

**LUCAS SEIJI SAITO
MARCOS FELIPE NUNES LINO RIBEIRO**

**VISUALIZAÇÃO DE DADOS PROCESSADOS
POR MACHINE LEARNING PARA
MANUTENÇÃO DA EFVM**

São Paulo
2021

LUCAS SEIJI SAITO
MARCOS FELIPE NUNES LINO RIBEIRO

**VISUALIZAÇÃO DE DADOS PROCESSADOS
POR MACHINE LEARNING PARA
MANUTENÇÃO DA EFVM**

Trabalho apresentado à Escola Politécnica
da Universidade de São Paulo para obtenção
do Título de Engenheiro de Computação.

São Paulo
2021

**LUCAS SEIJI SAITO
MARCOS FELIPE NUNES LINO RIBEIRO**

**VISUALIZAÇÃO DE DADOS PROCESSADOS
POR MACHINE LEARNING PARA
MANUTENÇÃO DA EFVM**

Trabalho apresentado à Escola Politécnica
da Universidade de São Paulo para obtenção
do Título de Engenheiro de Computação.

Área de Concentração:

Engenharia de Computação

Orientador:

Pedro Luiz Pizzigatti Correa

Co-orientadores:

Oswaldo Gogliano Sobrinho

Wellington Dias de Queiroz

São Paulo
2021

RESUMO

A Estrada de Ferro Vitória a Minas é uma via muito importante para o Brasil. Ela transporta cerca de 1 milhão de pessoas ao ano e 135 milhões de toneladas de carga. Em vista destes dados, é de extrema importância que a via não pare de funcionar devido a acidentes ou frequentes manutenções. Pensando nisso, a Vale, em parceria com a USP, está fazendo um projeto de ciência dos dados que visa implementar aplicações para manutenção preditiva da via, assim evitando acidentes e diminuindo o tempo e os custos com manutenções. Neste projeto é importante que os operadores possam ver facilmente os trechos da via ou os equipamentos que precisam de algum tipo de reparo. Com isto em mente, este trabalho terá o enfoque de criar uma visualização dos dados adequada ao contexto de manutenção preditiva da EFVM.

Palavras-Chave – Visualização de dados, Manutenção preditiva em ferrovias, EFVM.

ABSTRACT

The Vitória-Minas Railway is a very important road for Brazil. It transports about 1 million people a year and 135 million tons of material. Given these data, the road must not stop working due to accidents or frequent maintenance. With this in mind, Vale, in partnership with USP, is carrying out a data science project that aims to implement applications for predictive maintenance on the road, thus avoiding accidents and reducing maintenance time and costs. In this project, it is important that operators can easily see the stretches of rail or equipment that need some repair. With this in mind, this work will focus on creating a data visualization suitable for the EFVM predictive maintenance context.

Keywords – Data Visualization, Railroad Predictive Maintenance, EFVM.

LISTA DE FIGURAS

1	Mapa da Estrada de Ferro Vitória Minas	13
2	Alargamento e estreitamento de bitola	18
3	Desnívelamento longitudinal dos trilhos	18
4	Desnívelamento transversal dos trilhos	19
5	desenho esquemático do empeno	20
6	desenho esquemático do alinhamento	20
7	Django	23
8	Sphinx	23
9	Mapbox	24
10	React	24
11	Bootstrap	25
12	ChartJs	26
13	Turf	26
14	Diagrama da arquitetura final	27
15	Workflow de um projeto de Big Data Analytics	29
16	Tela inicial da aplicação	40
17	Painel de coordenadas do mouse	40
18	Switch de RH's ligado	41
19	Quatro estilos alternativos de mapa	41
20	Vista panorâmica inicial	42
21	Seleção de um trecho da ferrovia	43
22	Painel de estatísticas	44
23	Mini mapa para visualização automática de trecho selecionado	45

LISTA DE TABELAS

1	Requisitos funcionais	33
2	Requisitos não funcionais	34

LISTA DE SIGLAS

- ABNT** Associação Brasileira de Normas Técnicas. 17
- ANTT** Agência Nacional de Transportes Terrestres. 12
- AREMA** American Railway Engineering and Maintenance-of-Way Association. 17
- BDA** Big Data Analytics. 12, 13, 29, 30
- BDRA** Big Data Risk Analysis. 20–22, 31, 48, 53
- CC** Carro Controle. 12, 13, 50
- DOM** Document Object Model. 24, 39, 40
- EFVM** Estrada de Ferro Vitória Minas. 11–13, 31, 32, 37, 42, 47
- HTML** HyperText Markup Language. 23, 25
- ML** Machine Learning. 11, 20, 27, 50, 53
- MVC** Model-View-Controller. 23
- PDF** Portable Document Format. 23
- RH's** Reference Houses. 40–42, 47
- SPA** Single-Page Application. 24, 39
- SVG** Scalable Vector Graphics. 25, 26
- VI** Vagão Instrumentado. 11–13, 50

SUMÁRIO

Capítulo 1: INTRODUÇÃO	10
1.1 Objetivo	11
1.2 Motivação	12
1.3 Justificativa	14
Capítulo 2: REVISÃO BIBLIOGRÁFICA	16
2.1 Aspectos Conceituais	17
2.1.1 Defeitos da Geometria da Via Permanente	17
2.1.1.1 Alargamento ou estreitamento de bitolas	17
2.1.1.2 Desnivelamento Longitudinal	17
2.1.1.3 Desnivelamento Transversal	18
2.1.1.4 Empenamento	18
2.1.1.5 Alinhamento	19
2.1.2 Visualização de Big Data Risk Analysis	20
2.1.2.1 Análise Visual	21
2.1.2.2 Análise de dados	21
2.1.2.3 Storytelling	22
2.2 Tecnologias Utilizadas	23
2.2.1 Django	23
2.2.2 Sphinx	23
2.2.3 Mapbox	24

2.2.4 React	24
2.2.5 Bootstrap	25
2.2.5.1 Bootstrap Icons	25
2.2.6 ChartJs	25
2.2.7 Turf	26
2.3Arquitetura final	27
Capítulo 3: ESPECIFICAÇÃO/PROJETO	28
3.1Fase de estudos	29
3.1.1 Conceituação do problema	29
3.1.2 Contexto externo	29
3.2Projeto	31
3.3Requisitos Funcionais	32
3.4Requisitos não funcionais	34
Capítulo 4: IMPLEMENTAÇÃO	35
4.1Estrutura de diretórios	36
4.2Backend	38
4.2.1 Documentação	38
4.3Frontend	39
4.3.1 Página Principal	39
4.3.2 CSS	39
4.3.3 Componentes	39
4.3.3.1 App	40

4.3.3.2 Map	41
4.3.3.3 MapSelector	42
4.3.3.4 Panel	43
4.3.3.5 MapStatistics	43
4.3.3.6 MiniMap	45
Capítulo 5: DISCUSSÃO DOS RESULTADOS	46
5.1 Resultados	47
Capítulo 6: CONSIDERAÇÕES FINAIS	49
6.1 Trabalhos Futuros	50
6.2 Conclusão	52
6.2.1 Contribuições	52
Referências Bibliográficas	54
Referências Bibliográficas	54

CAPÍTULO 1

INTRODUÇÃO

1.1 OBJETIVO

Com o intuito de adotar ferramentas de Big Data na rotina diária de trabalho na área de Engenharia Ferroviária da Vale, a empresa “está desenvolvendo um projeto voltado para análise de dados da via permanente para otimização da gestão de inspeção e de manutenção”(1). Para tanto, está sendo implementada uma aplicação que, por meio de algoritmos de ML, processa informações coletadas pelo Vagão Instrumentado para caracterização das condições de infraestrutura da Estrada de Ferro Vitória Minas (EFVM). Nesse contexto, este trabalho visa a elaboração de um sistema que forneça ao usuário um meio intuitivo para utilizar, na prática, as informações funcionais extraídas desse processamento.

Para que isso seja executado da melhor maneira possível, um background consistente dos conceitos e aplicações de visualização de dados e storytelling são necessários. O primeiro possibilita a escolha dos componentes gráficos mais adequados para determinado objetivo, levando em consideração o tipo da informação, as finalidades/interações almejadas e os conhecimentos técnicos do usuário final. O segundo, por sua vez, consolida uma forma eficiente de mostrar como ou por que as informações são relevantes, a maneira como elas interagem entre si e a mudança que elas sofrem dentro de um período. Sendo assim, também é importante que este trabalho implemente mecanismos de visualização embasados na literatura existente sobre esses dois conceitos do processo de Big Data.

A aplicação utiliza características inferidas das condições de infraestrutura da via permanente para fornecer e visualizar informações úteis à sua manutenção preditiva. Adicionalmente é possível visualizar a via através de imagens de satélites. Com isso pretende-se agregar valor e eficiência aos serviços fornecidos nas ferrovias gerenciadas pela Vale, assim como reforçar a sua segurança, evitando-se incidentes que podem ser desde os mais leves, como pequenas interrupções de serviços, quanto os mais severos, como desastres ambientais.

1.2 MOTIVAÇÃO

Nos últimos anos o processo de Big Data Analytics (BDA) vem se tornando amplamente utilizado tanto no ambiente acadêmico quanto no industrial. Isso não poderia ser diferente, afinal, o desenvolvimento tecnológico e a produção constante e massiva de dados da Indústria 4.0 abriram espaço para uma série de aplicações até então inexploradas. BDA consiste em aplicar técnicas avançadas de análise de dados para extrair informações que não seriam acessíveis sem um processamento eficiente de uma grande quantidade de dados.

