



UNIVERSIDADE D
COIMBRA



Adriana Pereira Barbosa

AS CANÇÕES DO EXÍLIO

Dissertação no âmbito do Mestrado em Design e Multimédia orientada pela Professora Catarina Sofia Henriques Maças e pelo Professor Doutor Pedro José Mendes Martins e apresentada ao Departamento de Engenharia Informática da Faculdade de Ciências e Tecnologias da Universidade de Coimbra.

Setembro de 2018

Dissertação no âmbito do Mestrado em Design e Multimédia
Departamento de Engenharia Informática
Faculdade de Ciências e Tecnologias da Universidade de Coimbra.

AS CANÇÕES DO EXÍLIO

Adriana Pereira Barbosa

Orientação

Catarina Sofia Henriques Maçãs

Pedro José Mendes Martins

Júri

Maria Alice Barriga Geirinhas dos Santos

Fernando Amílcar Bandeira Cardoso

Setembro de 2018

Avó, vod.

AGRADECIMENTOS

Aos meus orientadores, a professora Catarina Maçãs e o professor Pedro Martins, pela orientação, partilha de ideias, motivação, paciência, ajuda e tempo disponibilizados, e ao Joshua Enslin pela oportunidade de participar neste projeto que me fascinou desde o primeiro dia.

À árEa65 pela concentração e pelos intervalos, esse equilíbrio foi importante.

À Filipa e à Cristiana por serem as minhas pessoas, pelas palavras certas nos momentos certos, por aguentarem os meus desabafos, o meu humor singular e por estarem sempre presentes mesmo quando eu não estive. À Elisabete pelos conselhos, conversas, opiniões e fugidas a rotina.

À minha irmã por todo o apoio e ajuda que me deu, por fazer da distância somente um número e lembrar-me do que é importante.

À minha mãe por acreditar em mim, por me deixar sonhar e ir sempre mais longe estando sempre do meu lado independentemente das minhas decisões.

Ao meu pai e a à minha avó pelo apoio nos momentos decisivos.

A todos os “não” que sempre me mostraram que é importante saber o que quero mas ainda mais importante é saber o que preciso.

Obrigada.

RESUMO

Visualização de Informação é uma área que tem estado presente no nosso quotidiano ao longo de vários séculos, sempre em constante transformação. Com a evolução da tecnologia foram descobertas novas formas de analisar dados, bem como os dados disponibilizados foram crescendo em quantidade. Para além disto, o fácil acesso a todo o tipo de dados permitiu que fossem explorados temas que de outra forma não seria possível.

O ponto de partida para esta dissertação é a exposição *Bird Watching: Visualizações da Influência de “Canção do Exílio”*, de Joshua Enslin e Alaina Enslin. Nesta exposição, foram apresentados alguns artefactos analógicos sobre os dados que agora vão ser explorados através de Visualização de Informação, aplicando-a, mais concretamente, na análise e visualização de textos. Estes dados concentram-se à volta do poema “*Canção do Exílio*”, de Antônio Gonçalves Dias, escrito em 1843 enquanto estudante de direito na Universidade de Coimbra. O poema tornou-se bastante popular desde a sua publicação. Essa popularidade levou a que também se tornasse um dos poemas mais reproduzidos na actualidade, com variações e paródias a surgirem todos os dias. Nesta dissertação foram considerados 500 textos (escritos entre 1843 e 2015) para o desenvolvimento de três artefactos visuais que dão a conhecer a história à volta de um dos poemas mais famosos do Brasil.

PALAVRAS-CHAVE

Canção do Exílio, Visualização de Informação, Análise de Texto, Visualização de Texto, Programação, Data Art

ABSTRACT

Information Visualization is an area that has been present in our daily lives for several centuries, and which has been always in transformation. With the evolution of technology, new ways of analyzing data have been discovered, as well as the available data has been growing in quantity. Furthermore, the easy access to all kinds of data allows us to explore topics that otherwise would not be possible.

The starting point for this dissertation was the exhibition *Bird Watching: Views of the Influence of “Canção do Exílio”* by Joshua Enslin and Alaina Enslin. The analog artifacts presented in this exhibition were based on the data that will now be explored through Information Visualization, applying it, more specifically, to the analysis and visualization of text. This data is centered around the poem of Antônio Gonçalves Dias “*Canção do Exílio*”, written in 1843 when the author was studying law in the University of Coimbra. The poem grew in popularity since its publication. This popularity made it become one of the most reproduced poems of these days, with variations and parodies created every day. For this dissertation 500 texts (written between 1843 and 2015) were considered with the aim to develop visual artifacts that transmit information and history around one of the most famous poems of Brazil.

KEYWORDS

Canção do Exílio, Information Visualization, Text Analysis, Text Visualization, Programming, Data Art

ÍNDICE

LISTA DE FIGURAS	xiii
INTRODUÇÃO	3
2 ESTADO DA ARTE	7
2.1 ANÁLISE E REPRESENTAÇÃO DE DADOS	9
2.2 VISUALIZAÇÃO DE INFORMAÇÃO	11
2.2.2 VISUALIZAÇÃO DE TEXTO	31
2.3 CONCLUSÕES DO ESTADO DA ARTE	35
3 A “CANÇÃO DO EXÍLIO”	37
3.1 O POEMA	39
3.2 A EVOLUÇÃO DO POEMA	40
3.3 <i>BIRD WATCHING: VISUALIZAÇÕES DA INFLUÊNCIA DE “CANÇÃO DO EXÍLIO”</i>	46
4 METODOLOGIA	51
4.1 PLANO DE TRABALHO	53
4.2 PROCESSO DE TRABALHO	56
5 PROJECTO	59
5.1 DADOS DO PROJECTO	61
5.2 FERRAMENTAS DE DESENVOLVIMENTO	62
5.3 PRIMEIROS ESBOÇOS	65
5.4 DESENVOLVIMENTO DO PROJETO	69
5.4.1 ESTUDO DE SÍMBOLO E CORES	69
5.4.2 EBOÇOS DAS VISUALIZAÇÕES FINAIS	76
5.4.3 DESENVOLVIMENTO DAS VISUALIZAÇÕES	80
5.4.4 FILTROS	92
5.5 PLATAFORMA FINAL	95
5.6 RESULTADO DA ANÁLISE DAS VISUALIZAÇÕES	100
6 CONCLUSÃO E TRABALHO FUTURO	105
BIBLIOGRAFIA	109
APÊNDICES	119
A - LISTA DE PALAVRAS	121
B - ESTUDOS DE FORMAS E CORES	123
C - INTERAÇÃO DA VISUALIZAÇÃO “O POEMA ...AO LONGO DO TEMPO”	127
D - INTERAÇÃO DA VISUALIZAÇÃO “O POEMA ...PRESENTE NOS TEXTOS”	131
E - INTERAÇÃO DA VISUALIZAÇÃO “O POEMA ...EM PALAVRAS”	139
F - ESBOÇO DIGITAL DA PLATAFORMA	143
G - PLATAFORMA EM INGLÊS	145

LISTA DE FIGURAS

Figura 2.1 Distinção das variáveis de imagem.	10
Figura 2.2 Representações hierárquicas.	13
Figura 2.3 “ <i>Paleontological Tree of Vertebrates</i> ” desenvolvido por Ernst Haeckel, 1879.	14
Figura 2.4 Visualização dos dados correspondentes a Portugal em 2016 no “ <i>The Observatory of Economic Complexity</i> ”.	15
Figura 2.5 “ <i>Freedom in Countries</i> ” de Sara Piccolomini para o jornal <i>la Lettura, Il Corriere della Sera</i> .	16
Figura 2.6 Circular Treemap desenvolvido por Wetzel que representa diferentes tipos de ficheiros.	17
Figura 2.7 Diagramas Relacionais	18
Figura 2.8 Visualização “ <i>The Real Montalbano</i> ” desenvolvido pela empresa Accurat	18
Figura 2.9 “ <i>Shape of Song</i> ” aplicado a um composição de Chopin.	19
Figura 2.10 “ <i>Shape of Song</i> ” aplicado à música “ <i>Like a Prayer</i> ” de Madonna.	19
Figura 2.11 Representação dos aditivos de cada grupo alimentar.	20
Figura 2.12 Representação presença dos aditivos.	20
Figura 2.13 Primeira semana da Stefanie Posavec para o projecto “ <i>Dear Data</i> ”.	21
Figura 2.14 “ <i>Carte Figurative des pertes successives en hommes de l’Armée Française dans la campagne de Russie 1812-1813</i> ” desenvolvido por Charles Minard em 1869.	22
Figura 2.15 Diagramas de representação temporal	22
Figura 2.16 Visualização “ <i>The Swedish Emigrants</i> ”.	23
Figura 2.17 Gráfico de Barras de William Playfair, 1786.	24
Figura 2.18 Diagrama das Causas da Mortalidade de Florence Nightingale.	25
Figura 2.19 Tipos de mapas.	26
Figura 2.20 “ <i>Transparency: Are the Richest Americans Also the Best Educated? America’s Richest Counties and Best Educated Counties</i> ” para a revista <i>GOOD</i> , 2011.	26
Figura 2.21 “ <i>Mapping the World, one Tweet at a Time</i> ”.	27
Figura 2.22 “ <i>Carte figurative et approximative des quantités de viandes de boucherie envoyés sur pied par les départements et consommés à Paris</i> ” desenvolvido por Charles Minard (Palsky, 1996).	28
Figura 2.23 “ <i>Carte figurative et approximative des quantités de coton brut importées en Europe en 1858, en 1864 et en 1865</i> ” desenvolvido por Charles Minard (Mason, 2017).	28
Figura 2.24 “ <i>EU Referendum, A Divided Kingdom</i> ”.	29
Figura 2.25 Representação visual do “ <i>Phantom Terrains</i> ”.	30
Figura 2.26 Imagem da visualização de “ <i>Text Arc: Alice’s Adventures in Wonderland</i> ”.	31
Figura 2.27 “ <i>The Evolution of the Origin of Spieces</i> ” de Stefanie Posavec e Greg MCInerney.	32
Figura 2.28 Cartaz e pormenor do cartaz do evento “ <i>Poetry On The Road</i> ” 2006.	33
Figura 2.29 Representações das obras de Shakespeare.	34
Figura 3.1 Retrato de Gonçalves Dias desenvolvido por Alaina Enslin. (Enslin & Enslin, 2016)	46
Figura 3.2 Ilustração do poema “ <i>Saudade da Pátria</i> ”. (Enslin & Enslin, 2016)	47
Figura 3.3 Gráfico de semelhanças. (Enslin & Enslin, 2016)	48
Figura 3.4 Gradiente de cores desenvolvido a partir de fotografias (Enslin & Enslin, 2016)	49

Figura 4.1 Plano de trabalho proposto.	53
Figura 4.2 Plano de trabalho desenvolvido.	53
Figura 5.1 Experimentação da ferramenta Raphaël.js.	63
Figura 5.2 Desenho das formas exploradas na biblioteca Raphaël.js.	63
Figura 5.3 Primeiros resultados obtidos com a exploração da ferramenta D3.js.	64
Figura 5.4 Protótipo em papel das origens das variações.	65
Figura 5.5 Storyboard em papel do funcionamento da visualização através de fotografias do Instagram.	66
Figura 5.6 Explicação em papel da visualização através de fotografias do Instagram.	67
Figura 5.7 Protótipo em papel das relações das variações com as frases do poema original.	68
Figura 5.9 À esquerda está representada a visualização com as cores normais, à direita está a visualização com as cores do tipo de daltonismo protanopia.	70
Figura 5.10 Estudo inicial de formas e tipo de linha	71
Figura 5.11 Estudo de cor por temas e variação dos sentimentos dos textos.	71
Figura 5.12 Estudo de variação de cores e textura.	72
Figura 5.13 Novo estudo de Símbolos	73
Figura 5.14 Estudo da utilização dos diferentes símbolos na primeira visualização	73
Figura 5.15 - Cores finais que diferenciam os tipos de dados a representar.	73
Figura 5.16 Cores finais que diferenciam os períodos governativos do Brasil.	74
Figura 5.17 Cores finais que representam os movimentos artísticos e literários.	74
Figura 5.18 Combinação dos dados Negativos com os movimentos artísticos.	74
Figura 5.19 Cores representativas dos eventos mundiais.	75
Figura 5.20 Cores representativas dos eventos mundiais aplicadas no artefacto.	75
Figura 5.21 Esquema da visualização “O Poema ...ao longo do tempo”.	76
Figura 5.22 Esboço em papel da visualização “O Poema ...ao longo do tempo”.	77
Figura 5.23 Esboço da visualização “O Poema ...em palavras”.	78
Figura 5.24 Esquema do planeamento da visualização “O Poema ...em palavras”.	78
Figura 5.25 Esquema do funcionamento da visualização “O Poema ...presente nos textos”.	79
Figura 5.26 Esboço da visualização “O Poema ...presente nos textos” utilizando os textos de Carlos Drummond de Andrade.	80
Figura 5.27 Etapa inicial do desenvolvimento da visualização “O Poema ...ao longo do tempo”.	81
Figura 5.28 Criação da delimitação dos períodos históricos na visualização “O Poema ...ao longo do tempo”.	82
Figura 5.29 Aplicação das áreas sombreadas dos períodos históricos na visualização “O Poema ...ao longo do tempo”.	82
Figura 5.30 Artefacto e filtros correspondentes a’ “O Poema ...ao longo do tempo”.	
Figura 5.31 Visualização “O Poema ...ao longo do tempo” finalizada.	83
Figura 5.32 Representação dos períodos históricos do Brasil.	84
Figura 5.33 Representação dos Movimentos artísticos literários	84
Figura 5.34 Representação dos eventos históricos mundiais.	84
Figura 5.35 Representação do resultado da pesquisa para o poema com o título “Fim de Festa” com mouseover sobre o poema para aparecer o tooltip no canto superior direito do gráfico.	85

Figura 5.36 Representação inicial da visualização “O Poema ...presente nos textos”.	86
Figura 5.37 Artefacto já com os elementos visuais e de interação definidos.	86
Figura 5.38 Versão final do artefacto “O Poema ... presente nos textos”.	87
Figura 5.39 Representação das variações correspondentes ao período Império.	87
Figura 5.40 Realce do resultado da procura	88
Figura 5.41 Primeiro modelo de representação da visualização “O Poema ...em palavras”.	89
Figura 5.42 Segundo modelo de representação da visualização “O Poema ...em palavras”.	89
Figura 5.43 Terceiro modelo de representação da visualização “O Poema ...em palavras” já com as palavras das variações.	90
Figura 5.44 Quarto modelo de representação da visualização “O Poema ...em palavras”.	90
Figura 5.45 Representação final da visualização “O Poema ...em palavras”.	91
Figura 5.46 Representação da funcionalidade de tooltip com mouseover sobre a palavra triste.	92
Figura 5.47 Esboço em papel da plataforma.	95
Figura 5.48 Estrutura da plataforma.	96
Figura 5.49 Página inicial da plataforma final.	97
Figura 5.50 Página com explicação sobre o poema de Gonçalves Dias e sobre o projeto de Joshua Enslen.	98
Figura 5.51 Página final da visualização “O Poema ...ao longo do tempo”.	98
Figura 5.52 Página final da visualização “O Poema ...presente nos textos”.	99
Figura 5.53 Página final da visualização “O Poema ...em palavras”.	99
Figura 5.54 Página descritiva do projeto.	99
Figura 5.55 Representação das variações positivas.	101
Figura B1 Aplicação das formas desenvolvidas.	123
Figura B2 Exploração de novas formas para os diferentes tipos de opção visual.	123
Figura B3 Novo estudo de formas mais complexas com distinção por sentimento.	124
Figura B4 Estudo da sobreposição de símbolos consoante o seu tipo e a sua classificação	124
Figura C1 “O Poema ...ao longo do tempo” com a representação das variações “positivas”.	127
Figura C2 Representação das variações “negativas” no artefacto “O Poema ...ao longo do tempo”.	127
Figura C3 Representação das variações “outras” no artefacto “O Poema ...ao longo do tempo”.	128
Figura C4 Representação das variações “negativas” e “outras”.	128
Figura C5 Representação das variações “positivas” e “negativas”.	128
Figura C6 Representação das variações “positivas” e “outras”.	129
Figura C7 Procura com sugestões a partir do texto escrito pelo utilizador.	129
Figura D1 Realce com mouseover de uma variação.	131
Figura D2 Representação dos poemas classificados como “positivos”.	131
Figura D3 Representação dos poemas classificados como “negativos”.	132
Figura D4 Representação dos poemas classificados como “outras”.	132
Figura D5 Representação da combinação dos poemas classificados como “positivos” e “negativos”.	132

