sendo ele exclusivo dessa tela;
3. Acidentes por veículo: esse gráfico mostra quais os veículos mais envolvidos com acidentes no Brasil;
4. Acidentes por sexo: mostrando quais os sexos mais presentes nos acidentes;
5. Acidentes por mês: indicando o número de acidentes em ordem cronológica dos meses. Nesse caso, quando existe algum mês selecionado, o componentes passam a mostrar a ordem cronológica dos dias do mês escolhido;

6. Acidentes por tipo: esse gráfico mostra quais os tipos (forma de acidente conduzida pelo veículo) de acidentes mais comuns no Brasil.



Figura 3. Tela inicial do **DASHCARE** (inferior)

A Figura 4 mostra as informações relacionadas a um único estado, quando o mesmo é selecionado diretamente no mapa do Brasil ou é escolhido a partir da barra de navegação. É possível encontrar quase todos os mesmos componentes da tela inicial, porém alguns gráficos mostrados neste nível de visualização são novos, como:

1. Mapa: esse componente renderiza um mapa, que plota os pontos das coordenadas de cada acidente ocorrido, considerando o estado selecionado;
2. Barra de navegação: esse componente, quando renderizada uma tela de estado, ganha uma terceira caixa de seleção, indicando a possibilidade de visualizar os acidentes por município;
3. *Ranking* dos municípios: indicando os cinco municípios, do estado selecionado, que têm o maior número de acidentes;
4. Acidentes por dia da semana: esse gráfico revela quais os dias da semana com maior número de acidentes ocorridos. Esse componente é exclusivo para as telas de estado e município.

A partir de um estado, é possível ter acesso aos acidentes que ocorreram em seus municípios. A Figura 5 mostra as alterações ocorridas quando é selecionado um município específico. Dentre as mudanças existentes, considerando este nível de visualização, é possível citar:

1. A presença de uma Tabela: que corresponde a um componente que mostra os dados agrupados por município;
2. Acidente por dia: esse componente é o mesmo que “Acidentes por mês”, mudando apenas os dados passados ao gráfico e o título;
3. Mapa: mostrando as coordenadas referentes ao município selecionado.