Diversas áreas podem se beneficiar da adoção de aplicações de BDA, e sistemas ferroviários não são uma exceção. Um estudo de 2018 (2) contém um levantamento que buscou analisar com qual finalidade estão sendo empregadas técnicas de BDA nesses sistemas. Para isso, os autores fizeram uma extensa revisão literária de aplicações recentes de Big Data em linhas ferroviárias, levando em consideração três subáreas: operações, manutenção e segurança. Ao todo, foram analisados 115 papers, dos quais 22% tratam de aplicações em segurança, 29% em operações e 49% em manutenção. Além disso, os autores observaram um crescimento considerável do número de publicações nos últimos anos. Isso demonstra uma tendência na utilização de BDA para a resolução de problemas e aumento da efetividade dos serviços da indústria ferroviária.

A Vale opera, desde o ano de 1997 (3), duas ferrovias sob o regime de concessão da Agência Nacional de Transportes Terrestres (ANTT) – a Estrada de Ferro Carajás e a Estrada de Ferro Vitória a Minas. Com 2141 km de linha total e operando desde 1904, a Estrada de Ferro Vitória Minas (figura 1) une as cidades de Belo Horizonte (MG) e Vitória (ES). Por ela circulam anualmente 119 milhões de toneladas de minério de ferro, 22 milhões de toneladas de produtos diversos, além de um milhão de passageiros (4). Acompanhando a tendência, o projeto da Vale analisa o expressivo volume de dados obtidos através do Carro Controle (CC) e do Vagão Instrumentado (VI) que percorrem periodicamente a EFVM.

“De tempos em tempos o Vagão Instrumentado, como parte da composição de um

Figura 1: Mapa da Estrada de Ferro Vitória Minas



Fonte: (Wikipedia, 2020) (5)

trem de minério de ferro, percorre a EFVM, dotado de um amplo conjunto de sensores destinados ao levantamento de dados que posteriormente permitem a definição de parâmetros que refletem a dinâmica veicular em função das condições da via permanente. Ou seja, o estado de conservação da via permanente se reflete nos vagões que a percorrem. Adicionalmente, em intervalos de tempo maiores, tem-se a passagem de um Carro Controle também dotado de um conjunto de sensores, coletando séries complementares de dados referentes à geometria da via permanente” (1).

Voltando ao estudo mencionado (2), ele demonstrou que 11% das aplicações de manutenção eram focadas em equipamentos de sinalização, 36% em trilhos e 53% em veículos. Isso fornece uma base teórica sólida para a inferência das condições de infraestrutura da via permanente a partir do processamento dos dados coletados pelos VI e CC.

Para poder tirar proveito dessas inferências, esse trabalho focará na visualização dos resultados. Ela tem um papel fundamental em Big Data Analytics (BDA) para permitir que os usuários entendam o problema, gerem hipóteses, e tomem decisões (6). Nesse contexto, visa-se fornecer à Vale um sistema que facilite a adoção de manutenções preditivas da EFVM. Ademais, uma análise de diferentes tecnologias de visualização de dados e storytelling, assim como a firmiação desses e outros conceitos de BDA, vem como consequência da pesquisa realizada.

1.3 JUSTIFICATIVA

A pesquisa realizada para a elaboração do projeto da Vale (1) diz:

Tradicionalmente, ações de manutenção são do tipo corretiva ou preventiva. A primeira ocorre após uma falha, por vezes catastrófica, do equipamento e é sempre a que apresenta maior custo, pois além da troca do material inutilizado tem-se os custos de interrupção do funcionamento da ferrovia e de reparação de eventuais consequências desastrosas da falha. Já a manutenção preventiva, que leva em conta conceitos estatísticos de confiabilidade como o tempo médio entre falhas, procura antecipar ações de manutenção, de maneira a garantir a não interrupção do serviço, proporcionar maior segurança à operação e previsibilidade nas ações de manutenção. Embora bastante utilizada, a manutenção preventiva também tem suas falhas – frequentemente descartam-se componentes que ainda poderiam apresentar vida útil considerável (7).

Um conceito que traz maior eficiência é o da manutenção preditiva – a partir de um monitoramento contínuo das condições de operação de um equipamento, programam-se atividades de manutenção apenas quando se detectar alguma degradação potencial do equipamento. Outro ganho significativo desta abordagem é a possibilidade da criação de modelos de degradação, baseados em dados históricos, que permitirão prever possíveis problemas futuros com maior precisão.

Com isso em mente, esse projeto tem expectativas de:

- Inovar as tecnologias adotadas pelo setor de Engenharia Ferroviária da Vale;
- Aumentar o atual mercado da empresa, por meio do incremento em ações de manutenção preditiva relacionadas à via permanente, com a consequente redução em períodos de indisponibilidade de uso da ferrovia;
- Reduzir custos de investimento, por meio do aumento da vida útil das composições ferroviárias e da via permanente;
- Impactar positivamente o meio ambiente e a sociedade, através do aumento da eficiência energética dos serviços providos e da regularidade em oferta de transporte às comunidades que dependem de suas operações.

Complementarmente à última expectativa mencionada, em um estudo sobre sustentabilidade nos sistemas ferroviários (8), o autor levanta os impactos ambientais negativos provenientes de três fontes: circulação do material rodante, manutenção da via permanente, atividades desenvolvidas nas unidades de apoio da ferrovia. As consequências desses três pontos podem ser mitigadas com uma predição eficiente da necessidade de manutenção da via permanente, aumentando os benefícios ambientais alavancados por este trabalho.

CAPÍTULO 2

REVISÃO BIBLIOGRÁFICA

2.1 ASPECTOS CONCEITUAIS

2.1.1 Defeitos da Geometria da Via Permanente

De acordo com a American Railway Engineering and Maintenance-of-Way Association (AREMA), a geometria da superestrutura ferroviária pode ser descrita como “A posição que cada trilho, ou a linha central, ocupa pelo espaço tridimensional” (9).

É de suma importância que a qualidade da via férrea seja garantida, pois é um fator importante para a velocidade e segurança dos trens que a trafegam. Nesse contexto, um dos principais problemas que deve ser considerado são as vibrações verticais que a passagem dos vagões irá gerar. Dependendo de suas frequências, elas acarretarão na deterioração da via ou o desconforto dos passageiros que a utilizam.

Para preservar a geometria da via é necessário contínuas inspeções e manutenções de toda a linha férrea (10). Dentre os parâmetros citados pela ABNT (11), destacam-se os defeitos abordados a seguir.

2.1.1.1 Alargamento ou estreitamento de bitolas

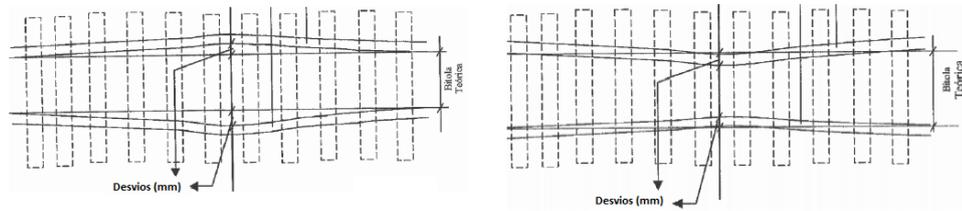
A bitola é um indicador de qualidade de construção e do estado da conservação da via. Quando este parâmetro não se apresenta em conformidade com a norma, traduz o reflexo do desgaste e degradação dos materiais ou a presença de anomalias.(10)

O defeito de alargamento ou estreitamento é quando, em um trecho da via, há uma diferença entre duas secções, como pode ser observado na imagem (figura 2). Esse fenômeno pode ocorrer por problemas em grupos de dormentes em condições ruins, desgastados ou por falta de fixação nas juntas (11)

2.1.1.2 Desnívelamento Longitudinal

O desnívelamento longitudinal se dá quando há pontos altos ou baixos de ocorrência simultânea (por igual em ambas as filas) ao longo de uma linha (13). Geralmente, os

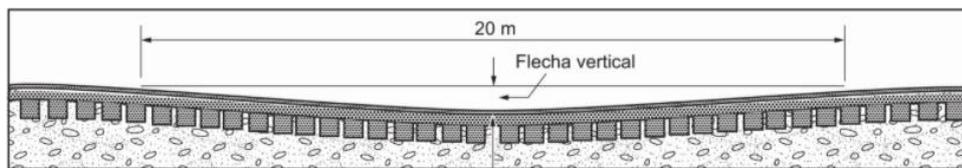
Figura 2: Alargamento e estreitamento de bitola



Fonte: (Silva, 2012) (12)

trilhos se deslocam para baixo formando uma bacia. Esta deformação pode ser medida através de uma flecha vertical a partir do topo do trilho em uma corda de 20m (11)(figura 3). Este fenômeno normalmente acontece como causa da drenagem do solo ou por juntas desniveladas. De modo geral, esta falha não é tao severa como as que veremos a seguir. Mas ainda assim dependendo de seu grau de severidade pode afetar o conforto dos passageiros e desgastar os equipamentos da via.

Figura 3: Desnivelamento longitudinal dos trilhos



Fonte: (ABNT, 2016) (11)

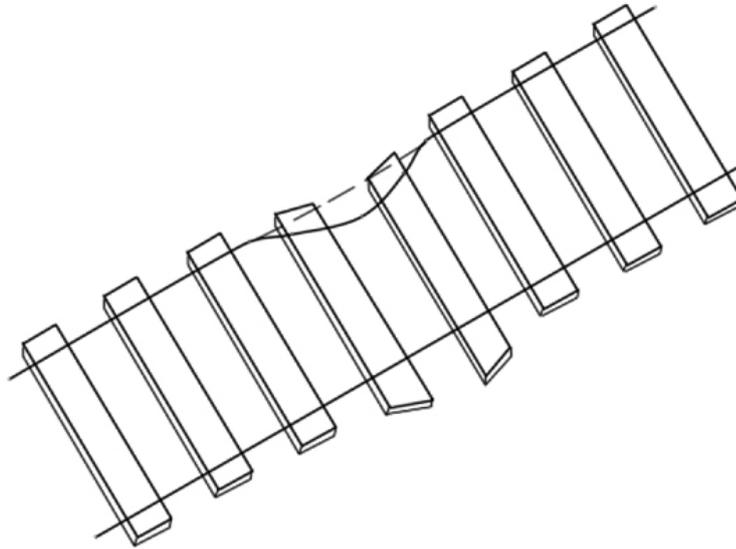
2.1.1.3 Desnivelamento Transversal

“O nivelamento transversal diz respeito às cotas dos boletos dos trilhos paralelos. É medido no sentido transversal, ou seja, perpendicularmente ao eixo da via” (13)(figura 6). Este defeito se dá pelo vazio que fica entre a superfície do solo e os trilhos, que pode acontecer por falta de compactação do aterro. Este tipo de falha é mais grave que a anterior pois ela afeta em grande intensidade a geometria da via. Assim, dependendo do grau de severidade da deformação, é preciso reduzir a velocidade dos trens para poder operar na via.