Figura D6 Representação da combinação dos poemas classificados como “positivos” e “outros”.	133
Figura D7 Representação da combinação dos poemas classificados como “outros” e “negativos”.	133
Figura D8 Representação das variações correspondentes ao período Primeira República.	133
Figura D9 Representação das variações correspondentes ao período Estado Novo.	134
Figura D10 Representação das variações correspondentes ao período Quarta República.	134
Figura D11 Representação das variações correspondentes ao período Ditadura Militar.	134
Figura D12 Representação das variações correspondentes ao período Sexta República no século XX.	135
Figura D13 Representação das variações correspondentes ao período Sexta República no século XXI.	135
Figura D14 Representação da função de pesquisa no artefacto.	135
Figura D15 Resultado da procura efectuada.	136
Figura D16 Realce de texto com informação correspondente	136
Figura E1 Representação da visualização “O Poema ...em palavras” no período Império.	139
Figura E2 Representação da visualização “O Poema ...em palavras” no período Primeira República.	139
Figura E3 Representação da visualização “O Poema ...em palavras” no período Estado Novo.	140
Figura E4 Representação da visualização “O Poema ...em palavras” no período Quarta República.	140
Figura E5 Representação da visualização “O Poema ...em palavras” no período Ditadura Militar.	140
Figura E6 Representação da visualização “O Poema ...em palavras” no período Sexta República do séc.XX.	141
Figura E7 Representação da visualização “O Poema ...em palavras” no período Sexta República do séc.XXI.	141
Figura G1 Página Inicial	145
Figura G2 Página sobre a “Canção do Exílio” e a exposição	145
Figura G3 Página da primeira visualização.	146
Figura G4 Página da segunda visualização.	146
Figura G5 Página da terceira visualização.	147
Figura G6 Página descritiva do projecto.	147

INTRODUÇÃO

INTRODUÇÃO

Escrito por Antônio Gonçalves Dias em 1843, a “*Canção do Exílio*” é um dos poemas que mais motivaram a criação de variações e paródias no Brasil. Desde que foi publicado, tem sido utilizado tanto como inspiração a variações de visão romântica sobre o Brasil, como a variações que criticam e mostram o lado menos positivo de momentos de crise económica e política do Brasil. Com a recolha de dados relativos às variações do poema, foi encontrada a necessidade de mostrar estes textos aos utilizadores, de forma a que eles consigam compreender a evolução das variações do poema, assim como a sua desconstrução e manipulação. Através de técnicas de Visualização de Informação, pretendemos transformar toda a complexidade existente nos 174 anos de informação relacionada com a “*Canção do Exílio*”, em artefactos visuais simples e acessíveis para o utilizador comum.

O objetivo da presente dissertação é, através do desenvolvimento de artefactos visuais, mostrar e permitir a exploração das variações do poema “*Canção do Exílio*” de forma a dar a conhecer estes textos. Desenvolveram-se um conjunto de visualizações que mostram a evolução das variações, a forma como se relacionam com a época em que foram escritas e que influências estão presentes nos textos. Para o desenvolvimento dos artefactos exploraram-se técnicas de visualização de texto e de análise textual.

Objectivo

Como já mencionado anteriormente, a “*Canção do Exílio*” é um dos poemas com mais variações no Brasil (Lítera, 2015). Graças às novas tecnologias, Joshua Enslin conseguiu reunir 500 variações que têm como ponto de partida o poema de Gonçalves Dias. Estas variações datam entre 1843 e 2015. Este conjunto de dados consegue contar a história do país através de poemas. Com a presente dissertação, pretendemos mostrar esses dados e a história neles existente através da visualização de texto. A visualização de texto é uma área da Visualização de Informação que ainda se encontra em expansão, sendo que o desenvolvimento do projeto é um desafio.

A motivação pessoal para a escolha desta dissertação parte do conhecimento prévio do poema “*Canção do Exílio*” bem como do desafio que é representar visualmente dados textuais. Ademais, existe uma enorme vontade em explorar e ampliar conhecimentos na área da Visualização de Informação sendo que a área de visualização de texto permitiu essa exploração e ampliação de conhecimentos.

Motivação

O resto desta dissertação está dividida nos seguintes capítulos: Estado da Arte, A “*Canção do Exílio*”, Metodologia, Projeto e Conclusão e Trabalho Futuro. No Estado da Arte, é feita uma revisão bibliográfica sobre a principal área de foco da presente dissertação — Visualização de Informação — começando com uma breve descrição sobre esta disciplina. A descrição serve como ponto de partida para a explicação dos diferentes tipos de visualização. No subcapítulo Visualização de Texto são abordadas formas de análise de grandes quantidades de dados textuais e são apresentados projetos

Estrutura

desenvolvidos com diferentes tipos de dados textuais. Este último tópico é de grande importância, pois é necessário analisar e compreender as melhores formas de representar os dados. A “Canção do Exílio” é o capítulo onde é feita uma breve descrição do poema “*Canção do Exílio*”, e onde apresentamos a evolução do poema ao longo dos anos com a seleção e análise temática de algumas das variações. Neste capítulo é também apresentado o projeto de Joshua Enslin sobre o mesmo tema. No quarto capítulo, Metodologia, é apresentado e descrito o plano de trabalhos e o processo seguido para o desenvolvimento do projeto. No quinto capítulo, Projeto, são identificados os dados que foram trabalhados e utilizados no desenvolvimento dos artefactos. Em seguida, mostramos as propostas apresentadas no primeiro semestre das variações. Apresentamos também as ferramentas que foram consideradas antes de se iniciar o desenvolvimento do projeto. No desenvolvimento do projeto começamos por apresentar o estudo de cores e símbolos feito de forma a encontrar a melhor solução para a representação dos dados. Em seguida, descrevemos o processo de desenvolvimentos dos novos esboços a partir da análise das visualizações e o desenvolvimento das mesmas. Na secção seguinte falamos da plataforma que engloba todo o resultado prático desta dissertação e, de seguida, das conclusões que retirámos das visualizações. No sexto capítulo, Conclusão e Trabalho Futuro, são identificadas as ideias finais sobre o trabalho que foi desenvolvido e é descrito o trabalho futuro, designadamente o que ainda poderá ser feito para melhorar e enriquecer o projeto.

2 ESTADO DA ARTE

Neste capítulo, abordamos a representação e análise de dados e descrevemos o que é a Visualização de Informação. São abordados diferentes tipos de visualização, dando maior ênfase à visualização de texto, o que inclui a apresentação de trabalhos desenvolvidos na área.

2.1 ANÁLISE E REPRESENTAÇÃO DE DADOS

Com a recolha e análise de dados surgiu a necessidade de criar representações visuais que auxiliam a compreensão dos dados analisados. Atualmente, o acesso a grandes quantidades de dados é simples e a tecnologia tornou-se uma componente necessária na análise e processamento da informação. Além do acesso simples, o tipo de dados disponíveis para análise também é bastante diversificado. Estes podem ser sobre a utilização de redes sociais, movimentos de pessoas ou evolução de doenças.

Para facilitar a compreensão do tipo de dados, estes podem ser inseridos em diferentes categorias: *Nominal*, *Ordinal* e *Quantitativa*. A categoria *Nominal*, segundo Ward et al. (2015) é para dados de valor não numérico. Estes podem ser extraídos de listas finitas seguindo uma ordem ou, no caso de grandes conjuntos de dados, não estarem dispostos segundo uma ordem. Para Card, Mackinlay e Shneiderman (1999) a categoria *Nominal* são conjuntos de valores não numéricos ordenados. Na categoria *Ordinal*, Ward et al. (2015) define como sendo a categoria onde se inserem os dados de valor numérico que podem ser binários, discretos ou contínuos. Na definição de Card et al. (1999) os valores ordinais são conjuntos ordenados de dados que estão relacionados entre si. Os dados podem ainda inserir-se na categoria *Quantitativa* (Card et al., 1999). Estes dados estão inseridos num intervalo numérico. Esta categoria auxilia a definição dos valores de cada eixo das visualizações.

Os dados podem ainda ser considerados como discretos ou contínuos. Os dados são considerados discretos quando são valores individuais e contínuos quando os valores que existem num intervalo numérico (Meirelles, 2013).

A análise de dados é importante para desenvolver representações visuais mais fidedignas. Por vezes os dados que estamos a trabalhar têm erros. É necessário eliminá-los para que não surjam erros no artefacto que vai representar os dados (Ward et al., 2015). A representação dos dados analisados implica também a aplicação das variáveis visuais de forma acertada (Bertin, 1999). Os elementos básicos representativos são: o ponto, a linha e o plano. O ponto é útil na representação de locais ou para marcar um valor no espaço. A linha serve para revelar relações, marcar distâncias ou indicar direcções. O plano ajuda a limitar o espaço e a aplicar a escala correta dos elementos. Jaques Bertin (1984) apresenta quatro formas de distinção de variáveis: *Associativa*, *Ordenada*, *Seletiva*, *Quantitativas*. As variáveis são *Associativas* quando podem ser inseridas num determinado grupo. São *Ordenadas* se o tipo de dados for possível de ordenar. Variáveis *Seletivas* são as que podem ser diferenciadas dentro de um grupo. Por fim, as variáveis *Quantitativas* são aquelas que têm uma leitura numérica.

As variáveis de imagem e diferenciação, que segundo Bertin (1999) podem ser aplicadas são: a *Posição* (x,y) no plano, o *Tamanho*, o *Valor*, *Textura*, *Cor*, *Orientação* e *Forma*. As variáveis têm uma organização

sequencial e devem ser adequadas ao tipo de dados que se está a trabalhar. Enquanto a variável de *Posição* pode ser aplicada em todo o tipo de dados, a *Forma* e a *Orientação* devem ser aplicados em tipos de dados *Associativos* (Bertin, 1999) (Figura 2.1).

	Associativa	Ordenada	Selectiva	Quantitativa
Posição (x,y)				
Tamanho				
Valor				
Cor				
Textura				
Forma				
Orientação				

Figura 2.1 Distinção das variáveis de imagem.

2.2 VISUALIZAÇÃO DE INFORMAÇÃO

Visualização de informação é uma ferramenta que permite a representação de grandes quantidades de dados. O seu principal objetivo é a transmissão de informação, de forma compreensível, ao público alvo. A visualização de dados é assim um meio para transmitir dados complexos e de grandes dimensões de forma acessível. Representar graficamente os dados serve dois propósitos, a comunicação de ideias e a descoberta dessas mesmas ideias ou de novas ideias (Card et al., 1999). As visualizações procuram ser a resposta, mas podem vir a gerar mais questões sobre os dados representados. Atualmente, os gráficos já não são apenas uma representação simplificada dos dados, são um ponto de partida para a exploração das simplificações. Graças à sua flexibilidade, os gráficos tornaram-se um instrumento de processamento de informação (Bertin, 1999). O papel de quem desenvolve as visualizações é encontrar a melhor solução para o tipo de dados que está a trabalhar.

Os primeiros diagramas mostram-nos que a humanidade sempre teve mais facilidade em compreender o que a rodeia através de imagens. Estes eram uma forma de facilitar a compreensão do mundo como espaço às sociedades antigas que procuravam aumentar o seu conhecimento relativo à posição das estrelas e ao desenvolvimento de mapas.

Ao longo dos anos as representações de dados têm acompanhado a evolução tecnológica. Inicialmente as representações eram estáticas. Atualmente, os sistemas de visualização são interativos e visualmente dinâmicos (Meirelles, 2013). Desta forma, as representações visuais de dados complexos permitem à sociedade a aquisição de conhecimento mais aprofundado do que existe no mundo, como a evolução do clima ou o movimento de pessoas.

As visualizações devem despertar no observador capacidades de reação como a memorização de informação e a pesquisa/manipulação de informação. A forma como a informação é apresentada vai influenciar a maneira como o utilizador compreende os dados, dependendo do destaque que é dado aos mesmos (Meirelles, 2013). Por este motivo, as variáveis visuais representadas devem estar de acordo com as propriedades dos dados (Mackinlay, 1986).

O desenvolvimento de uma visualização pode ser feita de diversas maneiras, porém, para que estas sejam consideradas boas visualizações dos dados, é importante seguir um conjunto de passos, como os sugeridos por Ben Fry (2008), antes mesmo de escolher o tipo de representação que vamos desenvolver. Para criar e desenvolver uma visualização é importante conhecer os dados que estamos a trabalhar, ou seja, fazer uma análise dos mesmos. O estudo e compreensão dos dados envolve as áreas de informática, estatística e design gráfico (Fry, 2008). É também necessário organizar a informação recolhida da análise dos dados para de seguida

definir o tipo de visualização que mais se adequa aos objetivos definidos previamente. A definição de objetivos permite que a visualização transmita a mensagem pretendida de forma eficiente aos utilizadores.

2.2.1 TIPOS DE VISUALIZAÇÃO

Como referido no capítulo anterior, um conjunto de dados pode ter várias representações como solução, sendo que algumas são mais apropriadas do que outras. Ao longo dos anos novas técnicas têm sido desenvolvidas e aperfeiçoadas. Nos primeiros séculos, as representações incidiam principalmente sobre dados demográficos e cartográficos, mas, nos dias de hoje existem mais áreas de interesse e, conseqüentemente, maiores quantidades de dados. Este crescimento no interesse e relevância da Visualização de Informação fez com que fossem desenvolvidas mais visualizações, que abordam diversos temas como consumo ou relações sociais (Tufté, 2006).

Existem muitas alternativas para representar o mesmo tipo de dados. De forma a escolher o tipo de visualização que melhor se adequa ao tipo de dados, devemos distinguir os seguintes tipos de visualização: Mapas, Diagramas, ou Redes (Bertin, 1999). Nos Mapas, os dados são posicionados no espaço/plano de acordo com informação fornecida. O desenvolvimento de mapas temáticos tem como objetivo a descoberta de padrões geográficos e temáticos (Robinson, 1982). Os Diagramas permitem a comparação de dados de uma categoria com as outras categorias presentes no gráfico, ou a visualização da evolução temporal de determinada categoria. As representações de dados através de Redes são compostas por um conjunto de nós (variáveis) e as suas ligações. Este tipo de visualizações pressupõem relação entre nós (e.g.: hierarquia). Não existem limites nas relações destas variáveis, ou seja, um nó pode estar ligado a todas as outras variáveis ou a nenhuma.

Os tipos de visualizações mencionados anteriormente abrangem um conjunto variado de modelos gráficos. Estes podem ser inseridos em diferentes categorias de organização espacial, dependendo do que representam. Existem cinco tipos de composição: *Linear*, *Ortogonal*, *Circular*, *Polares*, e de *Organização*. As *Composições Lineares* representam os dados apenas num plano, horizontal ou vertical. O mapeamento dos dados

é feita de forma proporcional entre si, fazendo com que o total esteja subentendido. As *Composições Ortogonais* utilizam os dois planos (x, y) para fazer a representação da evolução temporal de um tipo de dados. Este tipo de composição serve principalmente para fazer a comparação entre categorias ou espaços temporais, nunca existindo uma noção total dos dados representados. Nas *Composições Circulares* o valor dos dados são representados através do ângulo medido no centro da representação circular e o tamanho do arco. Nestas representações o total é apresentado, sendo fácil para o utilizador comparar a informação representada. As *Composições Polares* representam os dados, criando artefactos visuais que os distribuem de forma circular. Nestas representações os eixos (x, y) são usados para representar quantidade. As *Composições de Organização* representam os dados organizados, geralmente por hierarquias, dos mais importantes para os menos importantes. Estas representações podem facilitar o processo de compreensão e interpretação da informação.

A diversidade existente de gráficos serve o propósito de representar os mais variados tipos de dados. Os dados podem ser divididos da seguinte forma: dados hierárquicos; dados relacionais; dados temporais; dados geográficos. Para cada um destes existe um conjunto de visualizações mais apropriadas para os representar. As estruturas hierárquicas, adequadas a dados hierárquicos, organizam a informação da mais relevante para a mais elementar. Estas apresentam os relacionamentos hierárquicos e a influência dos dados. Alguns dos gráficos mais utilizados na representação de hierarquias são: *Diagramas de Árvore*; *Diagrama Treemap*; *Radial Treemap*; *Circular Treemap* (Figura 2.2).