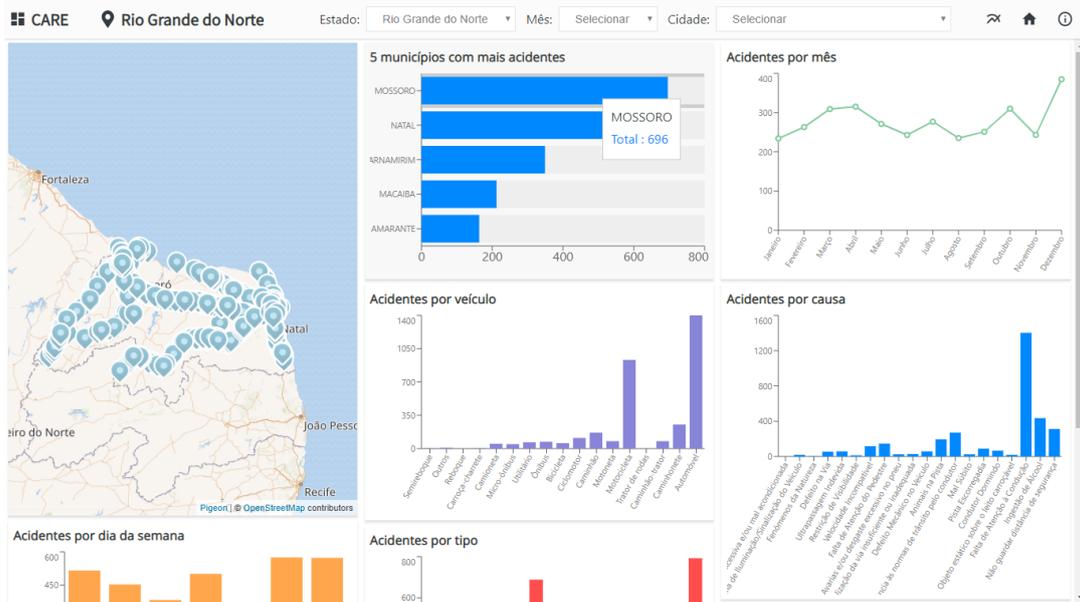


Figura 4. Tela por estado no *DASHCARE*

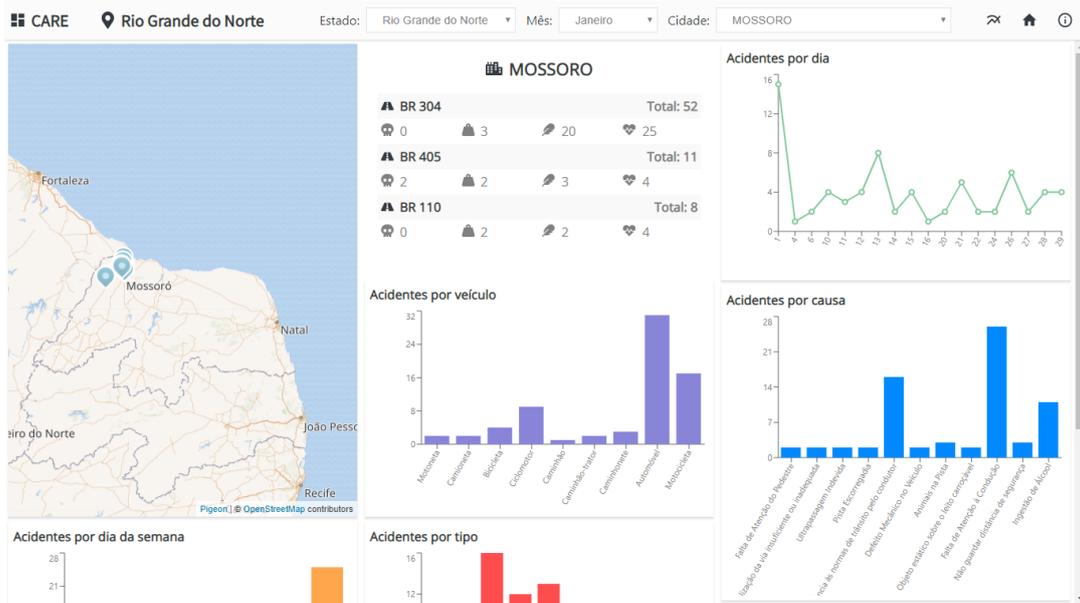


Figura 5. Tela por município de um estado do *DASHCARE*

Os componentes do *DASHCARE* foram projetados para encapsularem a lógica de comunicação com o servidor. Assim, cada componente, ao ser renderizado, faz uma requisição internamente, requerendo os dados específicos de seu domínio.

6. Conclusão

Este trabalho propôs um novo *dashboard*, chamado *DASHCARE*, que visa contribuir para um melhor entendimento dos DA da PRF. Ademais, indica uma colaboração para auxiliar no processo de tomadas de decisões, em soluções que possam vir a combater o alto índice de acidentes. Para resolução da proposta, foi pensado em um processo que envolveu o

desenvolvimento de três módulos: um para armazenamento dos dados, outro para atuar como um servidor e, um último, que permite a visualização das informações. O modelo básico de sua arquitetura seguiu uma estrutura baseada para tratamento de fluxos de dados. Como limitação, não é possível atestar sua total eficiência, pois não foi possível implantar a solução em um ambiente de utilização real, indicando o processamento de *stream* de dados. Para trabalhos futuros planeja-se criar outras modalidades para visualização dos dados e utilizar técnicas estatísticas para correlação das informações obtidas. Ademais, é pensada a construção de um módulo para mineração de dados sobre o universo de discurso abordado.

Referências

- Aggarwal, C. C. (2007). *Data streams: models and algorithms*, volume 31. Springer Science & Business Media.
- Andrade, Flávia Reis de e Antunes, J. L. F. (2019). Tendência do número de vítimas em acidentes de trânsito nas rodovias federais brasileiras antes e depois da década de ação pela segurança no trânsito. *Cadernos de Saúde Pública*, 35:e00250218.
- Brasil (2019). Institui o código de trânsito brasileiro.
- Chen, CL Philip e Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on big data. *Information sciences*, 275:314–347.
- G1 (2018). Acidentes em rodovias federais mataram mais de 83 mil pessoas no Brasil em 10 anos.
- Joazeiro, S. A. (2015). Desenvolvimento de dashboard dinâmico de projetos utilizando o processo de business intelligence sobre um sistema de solicitação de serviço em um ambiente governamental.
- Kopetz, H. (2011). *Real-time systems: design principles for distributed embedded applications*. Springer Science & Business Media.
- Marquesone, R. (2016). *Big Data: Técnicas e tecnologias para extração de valor dos dados*. Editora Casa do Código.
- PRF (2019). Dados abertos.
- Santos, L. M. d. (2017). Criação de um dashboard para monitoramento de perfis de qualidade de software.
- Schmidt, E. L. e. o. (2011). O sistema de transporte de cargas no Brasil e sua influência sobre a economia.
- Silva, R. F. d. (2018). Desenvolvimento de dashboard para análise de dados de agropênia. B.S. thesis, Universidade Tecnológica Federal do Paraná.
- Zuiderwijk, Anneke e Janssen, M. e. C. S. e. M. R. (2014). Design principles for improving the process of publishing open data. *Transforming Government: People, Process and Policy*, 8(2):185–204.