2.1.1.4 Empenamento

“O empeno pode ser definido como sendo uma variação do nivelamento transversal entre dois pontos de medição” (13), como mostrado na figura 6. Esta deformação da via férrea é muito séria pois ela pode acarretar na torção do trilho, que é a situação mais grave que pode acontecer. O empeno pode ser medido tomando-se como base 3 ou 5 metros

Figura 4: Desnívelamento transversal dos trilhos



Fonte: (Ribeiro, 2016) (13)

de distância entre as leituras de superelevação. Para tal, deve-se considerar um sinal de referência do empeno, podendo ser adotada como positiva a elevação do lado direito e, como negativa, a do lado esquerdo. O empeno é calculado pela diferença dos nivelamentos transversais em reta (13), como mostra a seguinte fórmula:

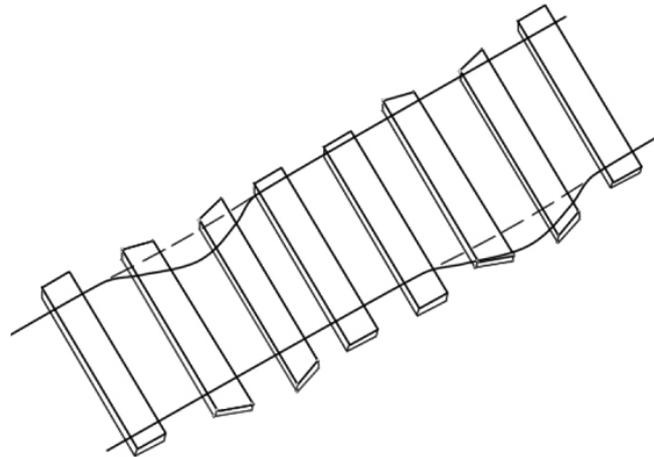
$$E = S_{frente} - S_{anterior}$$

Se o valor obtido for pequeno, deve-se diminuir a velocidade na via e iniciar uma manutenção preventiva dela. Já, se o grau for alto demais, a via deve ser interrompida e uma manutenção corretiva deve ser realizada. A Vale adota o valor de $\pm 9,1$ mm para o início da manutenção preditiva, e $\pm 12,0$ mm para a interrupção da via.

2.1.1.5 Alinhamento

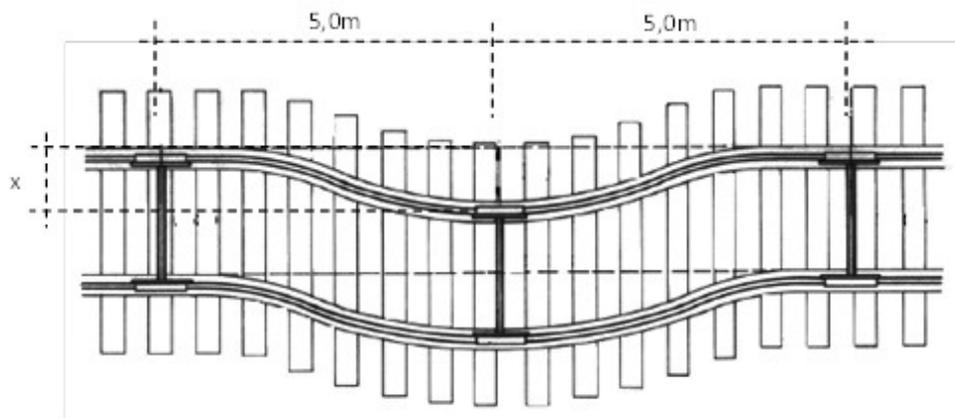
O problema de alinhamento se dá quando um ponto qualquer no trilho, situado entre outros dois pontos que formam uma reta, se desloca horizontalmente como pode ser observado na figura 6.

Figura 5: desenho esquemático do empeno



Fonte: (Ribeiro, 2016) (13)

Figura 6: desenho esquemático do alinhamento



Fonte: (Ribeiro, 2016) (13)

2.1.2 Visualização de Big Data Risk Analysis

Um dos maiores problemas do Big Data é o de se utilizar uma grande quantidade de dados de forma precisa e eficiente. Para resolvê-lo, alguns questionamentos devem ser feitos: O que é preciso saber? O que os dados mostram? O que o usuário entende?

Feito esse ponderamento, logo conclui-se que o processo de se criar a visualização dos dados é de suma importância para um projeto de BDRA, pois é ela quem irá fazer o elo entre os modelos preditivos de ML e o usuário.

A visualização de dados deve lidar não apenas com as técnicas e ferramentas disponíveis, mas também com os fatores humanos de compreensão dos dados apresentados

e os consequentes riscos que eles anunciam. Um entendimento dos fatores humanos é essencial para fornecer visualizações de riscos que levem em consideração os vieses inerentes e as heurísticas adotadas por analistas de segurança e tomadores de decisões.(14)

2.1.2.1 Análise Visual

A análise visual é uma área multidisciplinar que geralmente tenta obter conhecimento de dados massivos, incompletos, inconsistentes e conflitantes a fim de apoiar a avaliação e a tomada de decisão. Ela encaminha o processo de análise e apoia a comunicação das informações para o público-alvo, mas geralmente requer julgamento humano (14).

A visualização em projetos de BDRA relacionados a ferrovias enfrenta grandes desafios, pois, além de se ter uma grande variedade de informações a serem mostradas, ela também precisa permitir o usuário a fazer análises complexas e manejar os dados de diferentes formas e diversos níveis de qualidade.

Um outro fator importante da visualização que deve ser levado em conta é que ela será utilizada por usuários com diferentes níveis de conhecimento. Logo ela tem que ser flexível para que um usuário mais avançado possa fazer análises mais complexas, mas também tem que ser acessível para que uma pessoa com menor conhecimento técnico possa utilizar o sistema sem dificuldade.

2.1.2.2 Análise de dados

A análise de dados pode ser descrita como o processo investigativo usado para extrair conhecimentos, informações e percepções sobre a realidade por meio do exame de dados (14). Nela se emprega técnicas estatísticas automáticas e técnicas de aprendizado de máquina para a extração de informações valiosas e descobrimento de relações complexas entre os dados.

Entretanto, a análise de uma grande quantidade de dados utilizando técnicas e softwares complexos pode causar oclusão de dados, desorientação ou má interpretação. Isso afeta o controle dos dados e o entendimento que o usuário tem sobre eles.

A análise de dados é uma tarefa de sensemaking na qual se tem o objetivo de criar “ideias confiáveis sobre a realidade a partir dos dados observados” (14). O sensemaking é uma análise subjetiva dos resultados, o que faz com que pessoas com diferentes experiências tenham diferentes conclusões de um mesmo conjunto de dados. Porém, a visualização de dados pode ajudar os usuários a aumentar suas habilidades perceptivas,

estender a memória de trabalho e auxiliar na recuperação da memória de longo prazo (14), auxiliando na obtenção de resultados mais objetivos.

Em um projeto de BDRA aplicado a ferrovias, a análise de dados deve ser uma combinação de visualização, análise automática dos dados e interação humano-computador. Isso engloba uma análise exploratória e a escolha de testes estatísticos sobre a supervisão humana (14). O usuário também deve conseguir fazer quaisquer tipos de testes estatísticos independentemente dos tipos de dados, possibilitando-lo a ver o que o sistema faz, o que melhora o seu entendimento dos dados que foram apresentados.

2.1.2.3 Storytelling

O storytelling é um conceito abrangente que pode ser definido como a arte interativa de se usar palavras e ações para revelar os elementos e imagens de uma história enquanto se estimula a imaginação do ouvinte (15). Quando aplicado a dados, ele se torna a prática de combinar dados concretos com comunicação humana para criar uma narrativa envolvente ancorada em fatos. Ele usa técnicas de visualização de dados para ajudar a transmitir o significado deles de uma forma atraente e relevante para o público (16).

Histórias baseadas em dados são criadas por meio do processo de análise e filtragem de grandes conjuntos de dados para se descobrir insights e revelar maneiras novas ou diferentes de se entender as informações. Elas são personalizadas para um público específico e o contexto em que são consumidas. Isso pode ajudar a comunicar informações de maneira mais eficaz, ao mesmo tempo que gera a menor carga cognitiva, o que afeta a energia mental que o usuário precisa gastar para entender sua informação (16).

2.2 TECNOLOGIAS UTILIZADAS

2.2.1 Django



Figura 7: Django
via 17

O Django (17) é um web framework Python que utiliza o padrão de design Model-View-Controller (MVC). Ele separa os dados (model), a lógica da aplicação (controller) e a interface do usuário (view), de forma a obter uma estrutura de aplicações web sólida e intuitivamente manejável. Ele é amplamente utilizado no mercado devido à sua facilidade de desenvolver aplicações rápidas, seguras e escaláveis.

O backend do projeto da Vale foi implementado tendo o Django como base. No contexto da aplicação de visualização de dados, ele foi utilizado para mapear o endpoint do serviço e para disponibilizar a API que provém as informações a serem exibidas.

2.2.2 Sphinx



Figura 8: Sphinx
via 18

O Sphinx (18) é uma ferramenta para documentação de projetos em Python. Ele é versátil e fácil de utilizar, podendo gerar documentos em formatos PDF e HTML. Pode-se aplicar diferentes temas ao HTML gerado conforme o gosto/necessidade do desenvolvedor.