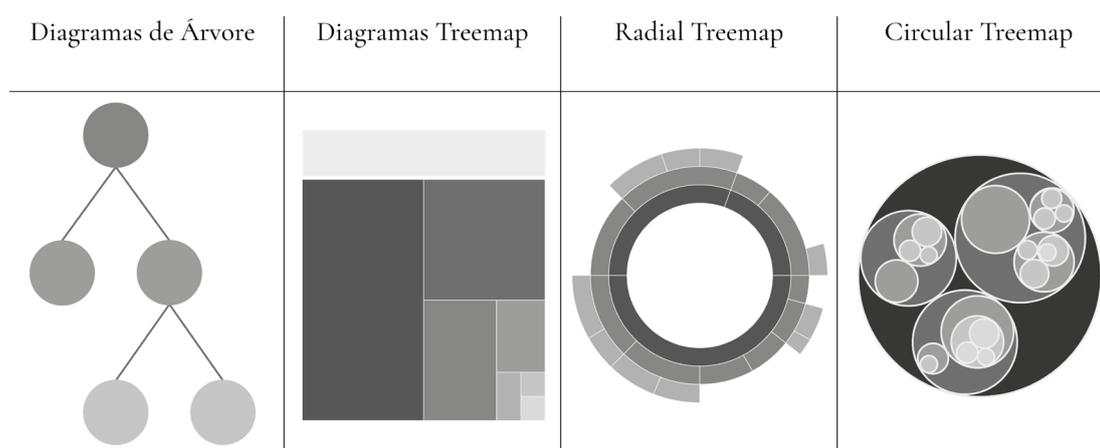


Figura 2.2 Representações hierárquicas.

Os *Diagramas de Árvore*, são um dos métodos mais representativos das estruturas hierárquicas e são bastante utilizados na representação de dados evolutivos. Os sistemas hierárquicos organizam os dados de forma a que os elementos apresentem uma dada relação entre si (Meirelles, 2013). As visualizações deste grupo inserem-se nas representações de Redes e pertencem às Composições de organização. Em 1879, Ernst Haeckel desenvolveu um dos primeiros diagramas de árvore. Com o título “*Paleontological Tree of Vertebrates*” (Figura 2.3) este diagrama representa a evolução dos seres vivos (Quarto Creates, 2014).

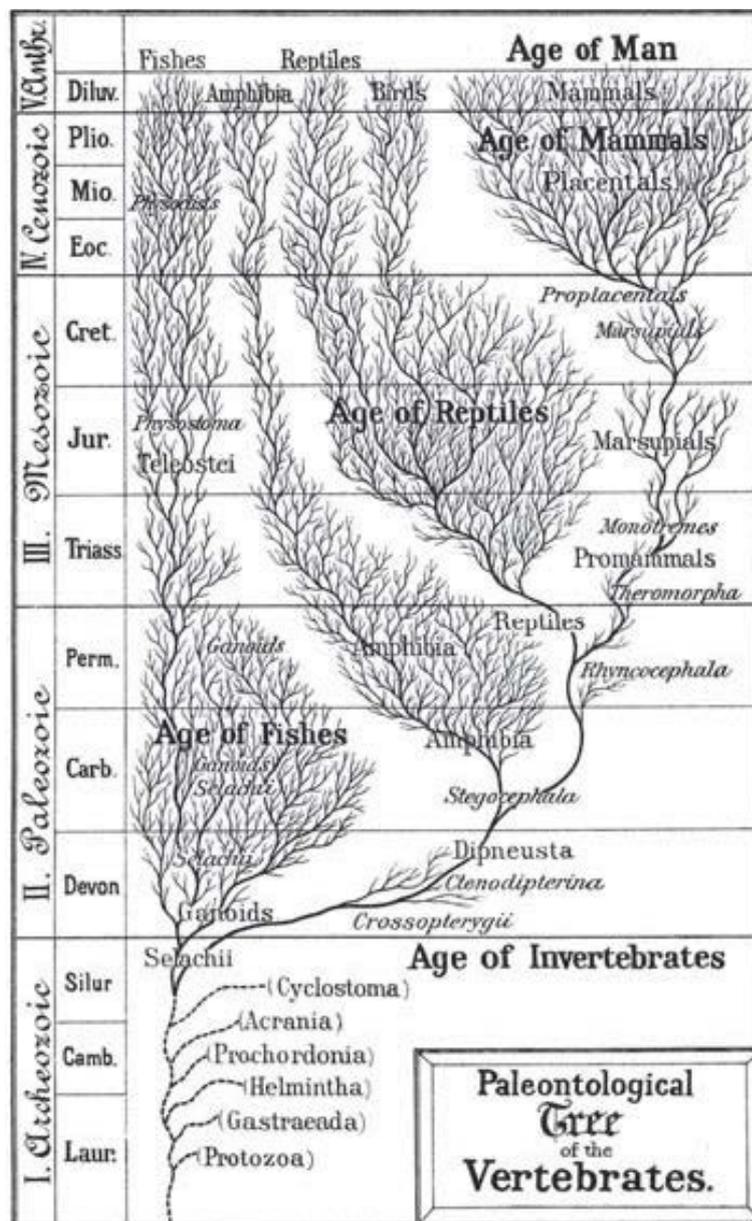


Figura 2.3 “*Paleontological Tree of Vertebrates*” desenvolvido por Ernst Haeckel, 1879.

O *Diagrama Treemap*, pode ser considerado uma evolução do anterior. Dispõe a informação em blocos retangulares, cujo tamanho varia consoante a informação dos dados. Este tipo de gráficos insere-se na categoria de Diagramas e de organização espacial Linear. “*The Observatory of Economic Complexity*” (Figura 2.4) desenvolvido por Alexander Simoes no MIT Media Lab é uma ferramenta que permite visualizar os dados sobre trocas de bens de vários países. O utilizador tem a possibilidade de escolher o ano e o país. É possível visualizar dados desde 1962 até 2016. Pode ser escolhido um único país e visualizar um ou mais anos. É também possível ver o conjunto de todos os países num ou mais anos. A ferramenta permite ainda que se vejam os dados com diferentes níveis de detalhe (Simoes & Hidalgo, 2011).

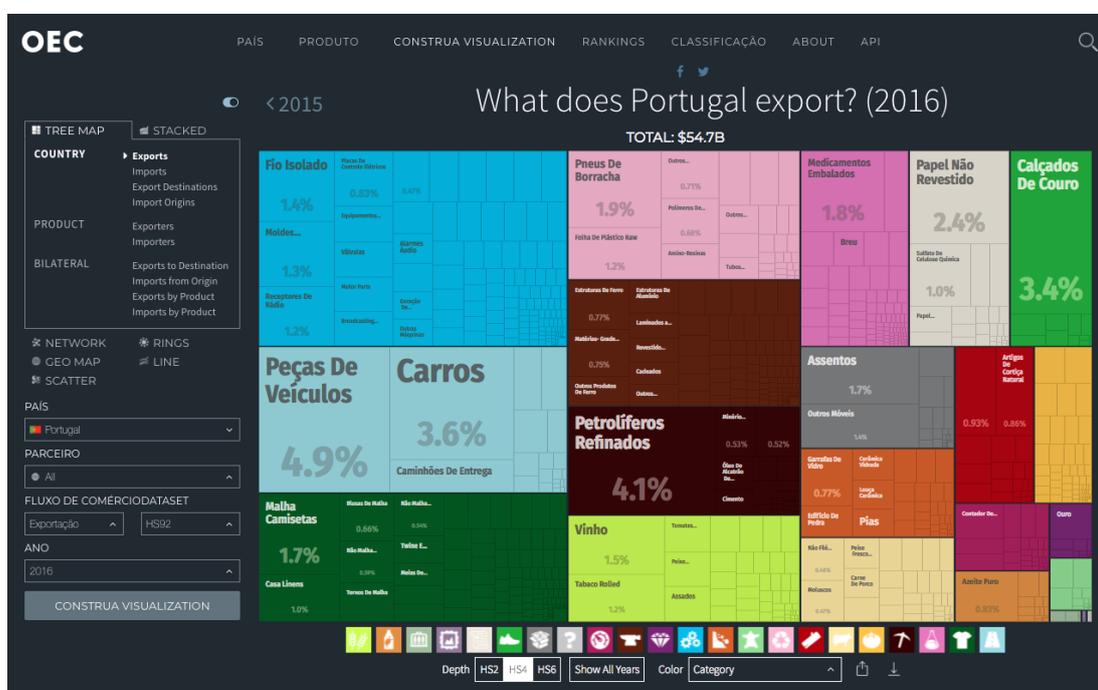


Figura 2.4 Visualização dos dados correspondentes a Portugal em 2016 no “*The Observatory of Economic Complexity*”.

O *Radial Treemap* funciona como o *Diagrama de Árvore* com a particularidade de ter forma circular. A sua informação é disposta em círculo, em que os níveis são colocados de dentro para fora do círculo. Estas representações inserem-se na categoria de Diagramas e de organização espacial Circular. A designer Sara Piccolomini desenvolveu em 2015 a visualização “*Freedom in Countries*” (Figura 2.5) para o suplemento cultural *La Lettura* do jornal italiano *Il Corriere della Sera*. Neste projeto é representado o grau de liberdade de cada país publicado pelo Freedom House. São representadas a liberdade de expressão e a liberdade civil (Piccolomini, 2015).

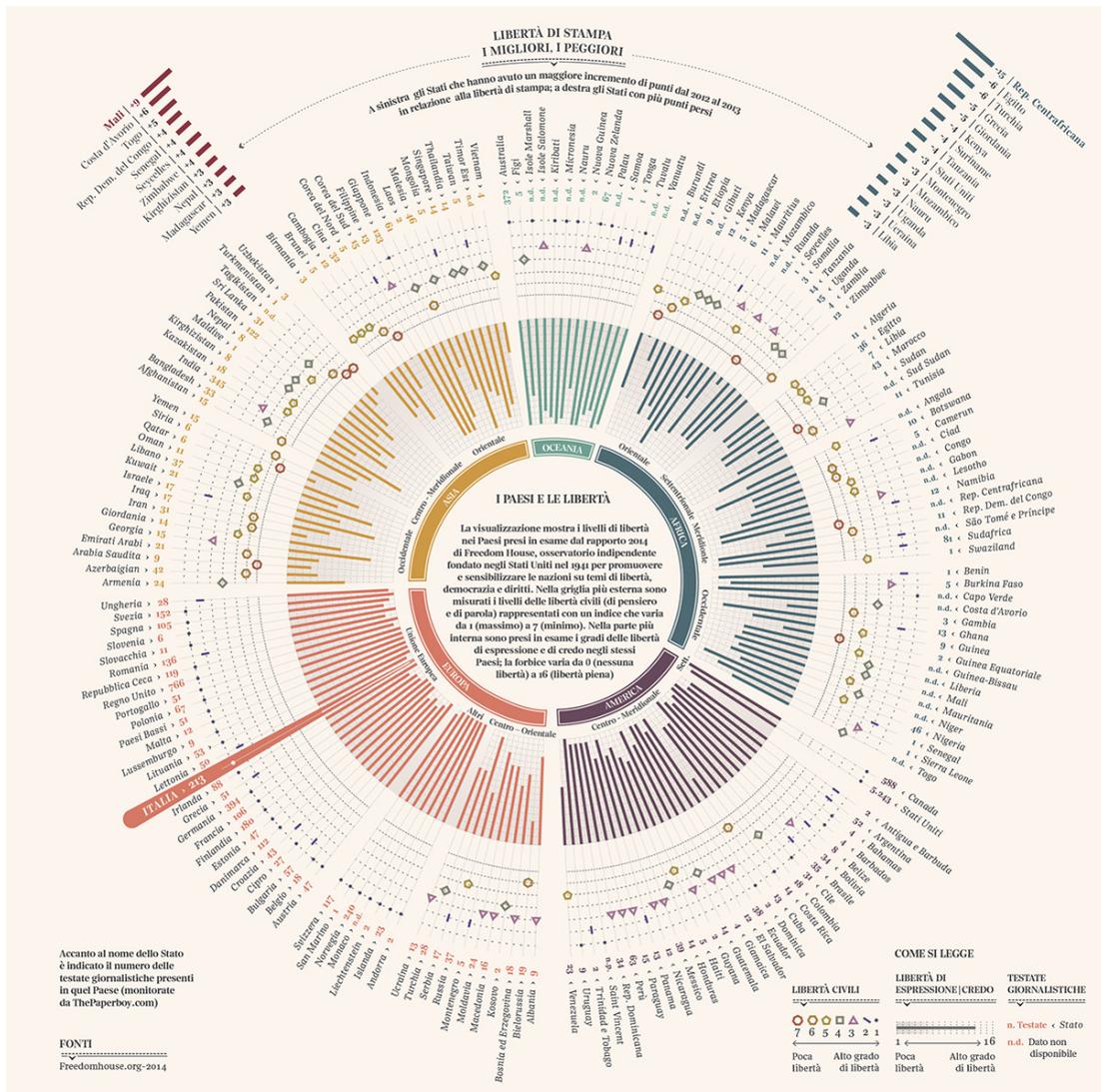


Figura 2.5 - "Freedom in Countries" de Sara Piccolomini para o suplemento cultural la Lettura, do jornal Il Corriere della Sera.

O *Circular Treemap*, é um diagrama circular que funciona como o Treemap, criando círculos dentro de círculos em que a área destes varia segundo os dados a representar (Ribecca, n.d.). Estes gráficos pertencem ao grupo dos Diagramas e de organização espacial Polar. Kai Wetzel explorou várias possibilidades de representação de dados através de Circular Treemaps, estas representações mostram o tamanho de ficheiros, a data de acesso e o tipo de ficheiros (Wetzel, n.d.) (Figura 2.6).

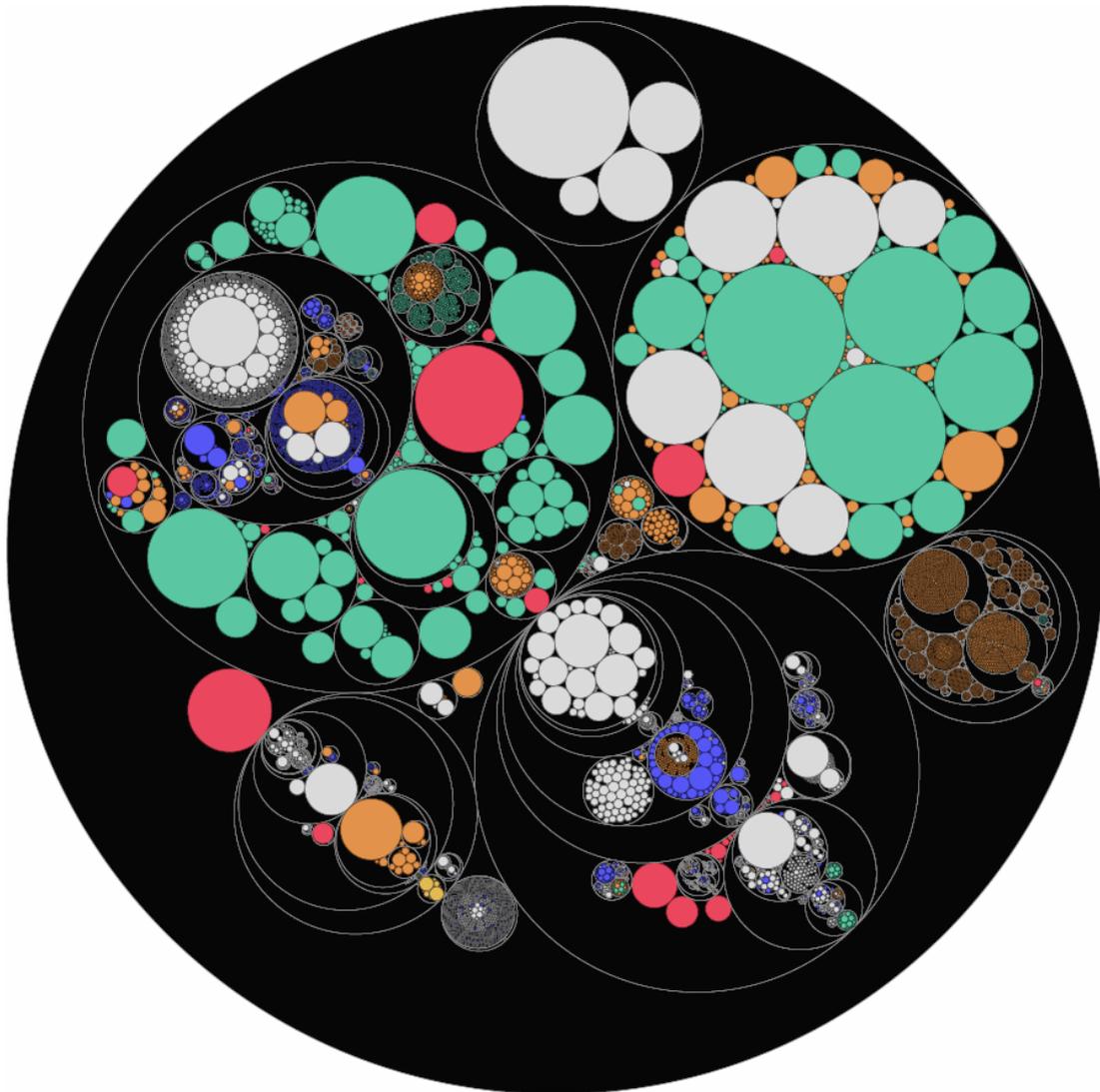


Figura 2.6 Circular Treemap que representa diferentes tipos de ficheiros.