Uma outra vantagem é a criação automática de um campo de pesquisa e de índices, facilitando a navegação pela documentação gerada.

Uma alternativa ao Sphinx para o escopo deste trabalho é o Swagger (19), outra ferramenta bastante utilizada que consegue documentar *endpoints* de forma clara e objetiva. Entretanto o Swagger é específico para APIs, sendo necessária uma segunda framework para a documentação de outros tipos de *endpoints*. Portanto, optou-se pelo Sphinx por sua maior flexibilidade.

2.2.3 Mapbox



Figura 9: Mapbox
via 20

O Mapbox (20) é uma ferramenta que possibilita a criação de mapas interativos. Ele é responsável pela renderização eficiente dos mapas, além de permitir uma série de customizações e a adição de camadas que podem ser atualizadas dinamicamente. Essas camadas servem diversos propósitos, como a localização de pontos específicos e a exibição de rotas.

O Mapbox é amplamente utilizado e sua documentação facilita bastante o desenvolvimento de projetos web baseados em mapas interativos. Ele foi o core para implementação das funcionalidades listadas na seção 3.3.

2.2.4 React

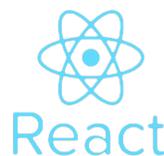


Figura 10: React
via 21

O React (21) é uma biblioteca JavaScript que auxilia na criação de interfaces de usuário. Ele é responsável pelo gerenciamento de estado e renderização de todos os elementos DOM em uma Single-Page Application (SPA) (embora também possa ser utilizado com outras arquiteturas).

Alternativas ao React são o Angular (22) e o VueJs (23). Todas são ferramentas poderosas para o desenvolvimento de interfaces web dinâmicas, mas optou-se pelo React pela sua presença no mercado de trabalho (24), que é ainda mais reforçada devido ao seu irmão React Native, tecnologia dominante no desenvolvimento mobile (25).

2.2.5 Bootstrap



Figura 11: Bootstrap
via 26

O Bootstrap (26) é um kit de ferramentas que auxilia na criação de páginas HTML responsivas. Através de classes CSS pré-estabelecidas, é possível criar sites inteiros com um ótimo resultado estético e esforço de estilização mínimo.

Uma alternativa ao Bootstrap é o Material (27), que é mantido pela Google e serve de base para uma série de frameworks. Ambos são bem avaliados e amplamente utilizados, possuindo vantagens e desvantagens bastante subjetivas. Para este projeto, o Bootstrap foi escolhido devido à familiaridade dos integrantes com a ferramenta.

2.2.5.1 Bootstrap Icons

Um subprojeto da equipe do Bootstrap é o Bootstrap Icons: uma série de SVG's que não só alavancam ainda mais a estética de um site, como também contribuem para a intuitividade da aplicação. Tal intuitividade se deve à associação feita pelo usuário de certos ícones a funcionalidades específicas, baseada em suas experiências anteriores (28). Os ícones dessa biblioteca foram utilizados em diversos botões do projeto.

2.2.6 ChartJs

O ChartJs (29) é uma biblioteca JavaScript especializada em visualização de dados. Ele disponibiliza 8 tipos diferentes de gráficos personalizáveis que atendem boa parte da demanda de aplicações de análise de dados. Neste projeto, ele foi utilizado para a exibição de estatísticas a respeito da estrutura de trechos da via.



Figura 12: ChartJs
via [29](#)

Uma alternativa ao ChartJs é o D3Js ([30](#)), outra biblioteca amplamente utilizada para análise de dados. Ele disponibiliza recursos mais genéricos para a criação e manipulação de SVG's. Optou-se pela utilização do ChartJs por ele atender as necessidades atuais do projeto, mas o D3 pode ser uma opção para trabalhos futuros que requeiram gráficos e/ou visualizações mais robustas e customizáveis.

2.2.7 Turf

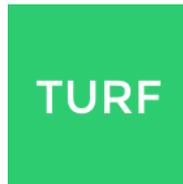


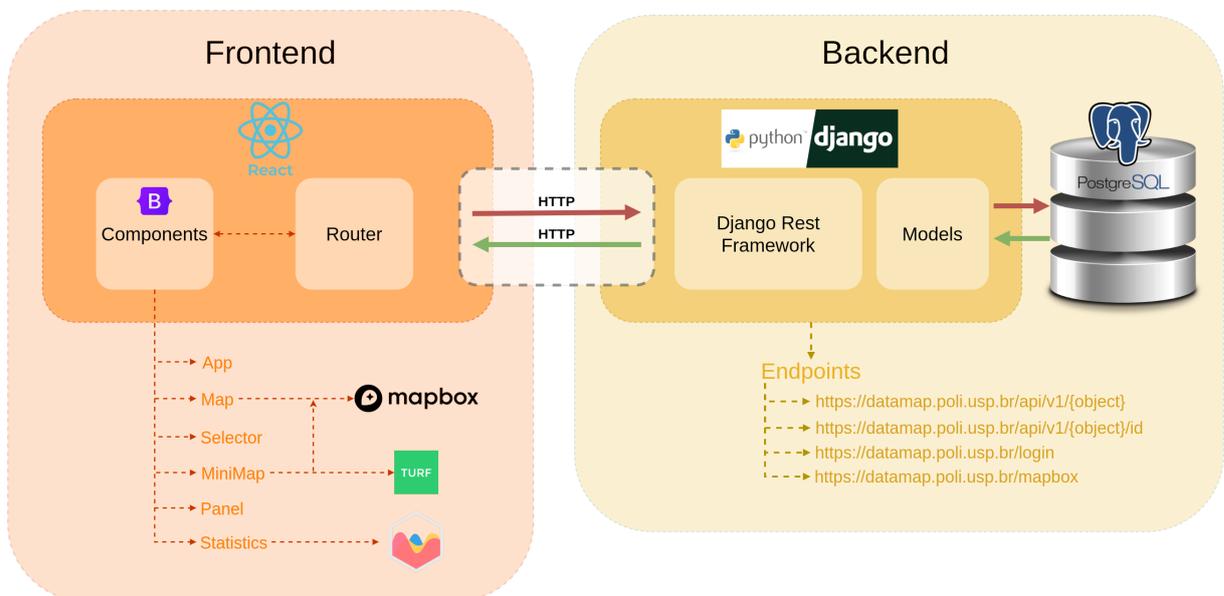
Figura 13: Turf
via [31](#)

O Turf ([31](#)) é uma ferramenta de análises geoespaciais avançadas para navegadores. Ele serviu para a implementação de algumas funcionalidades que serão melhor detalhadas na seção [4](#), e sua utilização é recomendada pela própria equipe do Mapbox. Sendo assim, nenhuma ferramenta alternativa foi revisada.

2.3 ARQUITETURA FINAL

O diagrama a seguir ilustra a arquitetura final da aplicação implementada. Detalhes de implementação serão explicados na seção 4.

Figura 14: Diagrama da arquitetura final



Fonte: (Autores, 2021)

É válido ressaltar que API's para o deploy dos modelos de Machine Learning desenvolvidos podem ser facilmente adicionadas conforme o desenvolvimento do projeto e a necessidade da aplicação.

CAPÍTULO 3

ESPECIFICAÇÃO/PROJETO

3.1 FASE DE ESTUDOS

3.1.1 Conceituação do problema

Para buscar a melhor maneira de solucionar o problema de visualização no contexto proposto, foi adotada uma metodologia de pesquisa exploratória. Motivado em desenvolver um projeto em ciências dos dados, o grupo começou o trabalho com um levantamento bibliográfico de artigos e relatórios técnicos que pudessem ajudar na questão. Tal levantamento incluiu estudos à respeito de Data Visualization, Storytelling e Big Data aplicado a sistemas ferroviários. Também foram feitas pesquisas a respeito da própria área de engenharia ferroviária para uma melhor familiarização com termos técnicos e melhor entendimento das causas, consequências e requisitos do projeto que está sendo desenvolvido pela Vale.

3.1.2 Contexto externo

Figura 15: Workflow de um projeto de Big Data Analytics



Fonte: (Kashyap, 2019) (32)

Este trabalho está focado na extremidade final de um projeto de BDA, a visualização. Isso faz necessário um entendimento das componente anteriores. Sendo assim, uma das etapas do projeto consistiu no estudo da análise de dados implementada, através da leitura de relatórios e examinação do código fonte.

Nessa parte, um dos principais elementos considerados foi o conjunto de dados resultantes do processamento realizado, pois ele serviu como entrada do sistema desenvolvido. Além disso, conforme levantado no RNF4 da tabela 2, a aplicação foi implantada num ambiente previamente configurado. Isso torna imperativa a compreensão da infraestrutura desse ambiente, atentando-se para possíveis restrições técnicas.

3.2 PROJETO

Como dito anteriormente, este trabalho visa possibilitar a tomada de decisões quanto a manutenção de componentes da via permanente da EFVM, o que o caracteriza como um projeto de Big Data Risk Analysis (BDRA). Dessa forma, um estudo (14) levantou 6 requisitos específicos para a parte de visualização de aplicações de BDRA em sistemas ferroviários que foram levados em consideração para projetar-se uma solução:

1. Visualização de BDRA para sistemas ferroviários deve ser tratado como um problema analítico visual. A visualização deve ajudar a análise de riscos no gerenciamento de dados, na análise de dados, na interação humano-computador e no processo de comunicação de riscos.
2. Deve-se introduzir análise visual nas técnicas de gerenciamento de dados para ponderar incertezas da análise de dados.
3. A análise de dados deve ser uma combinação de análise de dados automatizada, interação humano-computador e visualização. A visualização deve considerar os processos cognitivos na análise de dados.
4. A visualização de parâmetros de segurança de vias ferroviárias deve contemplar a interação com os dados e a análise colaborativa e simultânea de diversos usuários.
5. O processo de comunicação de risco deve ser adaptado a cada ator envolvido na segurança. Técnicas de visualização adequadas para representar riscos e incertezas devem ser usadas para reduzir os vieses e melhorar a compreensão das informações.
6. Deve-se implementar um sistema de visualização automatizado para aplicar análises visuais envolvidas em BDRA de sistemas ferroviários (gerenciamento de dados, interação humano-computador, análise de dados e comunicação de riscos).

3.3 REQUISITOS FUNCIONAIS

A tabela 1 lista as funcionalidades que a aplicação deve disponibilizar.

Requisito	Nome	Descrição
RF0	Mapa da EFVM	Sistema disponibiliza um mapa interativo que permite o usuário identificar a ferrovia a partir de pontos de referências e coordenadas
RF1	Exibição de layers e overlays	Sistema exibe camadas devidamente localizadas no mapa que mostram/informam: <ul style="list-style-type: none"> • RH's • Níveis Severidade • Falhas • Restrições • Observações • Análises e Estatísticas • Clustering (parametrizável) • Score e Preditividade
RF2	Vista panorâmica	Sistema exibe uma visualização tridimensional interativa da EFVM na qual é possível identificar a via e pontos de referência
RF3	Seleção de trecho	Sistema permite a seleção de um trecho específico da via interativamente através do mouse

RF4	Filtros de seleção	<p>Sistema possibilita a filtragem de layers/informações exibidas na tela através dos atributos:</p> <ul style="list-style-type: none"> • Data & timeline rolling • Viagem • Trecho • Elemento • Parâmetro específico (Severidade X, Aceleração Y, etc.)
RF5	Exportação	<p>Sistema permite exportar informações exibidas em arquivos do tipo imagem (PNG e JPG), KML e CSV</p>
RF6	Comentários e observações	<p>Sistema permite a inserção de comentários do usuário que devem ser salvos e estar disponíveis nas próximas seções que o usuário logar</p>
RF7	Alertas e mensagens	<p>Sistema exibe alertas e envia mensagens (email) a usuários quando algum evento importante (como detecção de altas severidades) ocorrer</p>

Tabela 1: Requisitos funcionais

3.4 REQUISITOS NÃO FUNCIONAIS

Seguindo as diretrizes estabelecidas pela ISO/IEC 25010 (33), a tabela 2 lista os requisitos não funcionais do sistema.

Requisito	Nome	Descrição
RNF1	Acessibilidade	Os dados devem ser exibidos de forma que usuários com diferentes níveis técnicos tenham acesso à informação necessária
RNF2	Usabilidade	As informações fornecidas pela aplicação devem ser exibidas da forma mais intuitiva e direta possível
RNF3	Confiabilidade	As informações exibidas devem corresponder aos dados mais recentes do servidor
RNF4	Co-existência	A aplicação deve ser compatível com a infraestrutura já desenvolvida pelo projeto da Vale sem prejudicar o funcionamento de outros serviços

Tabela 2: Requisitos não funcionais

CAPÍTULO 4

IMPLEMENTAÇÃO

4.1 ESTRUTURA DE DIRETÓRIOS

Como especificado no RNF4 (2), nossa aplicação compartilha o ambiente e a infraestrutura de outros serviços que compõem o projeto Cátedra Under Rail da Vale. Os códigos e arquivos desses serviços estão hospedados em um repositório privado do GitLab (34), em um subprojeto chamado DataMap. Sendo assim, para a implementação do nosso trabalho, foi criada uma branch própria. Todas as funcionalidades foram commitadas para essa branch ao longo do desenvolvimento do projeto.

A árvore de diretórios a seguir lista os arquivos da aplicação:

```

datamap/
├── src/api/
│   ├── views.py
│   ├── urls.py
│   └── static/
│       ├── css/
│       │   └── mapbox.css
│       └── js/
│           ├── map.jsx
│           ├── map.js
│           ├── map-selector.jsx
│           ├── map-selector.js
│           ├── map-statistics.jsx
│           ├── map-statistics.js
│           ├── mapbox.jsx
│           ├── mapbox.js
│           ├── mini-map.jsx
│           ├── mini-map.js
│           ├── panel.jsx
│           └── panel.js
│       └── MapaEFVM.geojson
├── templates/
│   └── mapbox.html
└── doc/
    ├── build/html/
    │   ├── endpoints/
    │   │   └── railway_status_visualization.html
    │   └── index.html
  
```

```
├── endpoints/
│   └── railway_status_visualization.md
├── index.rst
└── conf.py
```

Essa é uma versão simplificada da estrutura de diretórios do projeto que visa enfatizar os arquivos que foram criados ou modificados especificamente para o nosso projeto.

Um detalhe mostrado na árvore acima são os arquivos “duplicados” na pasta *js/*. Essa redundância se deve à utilização da sintaxe JSX, uma extensão do JavaScript que simplifica a implementação de React Components e torna o código mais organizado e, conseqüentemente, mais legível. Mas, como o navegador só interpreta JavaScript puro, os arquivos *.jsx* foram traduzidos para *.js*.

O arquivo MapaEFVM.geojson foi criado a partir de um arquivo KML contendo informações que possibilitam a geolocalização dos trilhos da EFVM e locais importantes para a operação da ferrovia, além de permitir a associação de propriedades arbitrárias a essas localidades. A conversão para geojson foi necessária pois esse é o padrão adotado pelo Mapbox, ao passo que o formato KML é utilizado pelo Google Earth, ferramenta com a qual a Vale está acostumada a trabalhar.

Os demais arquivos fazem parte da implementação propriamente dita da aplicação, e suas finalidades serão melhor detalhadas nas seções a seguir. Como o foco desse projeto está na visualização de dados, boa parte do código foi dedicada exclusivamente ao frontend da aplicação, enquanto que as modificações no backend visaram principalmente a adição do serviço no ambiente do projeto DataMap e a disponibilização das informações necessárias para a visualização.

4.2 BACKEND

Para disponibilizar o acesso à nossa aplicação no ambiente do DataMap, foram modificados os arquivos *views.py* e *urls.py*. Neles, o Django foi configurado para associar o endpoint `/mapbox` ao arquivo *templates/mapbox.html*. Também foi estabelecido que o usuário deve criar uma sessão para poder acessar o serviço, o que é necessário para poder associar os comentários e observações do RF6 (1) ao usuário que os fez.

Além disso, foram implementados *endpoints* para devolver informações requisitadas pelo frontend. A seção a seguir descreve o processo de documentação desses *endpoints*.

4.2.1 Documentação

A pasta `/doc` contém os arquivos que são utilizados para a geração da documentação do projeto. Dentre eles temos o *index.rst* e o *railway-status-visualization.md*. O primeiro é a página inicial, na qual constam os links para as demais páginas da documentação. O segundo contém o detalhamento do *endpoint* que estamos utilizando.

Pode-se perceber uma diferença entre os formatos dos arquivos. A extensão *.rst* é para arquivos do tipo Re-Structured Text e a extensão *.md* é para arquivos do tipo Markdown. Ambas são linguagens de marcação, e desempenham papéis parecidos.

Esses arquivos contém metadados, compostos por *tags* e textos que, após compilados, geram documentos apropriados para leitura. A extensão *.rst* é obrigatória para a compilação da página principal no Sphinx, porém o formato Re-Structured Text é menos conhecido e menos popular no mercado. Sendo assim, foi instalado um *parser* que faz com que o Sphinx aceite arquivos do tipo Markdown. Essa linguagem suporta algumas sintaxes adicionais, facilitando o processo de documentação do projeto.

4.3 FRONTEND

4.3.1 Página Principal

Feita a requisição ao endpoint `/mapbox`, o servidor devolve a página principal da aplicação (*mapbox.html*). Ela puxa todas as dependências utilizadas para a implementação do projeto: React, Bootstrap, Mapbox, ChartJs e Turf. A página contém apenas um nó DOM `<div>` dentro do `<body>`. Isso se deve à estruturação da aplicação baseada no conceito Single-Page Application (SPA), no qual os elementos da página são dinamicamente renderizados conforme o usuário interage com eles.

4.3.2 CSS

Para a estilização de boa parte dos elementos HTML foram utilizadas as classes padrão do Bootstrap. Entretanto, estilizações específicas (como por exemplo alturas, larguras e cores) de alguns elementos foram implementadas no arquivo *mapbox.css*.

4.3.3 Componentes

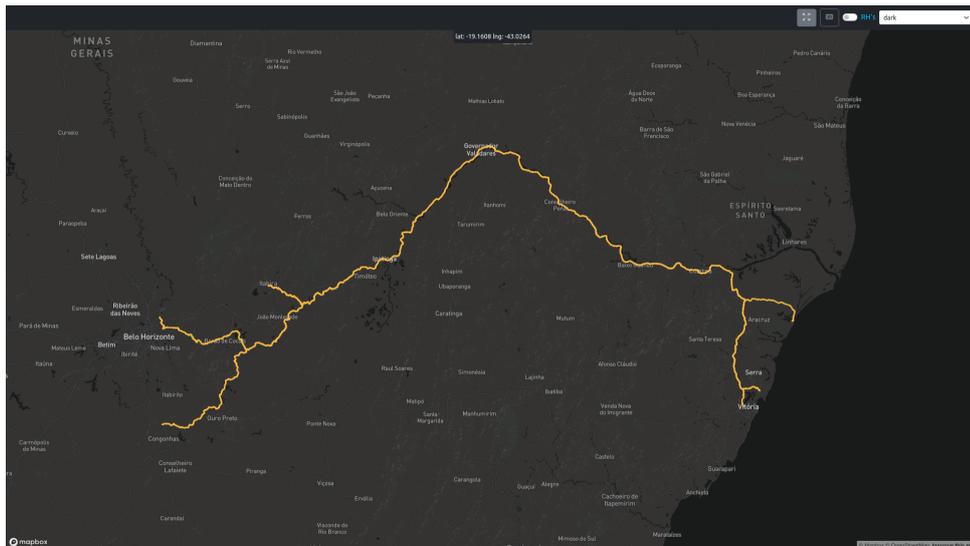
Conforme explicado na seção 2.2, a framework escolhida como base para a implementação do frontend foi o React. Sua arquitetura é baseada em ReactComponents, que têm como objetivo agrupar funcionalidades logicamente conectadas e passar informações entre si. Essas informações são guardadas em variáveis chamadas ReactStates e só podem ser passadas de cima para baixo na hierarquia de componentes (one-way data flow).