A representação das relações existentes nos dados torna mais intuitiva a compreensão e a descoberta de padrões (Meirelles, 2013). Os gráficos que permitem a visualização deste tipo de dados são: *Gráficos de Coordenadas Paralelas*; *Gráficos em Arco*; *Diagramas Circulares*; *Gráficos Polares ou Radiais*; e os *Diagramas de Sankey* (Figura 2.7).

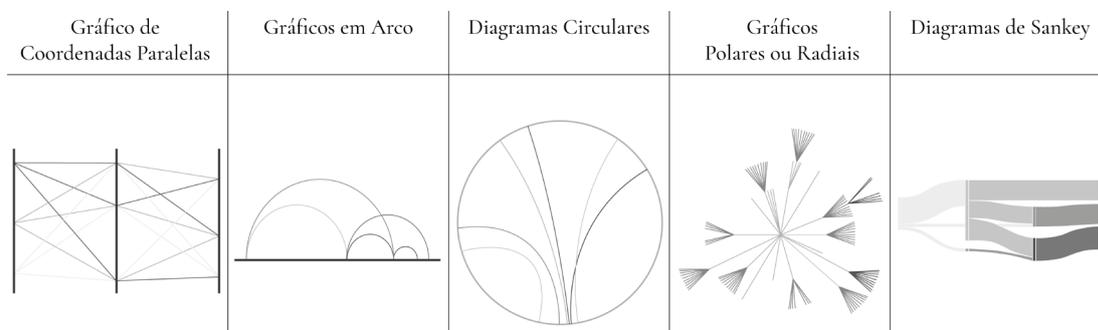


Figura 2.7 Diagramas Relacionais

Os Gráficos de *Coordenadas Paralelas* permitem estabelecer a comparação dos dados utilizando o eixo vertical, fazendo com que a informação presente numa coluna se ligue à seguinte e assim sucessivamente. Este tipo de gráfico apresenta os relacionamentos que os dados têm entre si, facilitando o estudo de padrões dos dados. “*The Real Montalbano*” (Figura 2.8) é um projeto desenvolvido pela empresa Accurat em 2012 para o suplemento cultural *La Lettura* do jornal italiano *Il Corriere della Sera*. A visualização tem como objetivo representar o desenvolvimento de uma personagem nos livros de Andrea Camilleri. A análise é feita sobre os 19 livros existentes, fazendo a representação das relações com outras personagens, hábitos, carácter e locais (reais ou fictícios) (Accurat, 2012).

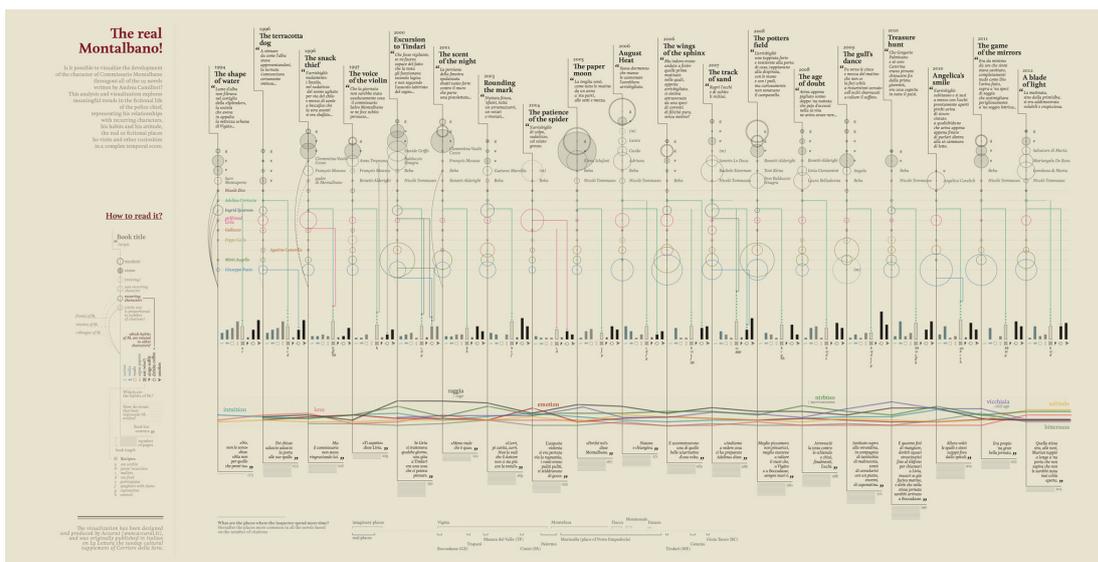


Figura 2.8 Visualização “*The Real Montalbano*” desenvolvida pela empresa Accurat

Os *Gráficos em Arco* têm todos os nós alinhados num único eixo. Cada nó relaciona-se com os restantes, através de um arco que os liga. Este tipo de gráfico funciona melhor com quantidades de informação pequenas (Meirelles, 2013). Estes gráficos inserem-se nas representações de Rede e na organização espacial linear. Desenvolvido em 2001 por Martin Wattenberg, o projeto “*Shape of Song*” (Figura 2.9 e Figura 2.10) analisa músicas e cria diagramas em arco com os dados presentes no ficheiro áudio. Os arcos são formados quando são detetadas passagens semelhantes na música. Quanto mais complexa for a composição, mais interessante se torna a visualização. Os diagramas permitem encontrar padrões nas músicas e ver como mudam dependendo do estilo musical (Wattenberg, 2001).

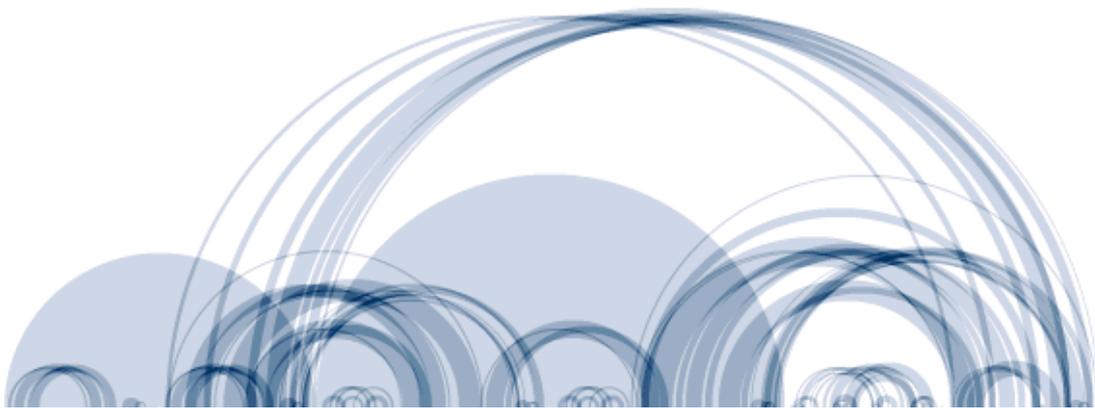


Figura 2.9 “*Shape of Song*” aplicado a um composição de Chopin.



Figura 2.10 “*Shape of Song*” aplicado à música *Like a Prayer* de Madonna.

Nos *Diagramas Circulares* os dados estão distribuídos à volta de um círculo e ligam-se através de linhas quando existe algo semelhante entre si. Estas representações pertencem à categoria de Redes e de organização espacial Circular. O projeto “*Axis of Additives*” (Figura 2.11 e Figura 2.12) utiliza um diagrama Circular para apresentar ingredientes que são adicionados a produtos embalados. O projeto foi desenvolvido em 2014 por Anna Becker, Martin Burch e Annie Gasparro para o WallStreet Journal. Para desenvolver o projeto foram selecionados 15 ingredientes que as empresas adicionam aos seus produtos. No diagrama é possível ver por grupo de alimentos que ingredientes estão presentes (Figura 2.11), e ver que produtos contém um ingrediente selecionado (Figura 2.12) (Becker, Burch & Gasparro, 2014).

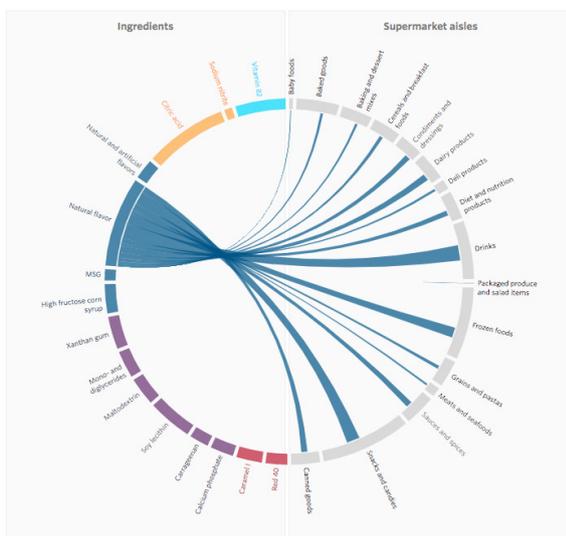


Figura 2.11 Distribuição de aditivos por corredor de supermercado.

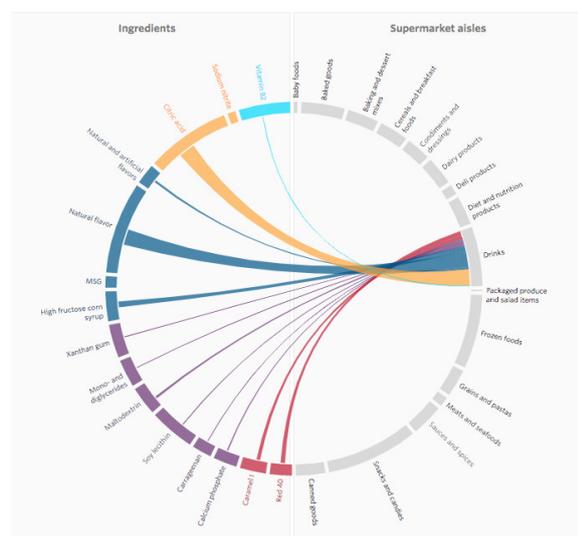
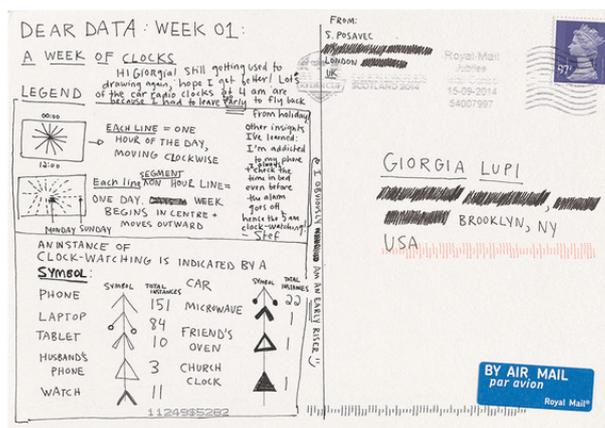


Figura 2.12 Representação da presença dos aditivos de uma seção de supermercado específica.

Os *Gráficos Polares ou Radiais* têm os nós organizados a partir do centro e do interior para o exterior. A sua posição está relacionada com o número de passos necessários até chegar a certo nó (Meirelles, 2013). No projeto “*Dear Data*” (Figura 2.13) de Stefanie Posavec e Giorgia Lupi (2015) foram recolhidos dados do seu dia a dia (segundo o tema da semana) e representados através de uma Visualização de Informação. Na primeira semana recolheram os dados de quantas vezes olhavam para o relógio. Para fazer a representação desses dados, Posavec desenvolveu um gráfico radial em que cada linha corresponde a uma hora. Essas linhas eram divididas pelos dias da semana. Os dados encontravam-se representados segundo o tipo de relógio para onde olhou, as horas que eram e o dia da semana (Posavec & Lupi, 2015).



This week Georgia and Stefanie tried gathering data in small notebooks (tedious), but soon switched to making notes on their phones (much easier). Stefanie's favourite clock to capture: a bell tolling the time in a town in Devon.

Figura 2.13 Primeira semana da Stefanie Posavec para o projecto "Dear Data".

Nos *Diagramas de Sankey* os dados são distribuídos verticalmente e ligados horizontalmente através de linhas. O tamanho da linha varia consoante a quantidade de ligações que passa de um nó para o outro. A representação feita por Minard em 1869 com o título "*Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813*" (Figura 2.14) mostra a tentativa de napoleão de invadir a Rússia (Landsteiner, 2013). A linha laranja representa os soldados que partiram de Paris. É possível ver a linha a diminuir cada vez mais até chegar ao destino. A preto está representado os soldados que regressaram, sendo possível verificar que mesmo no regresso ainda mais soldados perderam a vida (Gonzalez, 2012).

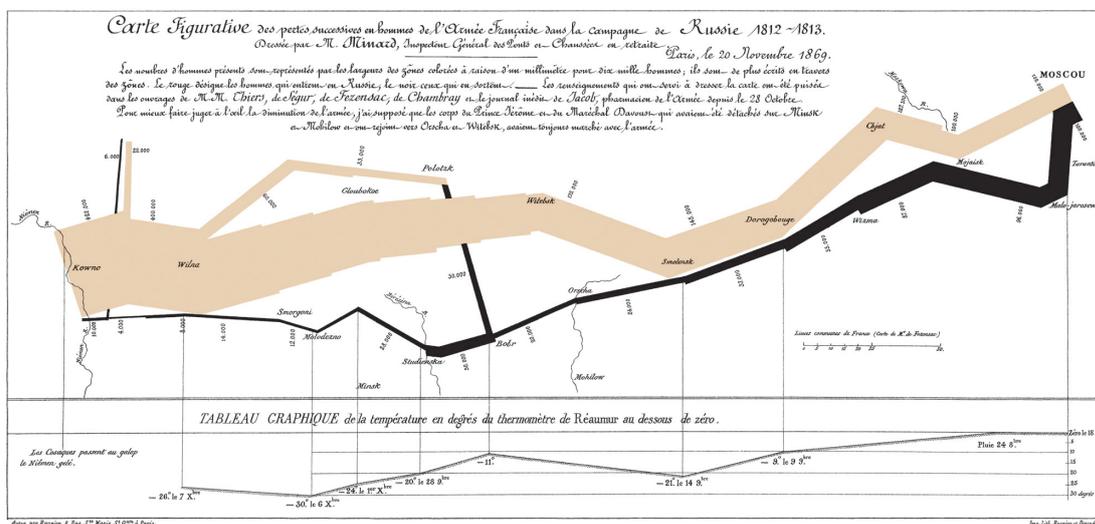


Figura 2.14 “Carte Figurative des pertes successives en hommes de l’Armée Française dans la campagne de Russie 1812-1813” desenvolvido por Charles Minard em 1869.