O código dos componentes implementados está completamente modularizado, o que significa que a página *mapbox.html* declara apenas o script principal (4.3.3.1), e os demais componentes importam uns aos outros conforme necessário. A seguir estão detalhados os componentes implementados para atender os requisitos descritos na seção 3.3.

4.3.3.1 App

Este componente está definido no arquivo *mapbox.jsx* e é responsável por instanciar e linkar os componentes Map (4.3.3.2) e Panel (4.3.3.4) da aplicação. A figura 16 mostra a tela inicial do programa, cuja estrutura DOM está descrita dentro do App. Na imagem é possível distinguir duas seções: o menu superior e o mapa interativo.

Figura 16: Tela inicial da aplicação

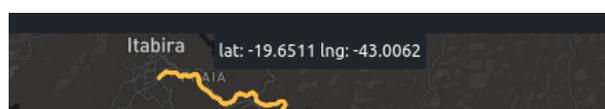


Fonte: (Autores, 2021)

O menu contém três botões e uma caixa de seleção. O primeiro botão simplesmente reseta a visualização do mapa para a visualização mostrada na figura 16. Isso é útil quando o usuário faz algumas manipulações no mapa (arrastar, dar zoom, girar, e etc) para obter informações de um região específica e quer ir atrás de outra. O segundo botão ativa a vista panorâmica tridimensional (figura 20), correspondente a uma das funcionalidades do RF2 1. O terceiro botão é um switch que possibilita localizar Reference Houses (RH's) (RF1 1), que são pontos estratégicos para a manutenção da ferrovia. A figura 18 ilustra o mapeamento desses pontos.

Finalmente, a caixa de seleção permite a escolha do estilo de mapa que melhor se adéque às preferências do usuário. Existem 8 estilos disponíveis e alguns exemplos são mostrados na figura 19.

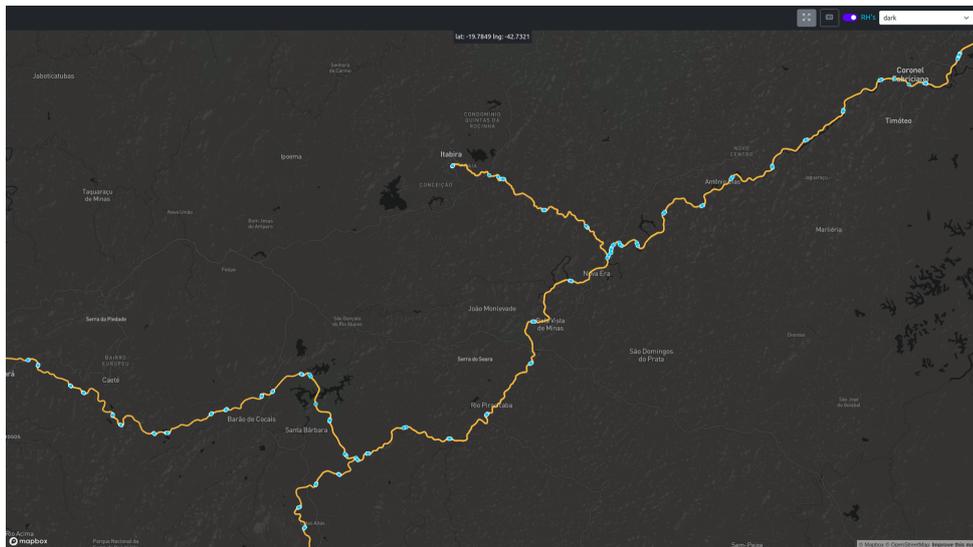
Figura 17: Painel de coordenadas do mouse



Fonte: (Autores, 2021)

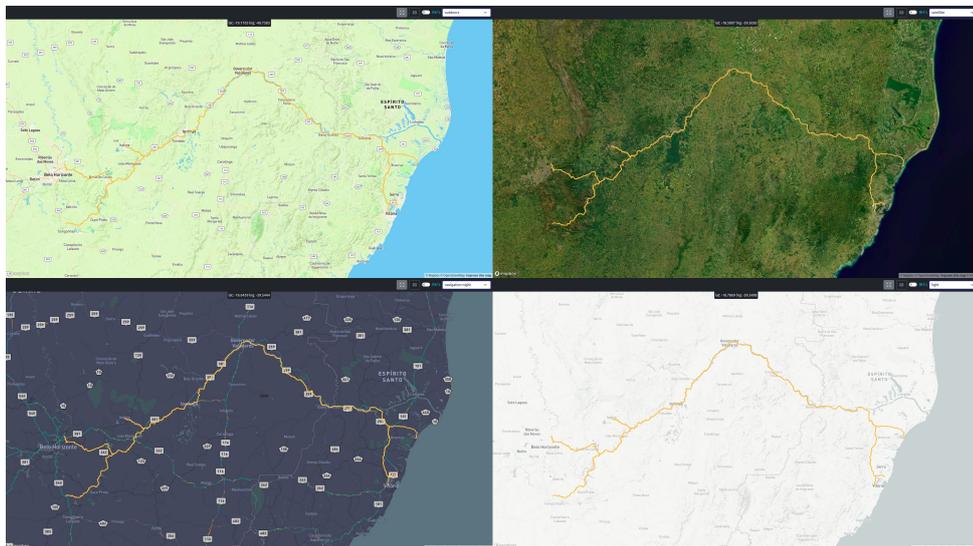
Uma última funcionalidade gerenciada por dois ReactStates do componente App é o painel da figura 17, localizado logo abaixo do menu no centro da página. Ele informa as coordenadas (latitude e longitude) do ponteiro do mouse conforme o usuário o move. Essa funcionalidade está relacionada ao RF0 1.

Figura 18: Switch de RH's ligado



Fonte: (Autores, 2021)

Figura 19: Quatro estilos alternativos de mapa



Fonte: (Autores, 2021)

4.3.3.2 Map

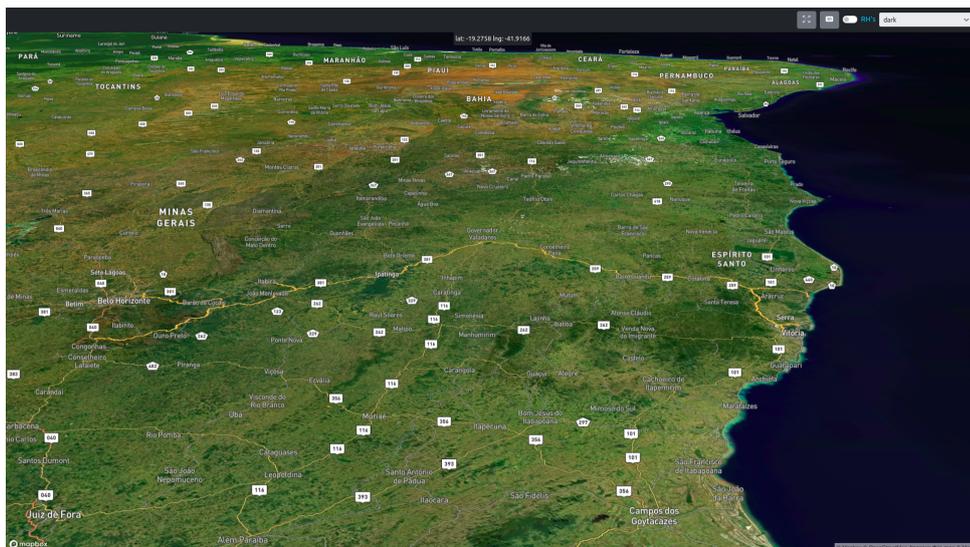
Este componente implementa o requisito base das funcionalidades da aplicação, o RF0 1, e pode ser visto como o core da aplicação. Ele importa os recursos disponibilizados pelo

Mapbox e define as camadas visualizadas no mapa, como a rota amarela correspondente à EFVM da figura 16 e os RH's da figura 18.

Embora os “sinais” enviados pelo usuário através do menu sejam capturados no componente App, é no Map que todas as alterações do mapa são implementadas. Por exemplo: os estilos mostrados na figura 19 são selecionados no menu, o que faz o App chamar a função `setStyle` do Map, que atualiza o mapa e chama uma callback para re-renderizar as camadas que estavam sendo exibidas.

Além disso, informações que serão contadas ao usuário são armazenadas neste componente. Sendo assim, ele é responsável por separá-las e realizar alguns cálculos em cima delas, o que resulta nos dados prontos para visualização. Esses dados são então passados para o MapStatistics (4.3.3.5), que os apresenta para análise.