As representações temporais mostram o comportamento dos dados ao longo de uma linha de tempo. É possível olhar para as representações temporais de duas formas distintas. Numa os dados movem-se ao longo do tempo e na outra é o tempo que se altera consoante os dados (Meirelles, 2013). O tempo em si não tem forma de ser medido nem tem direção (Gould, 1998). Os modelos de dados temporais são influenciados pela perceção de tempo de quem está a trabalhar os dados. As representações podem ser feitas de forma circular ou linear (Meirelles, 2013), sendo representados através de *Timelines*; *Histogramas* e *Gráficos de Área Polar*. (Figura 2.15)

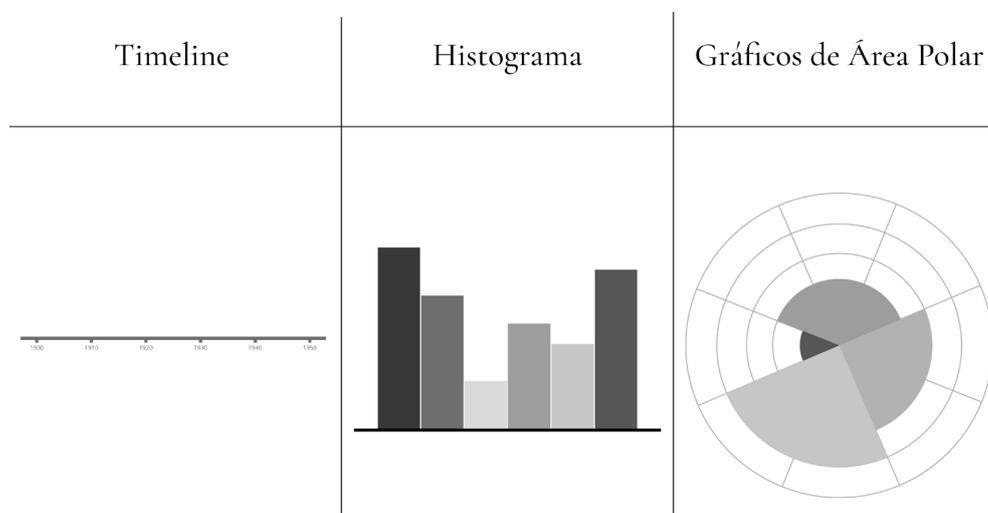


Figura 2.15 Diagramas de representação temporal

As *Timelines* organizam os dados de forma cronológica, podendo estes seguir diversos tipos de escalas temporais. Na maioria das vezes este tipo de gráficos são usados para relatar um conjunto de acontecimentos. Para além disso, tendem a ser acompanhados com outros diagramas para se transmitir mais informação. No projeto “*The Swedish Emigrants*” (Figura 2.16) de Madelene Wikskär foram representados os dados da emigração em massa entre 1850 e 1925. Esta timeline foi desenvolvida no eixo vertical e estão representados a branco os homens, a lilás as mulheres e a vermelho as vidas perdidas na travessia (Wikskär, 2015).

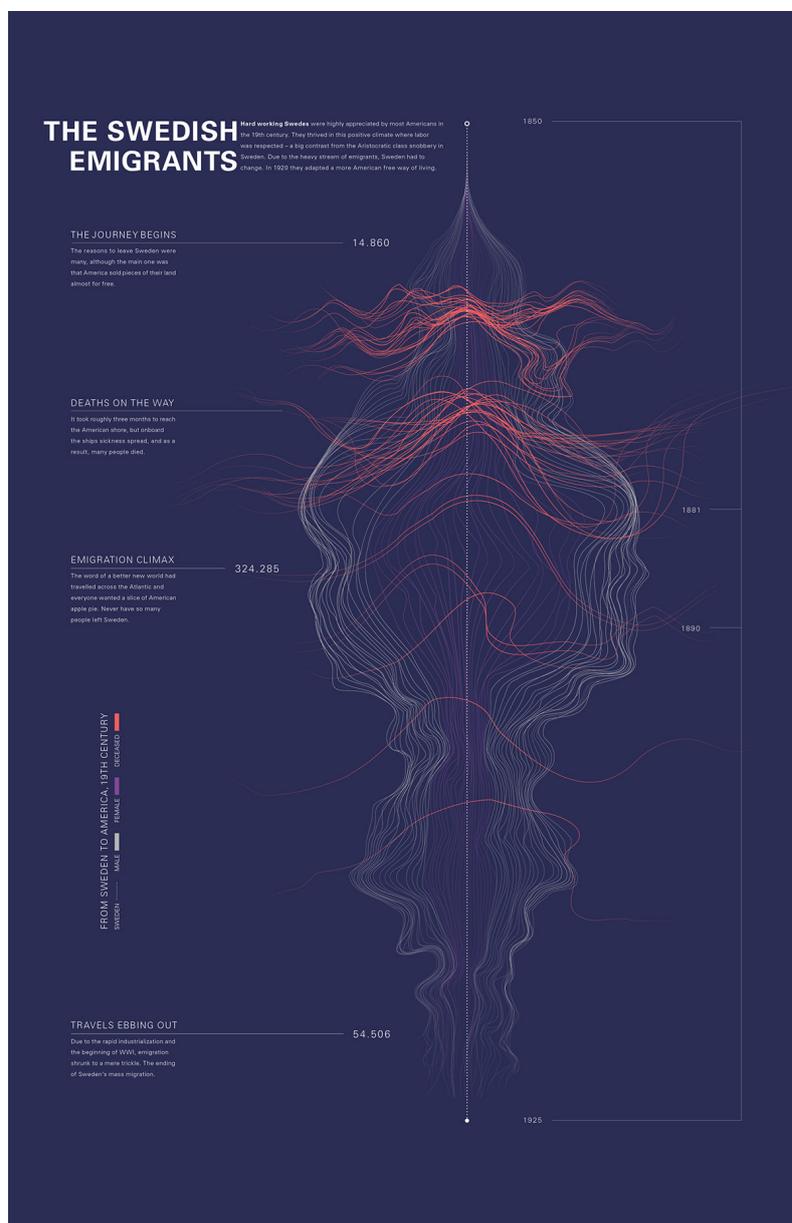


Figura 2.15 Visualização “*The Swedish Emigrants*”.

Os *Histogramas*, ou gráficos de barras, dão uso tanto ao eixo vertical como ao horizontal. Os dados são distribuídos por variáveis que são posicionadas ao longo de eixos temporais. Este tipo de gráficos facilita o estudo e comparação de dados. Em 1786, William Playfair apresentou um dos primeiros histogramas intitulado “*Exports and Imports of Scotland to and from different parts for one Year from Christmas 1780 to Christmas 1781*” (Figura 2.17). Neste gráfico representam-se as transações feitas com diversos países ao longo de um ano (AnyChart, 2015).

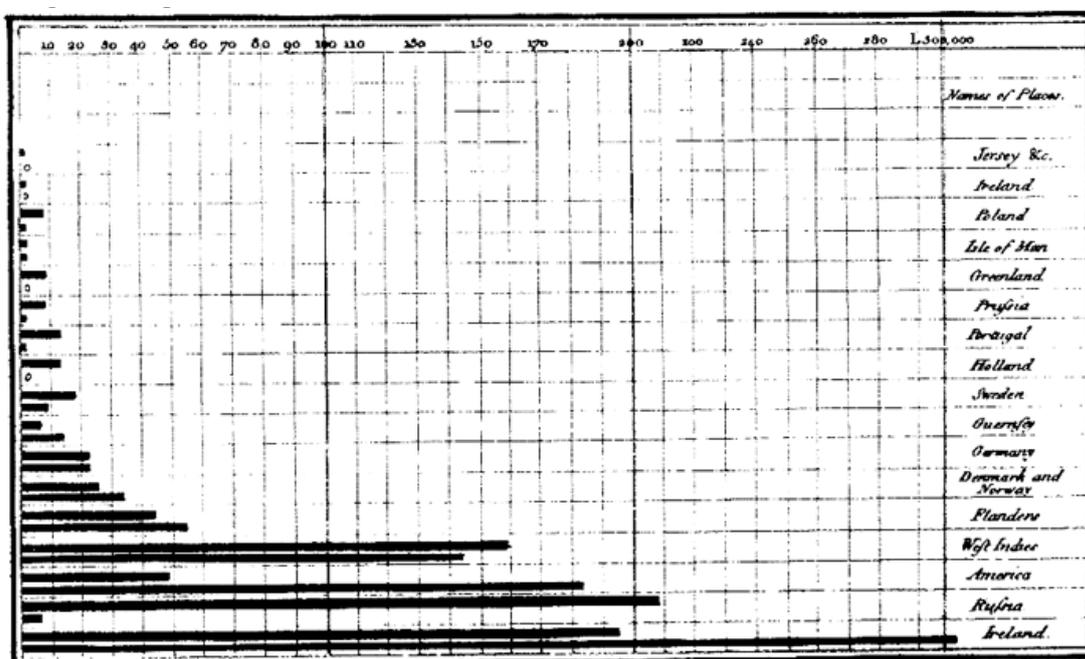


Figura 2.17 Gráfico de Barras de William Playfair, 1786.

Os *Gráficos de Área Polar* (também conhecidos por *Rose Chart*) são gráficos circulares em que é possível representar um período de tempo (não muito longo), dividindo o círculo no número de partes pretendida. A distribuição da informação é feita do centro para o exterior do gráfico, ou seja, o valor de cada variável é medido através da altura. Em 1858, Florence Nightingale desenvolveu um conjunto de diagramas “*Diagram of the Causes of Mortality in the Army in the East*” (Figura 2.18) para evidenciar que as mortes dos soldados britânicos na guerra da Crimeia eram provocadas, na sua maioria, pela pouca qualidade de vida e não pela guerra (ims25, 2008). Estes diagramas deram origem aos *Rose Charts*. Nos diagramas existem doze secções diferentes que representam cada mês do ano. Os dados referentes à mortalidade são representados através da

área. A cor azul representa o número de mortes causadas por doença, vermelho as mortes causadas por ferimentos de guerra e preto outras causas (Small, 1998). Com esta representação, Nightingale conseguiu comprovar a necessidade de criar serviços de qualidade para se diminuir o número de mortes.

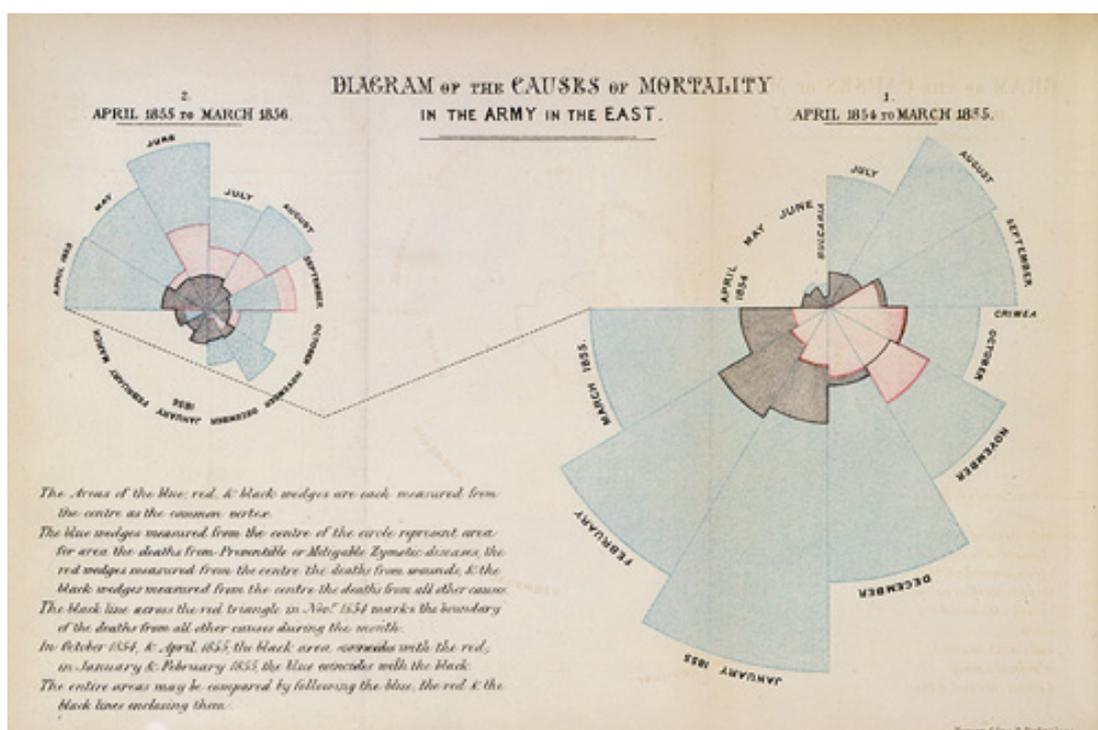


Figura 2.18 Diagrama das Causas da Mortalidade de Florence Nightingale.

Os mapas são dos modelos de representação de informação mais antigos. Os egípcios já registavam locais 200 anos AC, através de coordenadas semelhantes a altitude e longitude (Friendly, 2006). Com o avançar do tempo os mapas tornaram-se uma ferramenta de representação para as mais diversas áreas e não apenas para a georeferenciação. Os mapas temáticos foram desenvolvidos a partir do século XVII e têm sido utilizados e adaptados ao longo do tempo. Utilizar mapas para representar informação deve focar-se em três setores: a projeção, escala e simbologia (Meirelles, 2013). A representação georreferencial está dividida nas seguintes categorias: *Choropleth Maps*; *Dot Maps*; *Proportional/Graduated Symbol Maps*; *Flow/Network Maps*; *Area/Distance Cartograms*; e *Schematic Typographic Maps* (Figura 2.19).

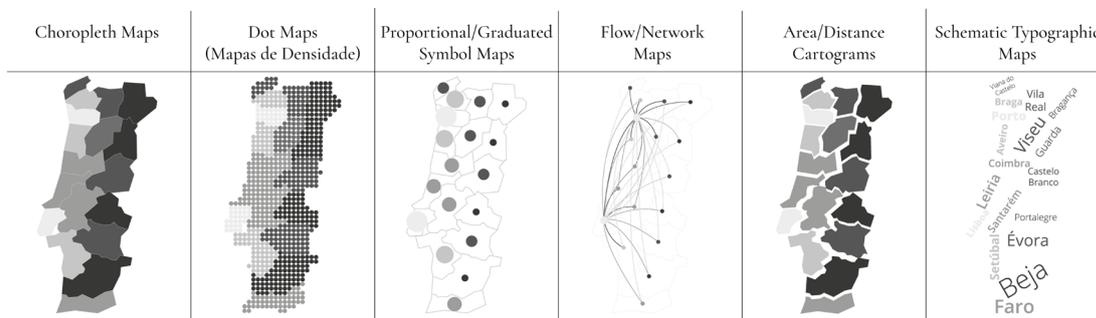


Figura 2.19 Tipos de mapas.

Os *Choropleth Maps* representam dados quantitativos numa determinada área do mapa. Esta representação é feita através de cor, que se altera de acordo com o valor dos dados. Geralmente é usado um gradiente de cor para mostrar a evolução dos valores. Os mapas temáticos têm sido utilizados ao longo dos séculos para apresentar dados de diversos acontecimentos (Meirelles, 2013). O projeto “*Transparency: Are the Richest Americans Also the Best Educated? America’s Richest Counties and Best Educated Counties*” (Figura 2.20) foi desenvolvido para a revista GOOD. Neste mapa são representados três tipos diferentes de dados: a cor de rosa, os dados relativos aos alunos que terminam o ensino secundário; a amarelo, os alunos que terminam o ensino superior; e a azul, o rendimento médio dos americanos. O mapa é a combinação dos três dados para cada cidade dos Estados Unidos. As cidades com resultados mais claros são as que têm valores mais baixos, ou seja, menos pessoas formadas e menos rendimentos. As regiões mais escuras têm níveis mais altos de educação e rendimentos (Hubacek, 2011).

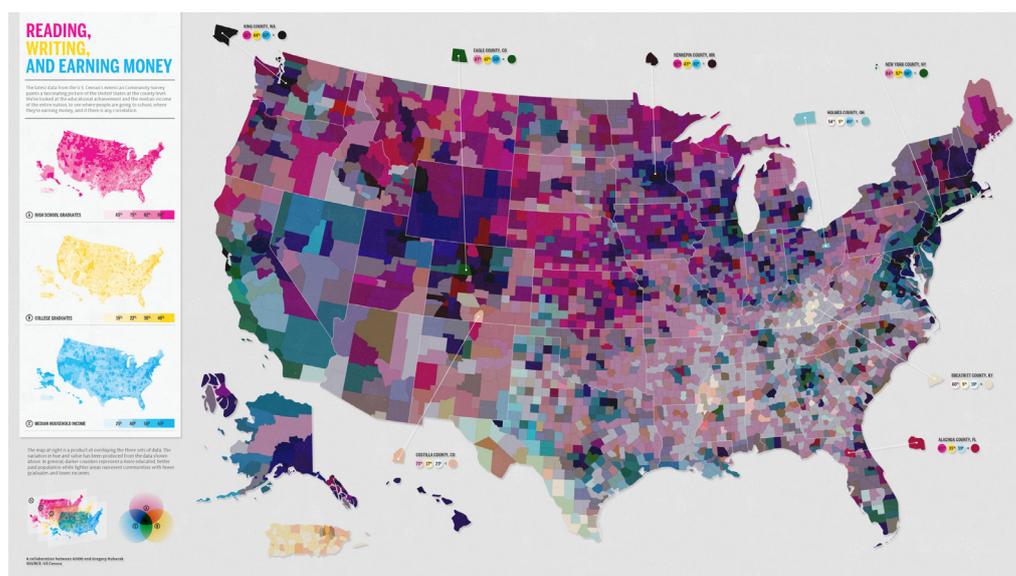


Figura 2.20 “*Transparency: Are the Richest Americans Also the Best Educated? America’s Richest Counties and Best Educated Counties*” para a revista GOOD, 2011.