Figura 20: Vista panorâmica inicial



Fonte: (Autores, 2021)

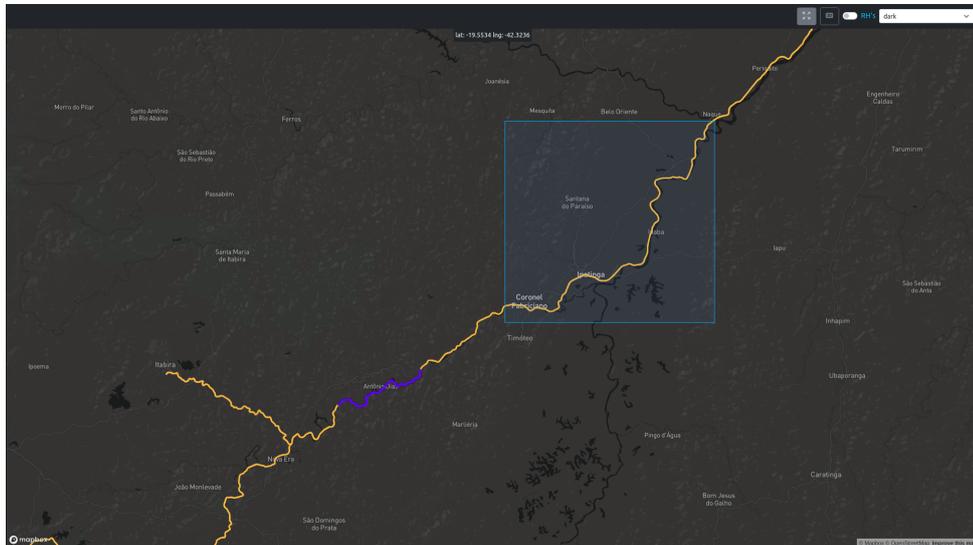
Uma outra funcionalidade importante implementada aqui corresponde ao RF3 1. Ela desabilita o drag do mapa quando o usuário aperta a tecla shift para que possa ser possível a seleção de um trecho da via. Essa seleção analisa quais elementos renderizados estão dentro da área delimitada pelo usuário. Esses elementos correspondem a um trecho da ferrovia e, feita a seleção, um painel (4.3.3.4) contendo estatísticas e informações sobre esse trecho é aberto na página.

4.3.3.3 MapSelector

O MapSelector implementa o retângulo que é visualizado durante a seleção de um trecho da via. Ele apenas instancia uma `<div>` com as posições recebidas do Map e com

as classes CSS adequadas. A figura 21 ilustra essa seleção.

Figura 21: Seleção de um trecho da ferrovia



Fonte: (Autores, 2021)

4.3.3.4 Panel

Este componente implementa um painel que é aberto após uma seleção de trecho e que pode ser fechado em qualquer momento. Ele é um contêiner no qual o único comportamento implementado é o posicionamento do painel (figura 22). Ademais, o Panel faz a ponte entre a interação do usuário com o mapa e os conteúdos exibidos pelos componentes MapStatistics (4.3.3.5) e MiniMap (4.3.3.6).

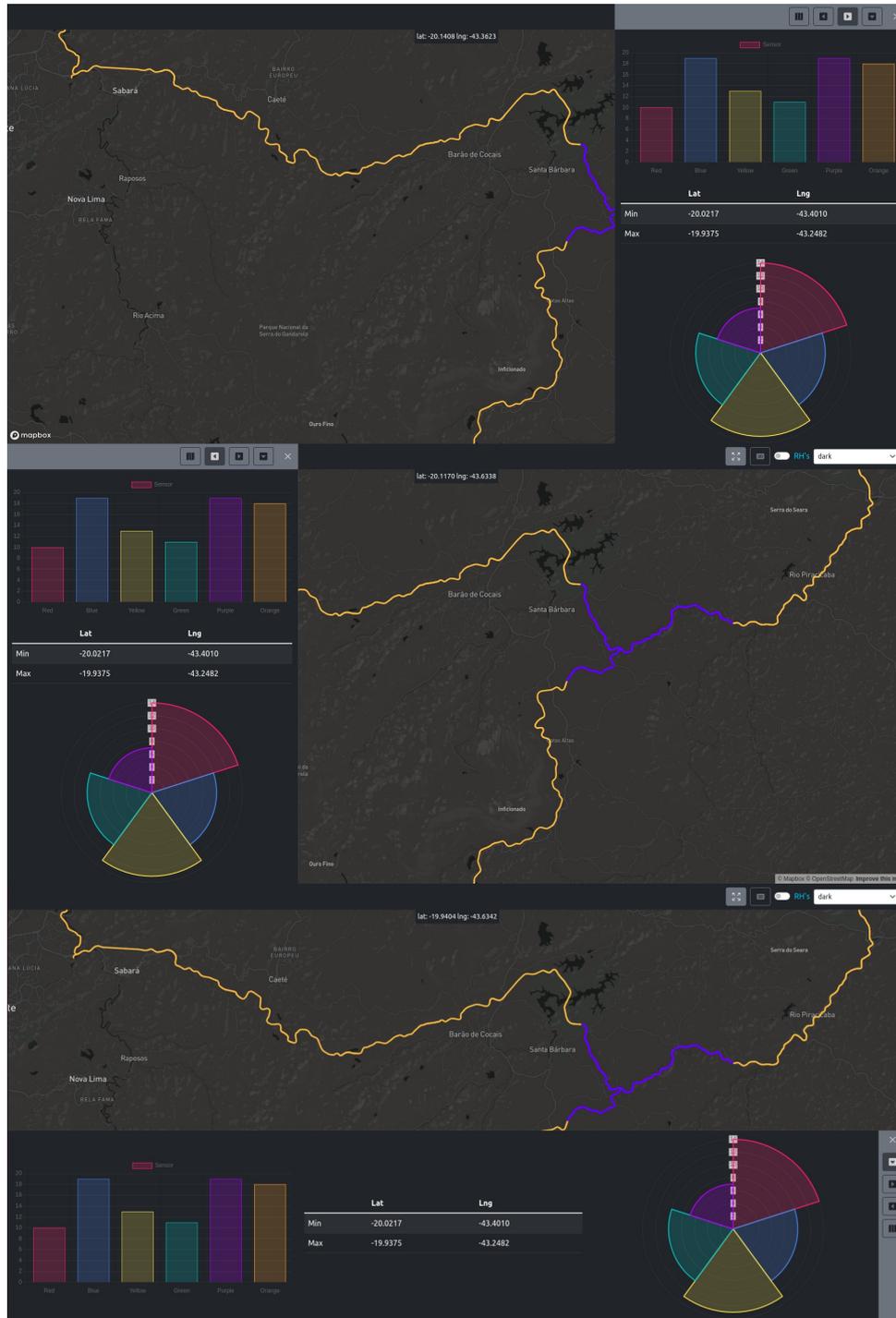
4.3.3.5 MapStatistics

O MapStatistics utiliza os recursos disponibilizados pelo ChartJs para prover ferramentas de análises estatísticas das informações inferidas pelos modelos de aprendizagem implementados no backend. Este componente é formado por:

- Um gráfico de barras;
- Uma tabela que mostra as coordenadas máximas e mínimas do trecho selecionado;
- Um gráfico de áreas polares.

A figura 22 mostra esses elementos dentro dos painéis.

Figura 22: Painel de estatísticas



Fonte: (Autores, 2021)

Atualmente os gráficos não exibem dados reais. Entretanto, todo o código de atualização dinâmica já foi implementado. Assim, conforme um novo trecho é selecionado, novos dados são gerados e exibidos para análise.

4.3.3.6 MiniMap

Este componente é ativado clicando no ícone de mapa presente no menu superior (ou lateral) do painel, e estende a funcionalidade de visualização 3D da ferrovia (RF2 1).

O MiniMap recebe o trecho selecionado pelo usuário e, utilizando as funções de análise geoespacial do Turf, calcula as informações necessárias para definir um caminho de coordenadas que percorre esse trecho. Feito isso, ele cria uma nova instância de um objeto do Mapbox. Esse objeto não é interativo, e é responsável por mostrar uma espécie de “vídeo de satélite” da via, que nada mais é do que o posicionamento da câmera nas coordenadas do caminho descrito.

Esse processo é feito através de uma função chamada periodicamente. Em cada chamada, uma espécie de timer é atualizado e seu valor é utilizado para definir a posição do caminho para a qual a câmera deve olhar. A figura 23 ilustra 3 momentos do vídeo de um trecho selecionado.

Figura 23: Mini mapa para visualização automática de trecho selecionado



Fonte: (Autores, 2021)

CAPÍTULO 5

DISCUSSÃO DOS RESULTADOS

5.1 RESULTADOS

Do planejamento do projeto, descrito na seção 3.1, à escolha das tecnologias (2.2), até a implementação (4), teve-se como prioridade a entrega de um sistema de visualização completo, útil e intuitivo. Nesse contexto, esta seção faz uma análise dos resultados obtidos a partir dos requisitos levantados nas seções 3.3 e 3.4.

A escolha do Mapbox implicou automaticamente na implementação do requisito funcional base do sistema, o RF0. Com o mapa interativo fornecido pelo framework, as informações contidas no arquivo *geoJson*, juntamente com o painel de coordenadas da figura 17, possibilitaram uma localização precisa dos trechos da EFVM. Sendo assim, todas as funcionalidades do RF0 foram atendidas.

Quanto ao RF1, a principal camada de exibição (os RH's) pode ser visualizada por meio de um dos botões do menu superior da aplicação. Adicionalmente, o painel de estatísticas da figura 22 provém as informações necessárias para as análises mencionadas no requisito. Ademais, com a possibilidade de seleção de trecho da maneira mais intuitiva possível, foi viável relacionar as funcionalidades dos requisitos RF1, RF3 e RF4 de forma a se obter o comportamento esperado por um usuário que quer analisar trechos específicos da via.

A visualização interativa tridimensional provém os meios necessários para se identificar determinada parte da via através de pontos de referência. Juntando isso à visualização em vídeo de um trecho selecionado provida pelo minimapa, tem-se o requisito RF2 completamente desenvolvido.

Todos os requisitos não funcionais foram atendidos: as informações providas pela aplicação são extraídas diretamente do backend, elas são úteis para uma série de análises, o sistema está intuitivo e ele foi implementado na infraestrutura do DataMap sem comprometer qualquer outro serviço. Além disso, a aplicação pode ser utilizada por diversos atores com diversas finalidades, e nenhuma funcionalidade implica no detrimento de outra.

Por fim, com o sistema de documentação implementado, demais serviços podem apro-

veitar das especificações de *endpoints* disponíveis. Sendo assim, os resultados obtidos foram satisfatórios e espera-se que a aplicação sirva de base para uma ferramenta de uso industrial extremamente versátil em um projeto de Big Data Risk Analysis para sistemas ferroviários. É válido comentar que ainda estão sendo implementadas algumas funcionalidades dos demais requisitos detalhados na seção 3.3, o que resultará na entrega de um sistema ainda mais completo.

CAPÍTULO 6

CONSIDERAÇÕES FINAIS

6.1 TRABALHOS FUTUROS

A fim de expandir as análises que um usuário pode fazer, são propostas a seguir seis funcionalidades para implementação futura.

Uma primeira consiste em informar quais defeitos (dos listados na seção 2.1.1) podem estar associados aos resultados das predições. Isso iria requerer do backend uma análise específica dos dados coletados pelos sensores do Vagão Instrumentado (VI) e do Carro Controle (CC), ou mesmo a implementação de um novo modelo de ML com esse objetivo. Uma vez obtida a informação, não seria difícil adicionar uma nova camada de exibição no mapa.

A segunda é a simulação dos impactos que as deteriorações causadas pelos defeitos calculados/preditos têm na movimentação de um vagão. Isso daria uma compreensão visual imediata do que pode acontecer com esse vagão quando ele passa por um trecho no qual o defeito está presente.

A terceira é a incorporação de dados meteorológicos/climáticos nas análises e modelos desenvolvidos. Isso seria interessante pela influência que fatores climáticos têm em outros dados coletados, como a velocidade do VI. É válido ressaltar que representantes da Vale comentaram que esses dados estariam facilmente disponíveis.

A quarta consiste em uma forma de se visualizar o estado real de trechos da via. O próprio VI captura imagens de alta resolução da via em cada viagem. Sendo assim, a aplicação poderia fornecer ao usuário uma forma de visualizar as imagens coletadas para um trecho específico, trazendo informações úteis para quem possui o conhecimento técnico necessário.

A quinta é a implementação de uma visualização que permita observar a velocidade com que cada trecho da via se deteriora. Assim seria possível observar regiões críticas que têm desgastes recorrentes, o que sugere um reforço estrutural de trechos específicos.

Finalmente, atentando-se ao 5º requisito da seção 3.2, uma análise dos diferentes atores que utilizariam o sistema poderia prover uma espécie de priorização das camadas e

filtros exibidos pela aplicação. Isso tornaria mais direto o acesso às informações analisadas com maior frequência por determinado usuário.

6.2 CONCLUSÃO

Como conclusão, tem-se que todos os objetivos descritos na seção 1.1 foram alcançados pelo grupo. Foi elaborado um sistema que permite dar uma finalidade prática para as informações coletadas e processadas no escopo do projeto da Vale. Em paralelo, o background teórico constituído ao longo do projeto permitiu um entendimento sólido do problema, resultando na utilização dos melhores meios para solucioná-lo.

Além disso, houve um reconhecimento da utilidade da aplicação pelos próprios funcionários da Vale. Com toda a estrutura apresentada, o resultado desse trabalho poderia:

- Aumentar a produtividade, por meio da viabilidade de se adotar manutenções preditivas na via permanente;
- Aumentar a confiabilidade dos ativos, como consequência da diminuição do risco de acidentes;
- Reduzir custos com a rotina de manutenção, que atualmente é embasada em manutenção corretiva e preventiva;
- Promover um maior conhecimento quanto ao comportamento mecânico global da via, por meio das ferramentas de análise disponibilizadas.

É válido ressaltar que nem todas as funcionalidades levantadas foram implementadas. Entretanto, todos os principais requisitos foram atendidos e uma série de análises interessantes pode ser realizada. Portanto, considera-se que o grupo pôde firmar uma aplicação coerente e satisfatória para o escopo do trabalho de conclusão de curso.

6.2.1 Contribuições

A principal contribuição do projeto de formatura realizado, decorrente do estudo bibliográfico feito, é o levantamento e a análise das características e funcionalidades necessárias

para um sistema de Big Data Risk Analysis (BDRA), especificamente no contexto de sistemas ferroviários. Ao longo deste documento foi mencionada uma série de estudos que mostram desde a tendência e as vantagens da implementação de modelos de ML para a adoção de manutenção preditiva da via permanente, até os requisitos comuns que todo engenheiro de software deve ter em mente na hora do planejamento de uma aplicação de análise de riscos para esses sistemas.

Indo mais além, a arquitetura, os frameworks e as bibliotecas utilizadas representam o estado da arte do cenário tecnológico atual. Sendo assim, outra contribuição do projeto é fornecer uma base para a infraestrutura de qualquer sistema de visualização e análise de dados que tenha como funcionalidade base a implementação de um mapa interativo.

REFERÊNCIAS BIBLIOGRÁFICAS

- 1 BERNUCCI, L. L. B. Big data analytics aplicado aos parâmetros de vagão instrumentado e carro controle no apoio da gestão de manutenção de via permanente. 2020.
- 2 GHOFRANIA, F. et al. Recent applications of big data analytics in railway transportation systems: A survey. 2018. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0968090X18303395>.
- 3 ANTT - Vale - Estrada de Ferro Vitória a Minas. ANTT, 2021. Disponível em: <https://portal.antt.gov.br/vale-estrada-de-ferro-vitoria-a-minas>.
- 4 VALE. *Estrada de Ferro Vitória a Minas completa 111 anos sobre os trilhos*. 2015. Disponível em: <https://www.vale.com/brasil/PT/aboutvale/news/Paginas/estrada-ferro-vitoria-minas-completa-111-anos-sobre-trilhos.aspx>.
- 5 WIKIPEDIA. *Estrada de Ferro Vitória a Minas*. 2020. Disponível em: https://pt.wikipedia.org/wiki/Estrada_de_Ferro_Vit%C3%B3ria_a_Minas.
- 6 GOLFARELLI, M.; RIZZI, S. A model-driven approach to automate data visualization in big data analytics. 2020. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/1473871619858933>.
- 7 FUMEO, E.; ONETO, L.; ANGUIA, D. Condition based maintenance in railway transportation systems based on big data streaming analysis. 2015. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1877050915018244>.
- 8 COIMBRA, M. do V. Modos de falha dos componentes da via permanente ferroviária e seus efeitos no meio ambiente. 2008. Disponível em: https://transportes.ime.eb.br/DISSERTA%C3%87%C3%95ES/Coimbra_Marcelo_Dissert_Mestrado.pdf.
- 9 AREMA. *Manual for Railway Engineering*. 2020.
- 10 LEITE, M. R. S. V. D. Especificações técnicas da via-férrea. 2017. Disponível em: <https://hdl.handle.net/10400.22/12054>.
- 11 ABNT. *NBR 16387: Via férrea - Classificação de vias*. 2016.
- 12 SILVA, T. S. dos S. Inspeção e reabilitação de infraestruturas ferroviárias. 2012. Disponível em: <https://hdl.handle.net/10362/8505>.
- 13 RIBEIRO, R. de Alvarenga Rosa e R. C. H. *Estradas de ferro : projeto, especificação construção*. Edufes, 2016. Disponível em: <https://repositorio.ufes.br/handle/10/11609>.
- 14 FIGUERES-ESTEBAN, M.; HUGHES, P.; GULIJK, C. V. The role of data visualization in railway big data risk analysis. 2015. Disponível em: <https://eprints.hud.ac.uk/id/eprint/25009/>.

- 15 WHAT Is Storytelling. National Storytelling Network, 2021. Disponível em: <https://storynet.org/what-is-storytelling/>.
- 16 DATA-DRIVEN Storytelling: 9 Techniques for Effective Visualization. kantalupe, 2020. Disponível em: <https://www.gokantaloupe.com/blog/best-techniques-for-data-driven-storytelling/>.
- 17 DJANGO. Django Software Foundation, 2021. Disponível em: <https://www.djangoproject.com/>.
- 18 SPHINX. Sphinx team, 2021. Disponível em: <https://www.sphinx-doc.org/>.
- 19 SWAGGER. SmartBear Software, 2021. Disponível em: <https://swagger.io/>.
- 20 MAPBOX. Mapbox, 2021. Disponível em: <https://www.mapbox.com/>.
- 21 REACT. Facebook Inc, 2021. Disponível em: <https://reactjs.org/>.
- 22 ANGULAR. Google, 2021. Disponível em: <https://angular.io/>.
- 23 VUEJS. Evan You, 2021. Disponível em: <https://vuejs.org/>.
- 24 MOST used web frameworks among developers worldwide, as of 2021. Statista, 2021. Disponível em: <https://www.statista.com/statistics/1124699/worldwide-developer-survey-most-used-frameworks-web/>.
- 25 CROSS-PLATFORM mobile frameworks used by software developers worldwide from 2019 to 2021. Statista, 2021. Disponível em: <https://www.statista.com/statistics/869224/worldwide-software-developer-working-hours/>.
- 26 BOOTSTRAP. Bootstrap team, 2021. Disponível em: <https://getbootstrap.com/>.
- 27 MATERIAL Design. Google, 2021. Disponível em: <https://material.io/develop/web/getting-started>.
- 28 HARLEY, A. Icon usability. 2014. Disponível em: <https://www.nngroup.com/articles/icon-usability/>.
- 29 CHARTJS. ChartJs team, 2021. Disponível em: <https://www.chartjs.org/>.
- 30 D3JS. Mike Bostock, 2021. Disponível em: <https://d3js.org/>.
- 31 TURF. Turf team, 2021. Disponível em: <https://turfjs.org/>.
- 32 KASHYAP, V. *Data Science Workflow: From Research Experiments to Business Use-Cases*. 2019. Disponível em: <https://www.progress.com/blogs/data-science-workflow-from-research-experiments-to-business-use-cases>.
- 33 ISO. *Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models*. [S.l.], 2017. Disponível em: <https://www.iso.org/obp/ui/#iso:std:iso-iec:25010:ed-1:v1:en>.
- 34 GITLAB. GitLab B.V., 2021. Disponível em: <https://about.gitlab.com/>.