Os *Dot Maps* (ou *Mapas de Densidade*) utilizam, como o nome indica, pontos para fazer a distribuição dos dados no espaço. Através deste tipo de mapas, é possível perceber padrões de distribuição espacial. Os pontos devem ter o mesmo tamanho e o mesmo significado para cada representação. O projeto “*Mapping the World, one Tweet at a Time*” (Figura 2.21) desenvolvido por Alan Mislove mostra tweets partilhados com localização. A recolha foi feita ao longo de 30 meses (entre janeiro de 2011 e abril de 2013) tendo sido guardada a localização de mais de 200 milhões de *tweets*. Todo o mapa é construído a partir de pontos que são a localização de onde foram feitos os *tweets* (Mislove, n.d.).



Figura 2.21 “*Mapping the World, one Tweet at a Time*”.

Proportional/Graduated Symbol Maps são representações que recorrem à variação do tamanho de símbolos para mostrar diferenças na representação. Pode haver problemas de perceção com o tamanho dos símbolos, estando sujeitos a uma interpretação errada, dependendo das variáveis que o rodeiam (Gruver, & Dutton, n.d.). Charles Minard desenvolveu em 1858 um mapa para representar os valores da carne fornecida a Paris de cada cidade (Figura 2.22). As cidades que fazem o fornecimento estão marcadas a amarelo. Nos círculos, o tamanho do círculo está relacionado com a quantidade total de carne fornecida, a preto é representada a percentagem de carne bovina, a vermelho vitela e a verde carneiro (Meirelles, 2013).

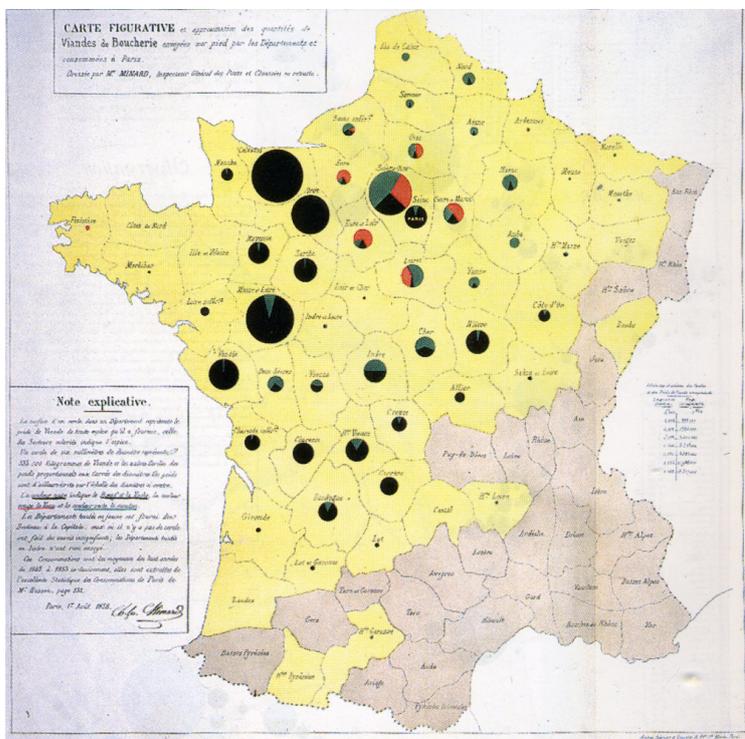


Figura 2.22 “Carte figurative et approximative des quantités de viandes de boucherie envoyées sur pied par les départements et consommés à Paris” de Charles Minard (Palsky, 1996).

Flow/Network Maps representa o movimento dos dados no espaço, o caminho em si pode não representar locais exatos, sendo o ponto de partida e o ponto de chegada que de facto interessam para efetuar a análise geral da informação. Minard criou também um dos primeiros mapas de fluxo, “Carte figurative et approximative des quantités de coton brut importées en Europe en 1858, en 1864 et en 1865” (Figura 2.23). Nestes mapas é possível ver a origem da importação de algodão para a Europa através de linhas de fluxo. Cada cor representa um país e a direção do fluxo é de fora para dentro da Europa (Meirelles,2013).

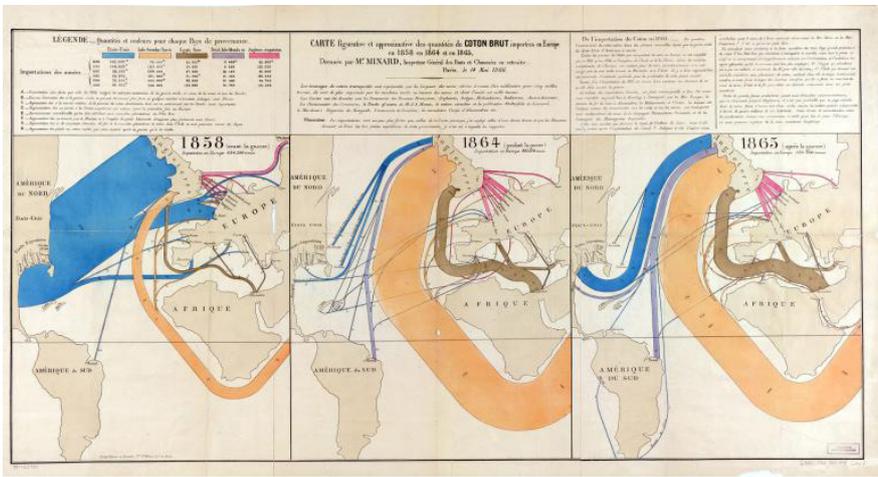


Figura 2.23 “Carte figurative et approximative des quantités de coton brut importées en Europe en 1858, en 1864 et en 1865” desenvolvido por Charles Minard (Mason, 2017).

Os *Area/Distance Cartograms*, ao contrário das representações comuns, distorcem os mapas, tendo por base os valores dos dados que estão a representar. Benjamin Hennig (2016) desenvolveu um cartograma com distorção sobre os dados do referendo do Reino Unido realizado em 2016 (Figura 2.24). A representação dos dados é feita através da distorção das cidades de acordo com a população elegível para votar. As cidades que tiveram mais votos a favor da saída do Reino Unido da União Europeia têm a cor vermelha e os votos contra cor azul (Hennig, & Dorling, 2016).

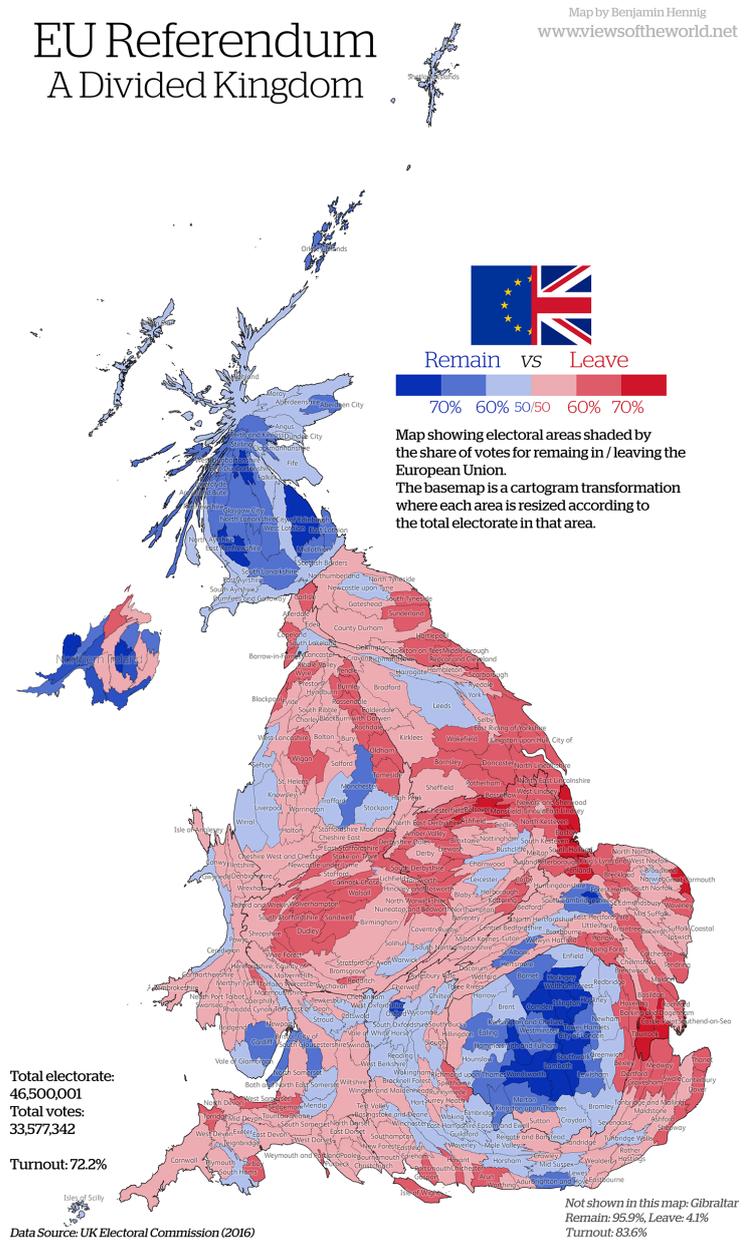


Figura 2.24 “EU Referendum, A Divided Kingdom”.

Os *Schematic Typographic Maps* são mapas que representam um espaço ou uma área, utilizando apenas a tipografia. Este tipo de mapa acaba por ter pouca utilização para a representação de dados, tendo sido desenvolvidos para fins mais artísticos.

Existe ainda a possibilidade de representar a evolução de um espaço num determinado período de tempo. Estes dados podem ser sobre os mais diferentes temas, desde que exista a possibilidade de visualizar uma evolução, positiva ou negativa, ao longo do tempo. Este tipo de mapas são importantes para compreender padrões (Meirelles, 2013). Estas representações tornam-se complexas quando é necessário representar a evolução do tempo numa imagem estática. É frequente o uso de sequências de mapas para ser possível observar os dados. Os mapas de Fluxo são umas das soluções escolhidas para estas representações, por ser o mais próximo de representação de movimento numa imagem estática. “*Phantom Terrains*” (Figura 2.25) é um trabalho desenvolvido por Frank Swain e Daniel Jones. O seu objetivo é transformar os sinais das redes wireless em som. Com esta transformação o utilizador consegue perceber mais rapidamente quando ocorrem mudanças nos dados. Na representação visual, as cores ajudam a distinguir a qualidade das redes e através de padrões visuais o tipo de segurança da rede (Swain, & Jones, n.d.).

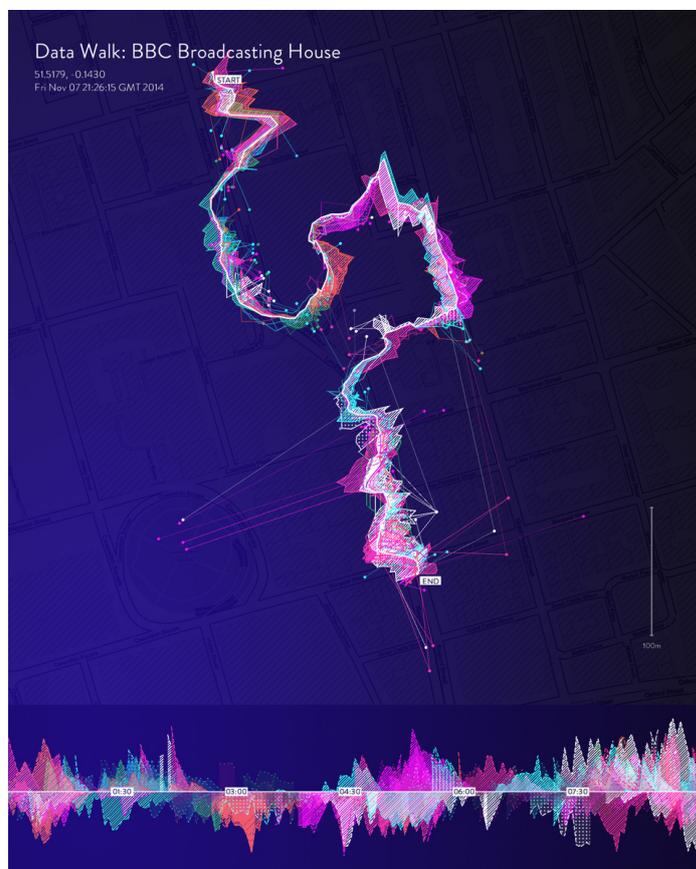


Figura 2.25 Representação visual do “*Phantom Terrains*”.

2.2.2 VISUALIZAÇÃO DE TEXTO

Nos dias de hoje o acesso a documentos de texto é bastante simples, mas as técnicas de representação e análise de texto ainda estão em desenvolvimento (Meirelles, 2013). Qualquer documento é composto por relacionamentos complexos entre as palavras nele existentes (Lima, 2011). As visualizações de texto podem ainda ser combinadas com outro tipo de representações, tornando-se mais informativas para o utilizador (Meirelles, 2013).

Existem diversas técnicas de visualização que podem ser utilizadas para representar os conteúdos dos documentos. A pré análise destes textos pode incidir sobre semelhanças, repetições, relações entre textos do mesmo documento ou de documentos diferentes do mesmo autor e documentos de diferentes autores que têm por inspiração um só autor. O principal objetivo da visualização de texto é o de simplificar a compreensão de texto para os utilizadores.

Podemos fazer uma análise quantitativa num texto, calculando quantas vezes uma letra ou uma palavra aparece no documento. No projeto “Text Arc: Alice in Wonderland” (Figura 2.26) de W. Bradford Paley é feita uma análise ao livro de Lewis Carroll. O texto do livro é apresentado, por completo, à volta de uma circunferência a começar nas 12:00 horas e a seguir no sentido dos ponteiros do relógio. Dentro da circunferência as palavras são posicionadas o mais perto da parte do livro em que aparecem, palavras que estão a aparecer constantemente ficam mais próximas do centro do objeto, o seu tamanho e brilho também é influenciado pela frequência com que aparecem no texto. Ao seleccionar uma palavra é ainda possível ver mais pormenorizadamente onde esta aparece no texto (Paley, 2002).

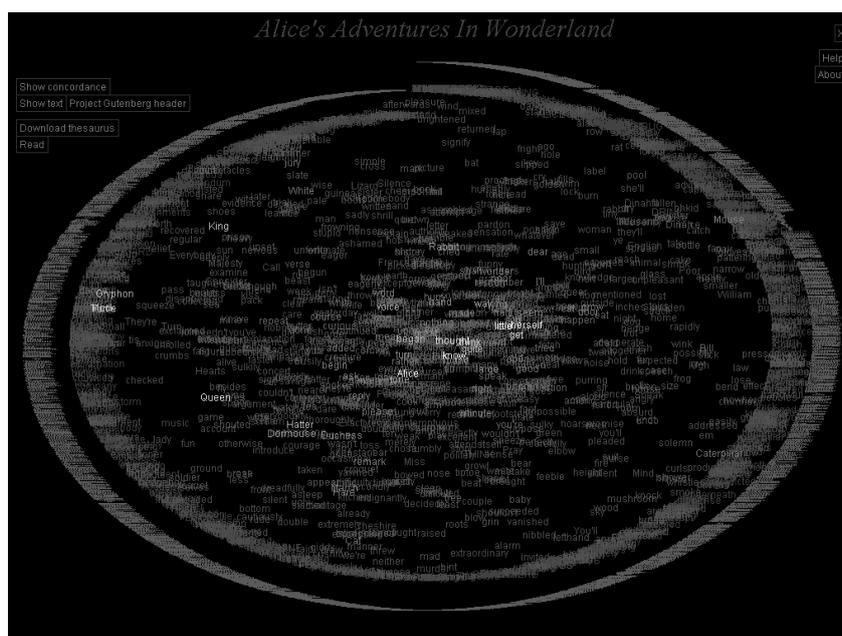


Figura 2.26 Imagem da visualização de “Text Arc: Alice’s Adventures in Wonderland”.

Através da visualização de texto é possível analisar as alterações entre edições do mesmo livro, como fez Stefanie Posavec e Greg McInerney no projeto *“The evolution of the Origin of Species”* (Figura 2.27), em que são feitas representações visuais para cada edição do livro de Charles Darwin. Cada artefacto representa as alterações feitas entre cada edição. Em cada representação são exibidos os capítulos do livro, que são divididos em subcapítulos, em que vão ser representados os parágrafos e para cada parágrafo as frases nele presentes. A laranja são representadas as frases que estão alteradas na edição seguinte e a azul as partes do livro que se mantém (Posavec, & McInerney, n.d.).

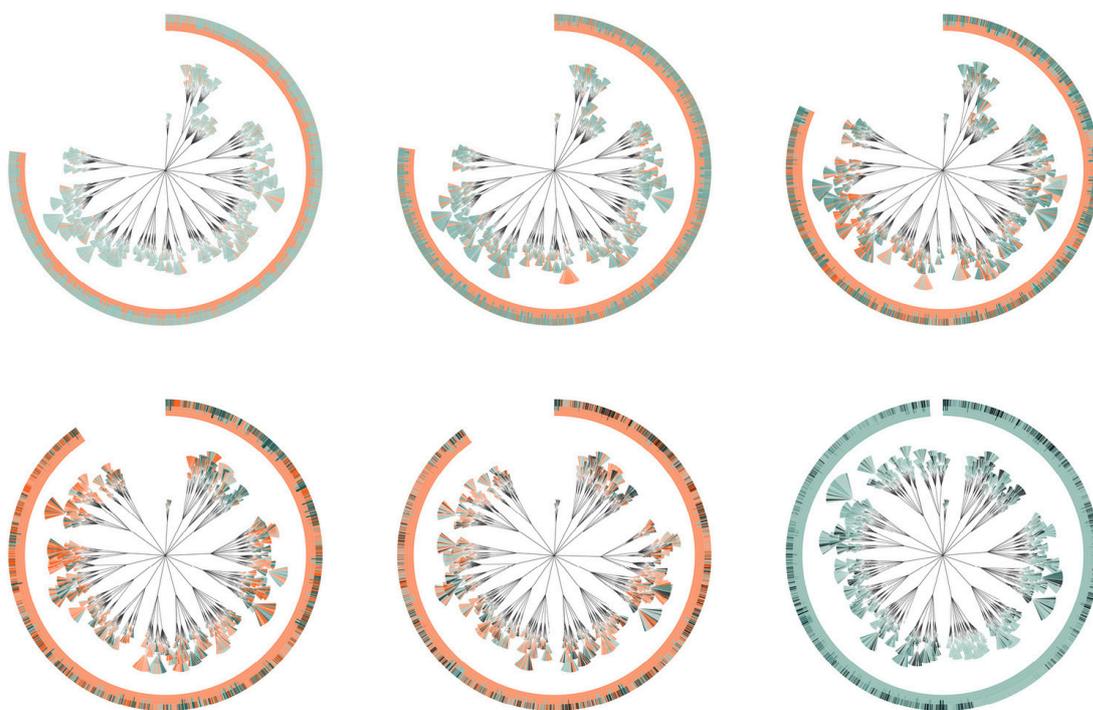


Figura 2.27 *“The Evolution of the Origin of Species”* de Stefanie Posavec e Greg McInerney.

Podendo também ter uma vertente de linguagem visual (Meirelles, 2013), a visualização de texto pode apenas transmitir uma mensagem visual dos dados textuais. Nestas representações, os dados são analisados para serem apresentados de forma abstrata, podendo ser submetidos a análises estatísticas ou geográficas para o desenvolvimento de padrões visuais (Meirelles, 2013). O projeto visual de *“Poetry On The Road”* (Figura 2.28) por Boris Müller, em 2006, fez uma representação abstrata de poesia. Para o desenvolvimento da imagem do evento, a cada letra do alfabeto foi atribuído um número, e para descobrirem o valor de cada palavra somam

os números de cada letra dessa palavra. Os poemas mais curtos têm uma circunferência de diâmetro mais pequeno e os poemas mais longos diâmetro maior (Schaffors, Müller, & Pfeffer, 2006).

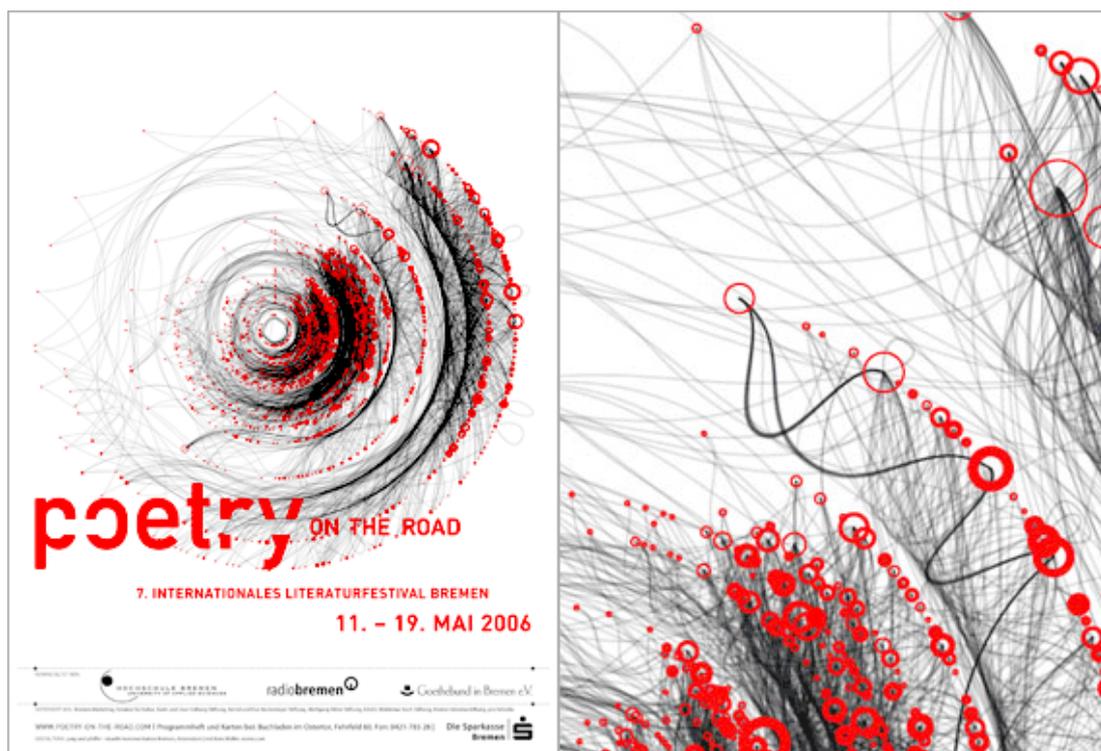


Figura 2.28 Cartaz e pormenor do cartaz do evento “Poetry On The Road” 2006.

De acordo com Ward, Grinstein e Keim (2015), para a realização de uma análise linguística existem três níveis. No primeiro nível faz-se uma análise Lexical dos textos. Cada palavra presente no documento é transformada num token. Os *tokens* vão ser submetidos a uma análise seguindo um conjunto de regras. O objetivo é gerar novas sequências de *tokens* (palavras). O segundo nível é o Sintático. Neste nível são analisadas as sequências resultantes da análise Lexical. As sequências vão ter registos atribuídos sobre a informação que contém (por exemplo, podem ser palavras no plural ou números). A análise Semântica é a terceira a ser realizada. Os objetos analisados são os registos atribuídos às sequências no nível anterior. O objetivo neste nível é fazer a interpretação analítica global do documento em análise.

A Visualização de Informação aplicada à literatura facilita a compreensão de textos, permitindo ao utilizador uma melhor compreensão das alterações entre edições ou dos relacionamentos existentes em livros do mesmo autor. Rody Zakovich desenvolveu um conjunto de representações sobre as obras de Shakespeare (Figura 2.29). Cada diagrama circular representa uma obra de Shakespeare. Os dados recolhidos são sobre as personagens que têm falas nos livros. Os círculos representam a quantidade de falas numa determinada cena e as linhas mostram a distância da presença de cada personagem entre cenas (Zakovich, n.d.).

Shakespeare: A Data Visualization

How to read: The entire circle \bigcirc represents every character who had a speaking part in each scene throughout the play (clockwise from 12 O'clock). This could consist of a single line, or series of lines. Each line \diagup around the circle is the number of words spoken by a character during a particular scene. The inner filled circles \bullet represent the number of lines the character had in the scene, larger circles equals more lines. The Chords \frown are determined by the distance between a character's presence in a scene to the next. Marks are ordered by total words said in the scene in ascending order.

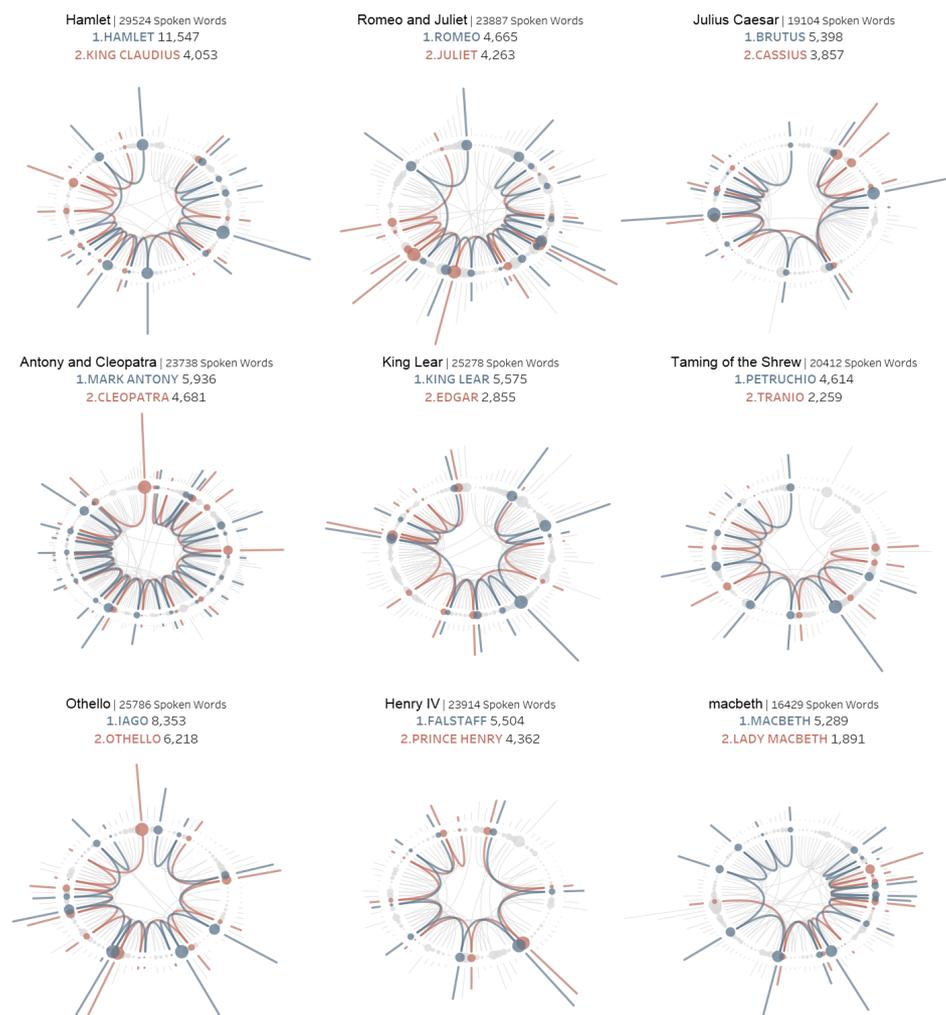


Figura 2.29 Representações das obras de Shakespeare.

2.3 CONCLUSÕES DO ESTADO DA ARTE

O desenvolvimento deste capítulo permitiu compreender as diversas formas de representações visuais que auxiliam a compreensão dos diferentes tipos de dados. Concluímos que a categorização e a escolha correta das variáveis visuais para a representação dos dados leva à construção de artefactos mais fidedignos e mais facilmente compreendidos pelos observadores. A informação tem que ser transmitida de forma compreensível e acessível sendo que para isso, é preciso analisar e encontrar a melhor representação dos dados que estão a ser trabalhados. Para além disso, é importante saber como provocar o interesse no observador para a análise os artefactos, dando a possibilidade de interação e manipulação dos mesmos.

A exploração e apresentação de vários tipos de visualização de informação é importante para aprofundar o conhecimento dos diversos tipos de visualizações existentes, com exemplos aplicados ao longo dos séculos. A procura de trabalhos desenvolvidos com dados textuais abriu ainda caminho para o desenvolvimento de novas ideias de exploração dos dados da presente dissertação através.

3 A “CANÇÃO DO EXÍLIO”

Neste capítulo, é apresentado o poema escrito por Gonçalves Dias, é abordada a evolução do poema ao longo dos anos e são realçados alguns textos que relatam momentos importantes da história do Brasil. Por fim, é abordado o trabalho desenvolvido por Joshua Enslen que culminou na exposição *Bird Watching: Visualizações da Influência de “Canção do Exílio”*.

3.1 O POEMA

O poema “*Canção do Exílio*” foi escrito por Antônio Gonçalves Dias, em 1843, enquanto era estudante de Direito na Universidade de Coimbra. Vis- to como sendo uma declaração de amor à pátria, desde a sua publicação em 1846, tornou-se um dos poemas mais populares no Brasil, fazendo com que seja um dos textos mais imitados (Toda Matéria, 2017). O poema é um dos mais importantes do início do Romantismo brasileiro.

Canção do Exílio
Antônio Gonçalves Dias, 1843

*Minha terra tem palmeiras,
Onde canta o Sabiá;
As aves, que aqui gorjeiã,
Não gorjeiã como lá.*

*Nosso céu tem mais estrelas,
Nossas varzeas tem mais flores,
Nossos bosques têm mais vida,
Nossa vida mais amores.*

*Em scismar – sósinho – á noite –
Mais prazer encontro eu lá;
Minha terra tem palmeiras,
Onde canta o Sabiá.*

*Minha terra tem primores,
Que taes não encontro eu cá;
Em scismar – sósinho – á noite –
Mais prazer encontro eu lá;
Minha terra tem palmeiras,
Onde canta o Sabiá.*

*Não permitta Deos que eu morra,
Sem que eu volte para lá;
Sem que desfructe os primores
Que não encontro por cá;
Sem qu’inda aviste as palmeiras,
Onde canta o Sabiá.
(Dias, 1846)*

O poema tem servido como inspiração para o desenvolvimento de novos poemas ou textos, que glorificam e criticam o Brasil. As palavras escritas por Gonçalves Dias, têm servido como meio de expressão dos sentimento vividos em relação ao Brasil, possibilitando a descrição do país ao longo dos anos através dos olhos da população. É ainda possível encontrar a presença do poema de Gonçalves Dias no hino brasileiro (Morfeus, n.d.).

Alguns dos autores de variações mais conhecidos são: Carlos Drummond de Andrade, Jô Soares, Oswald de Andrade, Tom Jobim e Chico Buarque (Dias, Bueno, & Bandeira, 1998).

3.2 A EVOLUÇÃO DO POEMA

Ao longo dos 174 anos de existência, a “*Canção do Exílio*” tem servido como inspiração para as mais diversas adaptações. Estas adaptações muitas vezes acabam por refletir o sentimento do povo brasileiro para com a situação política e econômica do país, podendo estas ser positivas ou negativas. Inicialmente, as novas versões eram feitas e publicadas em livros ou jornais. Hoje em dia são muitas vezes difundidas na web, como músicas ou textos em prosa. Enquanto que no poema original de Gonçalves Dias era feita uma declaração de amor ao Brasil, e o autor tentava mostrar que mais nenhum lugar no mundo era comparável ao seu país, nas paródias mais recentes o tom que impera é a crítica. Se antes tudo servia para mostrar o lado positivo do Brasil, hoje em dia todas as polêmicas servem para criticar o país. Nos 20 anos após a publicação do poema, é possível observar nos dados fornecidos por Joshua Enslin, que a maioria das variações eram ainda positivas sobre o Brasil. Existe um período em que as variações não revelam nada de positivo mas também não apresentam uma imagem negativa sobre o Brasil (entre 1868 e 1873). Após esse período neutro, as paródias negativas e de protesto foram aumentando. Esta mudança é interessante de se ver pois é possível relacioná-la com períodos históricos que marcaram o Brasil. Quando o poema foi publicado ainda existia a euforia da independência recente do Brasil como colônia portuguesa. Este sentimento foi desaparecendo e recentemente as crises políticas e econômicas incentivaram o uso da palavra como forma de protesto.

No início do projeto foi feita uma breve pesquisa em que foram encontradas 18 variações da “*Canção do Exílio*”. O período de publicação destas variações é de 1855 a 2000. Em 1855, Casimiro de Abreu escreveu a variação “*Eu Nasci Além dos Mares*”. Neste poema de Casimiro de Abreu também é feita uma declaração de amor ao Brasil como no poema de Gonçalves Dias. Nesta variação é possível encontrar referências à “*Canção do Exílio*” de Gonçalves Dias através do uso de palavras como: sabiá, palmeiras, céu e exílio.

Eu Nasci Além Dos Mares
Casimiro de Abreu, 1855

*Os meus lares,
Meus amores ficam lá!
— Onde canta nos retiros
Seus suspiros,
Suspiros o sabiá!*

*Oh! Que céu, que terra aquela,
Rica e bela
Como o céu de claro anil!
Que seiva, que luz, que galas,
Não exalas,
Não exalas, meu Brasil!*

*Oh! Que saudades tamanhas
Das montanhas,
Daqueles campos natais!
Que se mira,
Que se mira nos cristais!*

*Não amo a terra do exílio
Sou bom filho,
Quero a pátria, o meu país,
Quero a terra das mangueiras
E as palmeiras
E as palmeiras tão gentis!
Como a ave dos palmares
Pelos ares
Fugindo do caçador;
Eu vivo longe do ninho;
Sem carinho
Sem carinho e sem amor!*

*Debalde eu olho e procuro...
Tudo escuro
Só vejo em roda de mim!
Falta a luz do lar paterno
Doce e terno,
Doce e terno para mim.*

*Distante do solo amado
— Desterrado —
a vida não é feliz.
Nessa eterna primavera*

*Quem me dera,
Quem me dera o meu país!*
(Helena & Telles, 2011)

Quando o Brasil entrou no período Modernista, em 1925, Oswald de Andrade escreveu o poema “*Canto do Regresso à Pátria*”. Tendo referências claras ao poema de Gonçalves Dias (e. g.: “*Não permita Deus que eu morra*”) esta variação não faz uma descrição idílica do Brasil, mas sim uma paródia do poema original.

Canto Do Regresso À Pátria
Oswald de Andrade, 1925

*Minha terra tem palmares
onde gorjeia o mar
Os passarinhos daqui
Não cantam como os de lá*

*Minha terra tem mais rosas
E quase que mais amores
Minha terra tem mais ouro
Minha terra tem mais terra*

*Ouro terra amor e rosas
Eu quero tudo de lá
Não permita Deus que eu morra
Sem que volte para lá*

*Não permita Deus que eu morra
Sem que volte pra São Paulo
Sem que veja a Rua 15
E o progresso de São Paulo*
(Ao Pé da Letra, 2011)

O Brasil atravessou um período ditatorial, entre 1964 e 1985, que inspirou poetas a usarem a escrita como forma de expressão. Torquato Neto escreveu em 1968 o poema “*Marginália II*”, no qual faz referência à “*Canção do Exílio*” através do uso da frase “*minha terra tem palmeiras*”. No poema de Torquato Neto é possível ler a desilusão, o sentimento de culpa e o medo da situação que o Brasil atravessava.

Marginália II
Torquato Neto, 1968

*eu, brasileiro, confesso
minha culpa meu pecado
meu sonho desesperado*

*meu bem guardado segredo
minha aflição
eu, brasileiro, confesso
minha culpa meu degredo
pão seco de cada dia
tropical melancolia
negra solidão:
aqui é o fim do mundo
aqui é o fim do mundo
ou lá
aqui o terceiro mundo
pede a bênção e vai dormir
entre cascatas palmeiras
araçás e bananeiras
ao canto da juriti
aqui meu pânico e glória
aqui meu laço e cadeia
conheço bem minha história
começa na lua cheia
e termina antes do fim
aqui é o fim do mundo
aqui é o fim do mundo
ou lá
minha terra tem palmeiras
onde sopra o vento forte
da fome do medo e muito
principalmente
da morte
o-lelê, lalá
a bomba explode lá fora
e agora, o que vou temer?
yes: nós temos banana
até pra dar,
e vender
aqui é o fim do mundo
aqui é o fim do mundo
ou lá
(Escritas.org, n.d)*

A partir dos anos 90, a instabilidade econômica do Brasil foi crescendo levando Jô Soares a escrever “*Canção do Exílio Às Avestas*” inspirada na “*Canção do Exílio*” de Gonçalves Dias, sendo a expressão “*Não permita Deus*” comum aos dois poemas. Na sua paródia, Jô Soares critica os escândalos políticos que envolveram o governo da época.

Canção do Exílio às Avelãs
Jô Soares

*Minha Dinda tem cascatas
Onde canta o curió
Não permita Deus que eu tenha
De voltar pra Maceió.
Minha Dinda tem coqueiros
Da Ilha de Marajó
As aves, aqui, gorjeiam
Nào fazem cocoricó.*

*§O meu céu tem mais estrelas
Minha várzea tem mais cores.
Este bosque reduzido
deve ter custado horrores.
E depois de tanta planta,
Orquídea, fruta e cipó,
Não permita Deus que eu tenha
De voltar pra Maceió.*

*Minha Dinda tem piscina,
Heliporto e tem jardim
feito pela Brasil's Garden:
Não foram pagos por mim.
Em cismar sozinho à noite
sem gravata e paletó
Olho aquelas cachoeiras
Onde canta o curió.*

*No meio daquelas plantas
Eu jamais me sinto só.
Não permita Deus que eu tenha
De voltar pra Maceió.
Pois no meu jardim tem lagos
Onde canta o curió
E as aves que lá gorjeiam
São tão pobres que dão dó.*

*Minha Dinda tem primores
De floresta tropical.
Tudo ali foi transplantado,
Nem parece natural.
Olho a jabuticabeira
dos tempos da minha avó.*

*Não permita Deus que eu tenha
De voltar pra Maceió.*

*Até os lagos das carpas
São de água mineral.
Da janela do meu quarto
Redescubro o Pantanal.
Também adoro as palmeiras
Onde canta o curió.
Nào permita Deus que eu tenha
De voltar pra Maceió.*

*Finalmente, aqui na Dinda,
Sou tartado a pão-de-ló.
Só faltava envolver tudo
Numa nuvem de ouro em pó.
E depoies de ser cuidado
Pelo PC, com xodó,
Nãõ permita Deus que eu tenha
De acabar no xilindró.
(Recantos Poéticos, 2012)*

As variações do poema apresentadas, permitem ter uma ideia de como o poema influenciou outros poetas a expressarem os momentos históricos do Brasil. Cada um com o seu próprio estilo de expressão e interpretação do poema original de Gonçalves Dias, mostram o Brasil pelas suas palavras.

3.3 BIRD WATCHING: VISUALIZAÇÕES DA INFLUÊNCIA DE “CANÇÃO DO EXÍLIO”

A exposição *Bird Watching: Visualizações da Influência de “Canção do Exílio”* desenvolvida por Joshua Enslin e Alaina Enslin, mostra diversas visualizações das variações do poema de Gonçalves Dias. O material presente na exposição foi recolhido e analisado por Joshua Enslin no decorrer do seu projeto de pós-doutoramento (Enslin, & Enslin, 2016).

A exposição tem várias peças que pretendem retratar a influência da “*Canção do Exílio*”. Uma das principais peças, desenvolvida por Alaina Enslin, é o retrato de Antônio Gonçalves Dias construído a partir de fotografias do Instagram, de textos de diversas origens e da representação dos elementos principais do poema: palmeiras e o sabiá. Ainda neste retrato representativo do poeta está representado um gráfico onde é apresentada a relação dos textos com o poema original (Figura 3.1) (Enslin & Enslin, 2016).

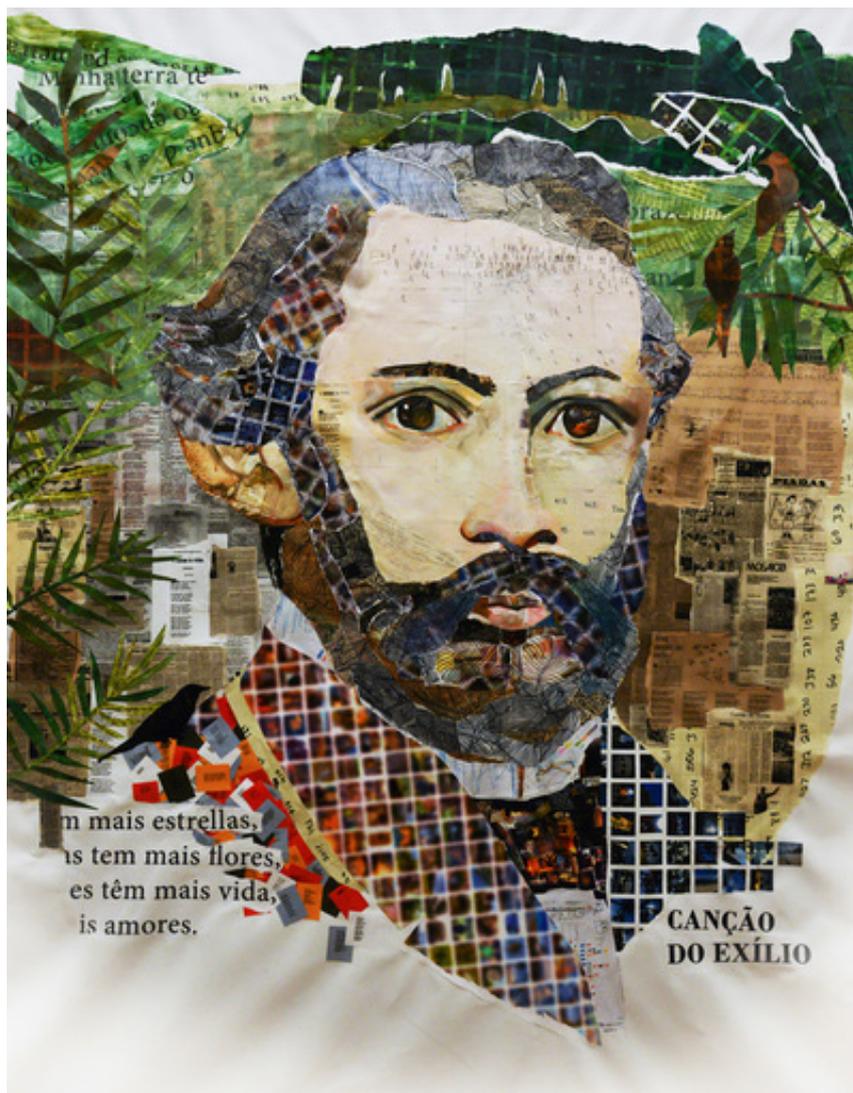


Figura 3.1 Retrato de Gonçalves Dias desenvolvido por Alaina Enslin. (Enslin & Enslin, 2016)

Foram desenvolvidos diversos objetos que mostram as variações. Por exemplo, na Figura 3.2 é possível ver uma ilustração da variação “Saudade da Pátria” escrita por António José Ferreira, um português no Brasil que compara as maravilhas de Gonçalves Dias às suas. Nesta ilustração, as palmeiras estão representadas através da imagem de loureiros e em vez do sabiá existe um rouxinol.



Figura 3.2 Ilustração do poema “Saudade da Pátria”. (Enslén & Enslén, 2016)

Joshua Enslén e Alaina Enslén organizam cronologicamente (de 1846 a 2015) os 500 textos através de um gráfico (Figura 3.3). Através deste, mostram a semelhança textual com o poema original de Gonçalves Dias— quanto mais elevado no eixo vertical maior é a semelhança. No gráfico ainda é possível ver as três expressões mais frequentes nos textos: terra, palmeiras e sabiá. Para fazer a distinção entre autores do sexo feminino e do sexo masculino foram utilizados papéis de cores diferentes (Enslén & Enslén, 2016).

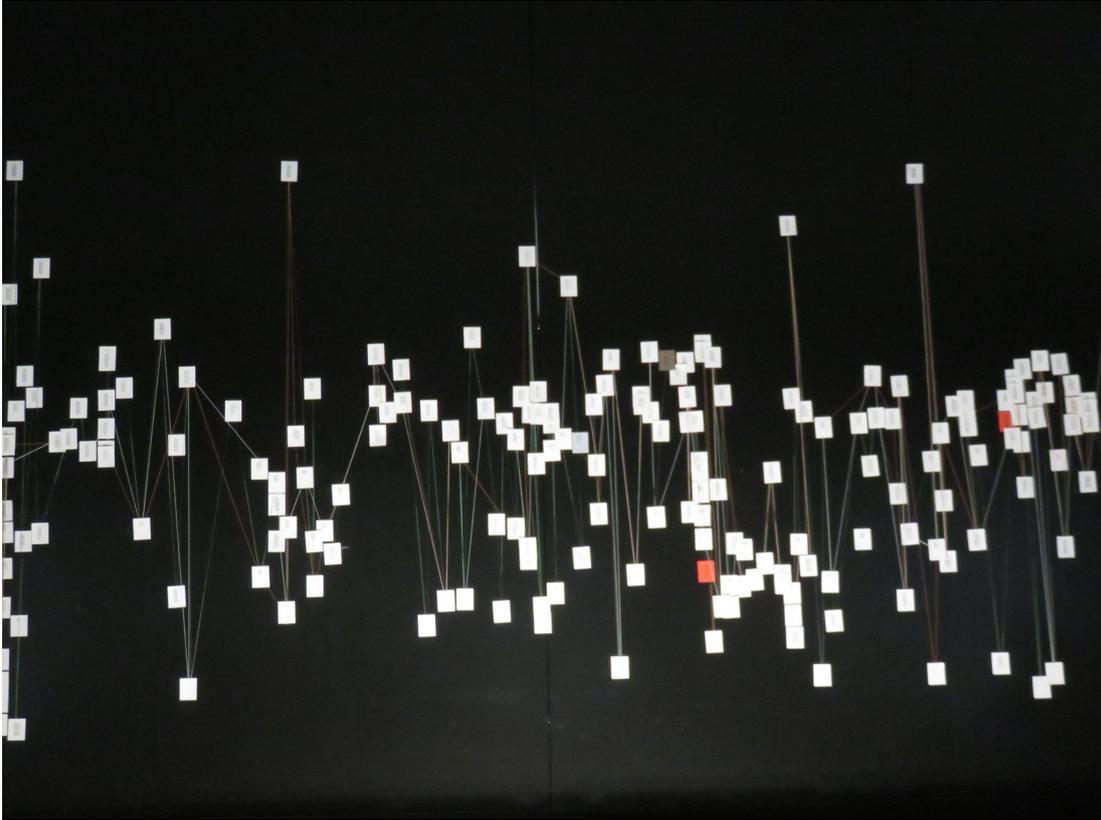


Figura 3.3 Gráfico de semelhanças. (Enslen & Enslen, 2016)

Outro objeto apresentado na exposição de Joshua Enslen e Alaina Enslen foi um quadro com as fotografias presentes na web que fazem referência ao poema “*Canção do Exílio*”. Essas fotografias foram recolhidas, analisadas e organizadas pelas suas cores principais, de forma a criar um objeto que represente a influência do poema através do gradiente de cores. (Figura 3.4)

A exposição esteve presente no Museu da Ciência da Universidade de Coimbra de 3 a 26 de junho de 2016, foi selecionada para ir ao festival FOLIO em 2016 como representante do Museu da Ciência da Universidade de Coimbra. e esteve no Edifício Caleidoscópio do Jardim do Campo Grande de 27 de julho a 15 de setembro de 2017.

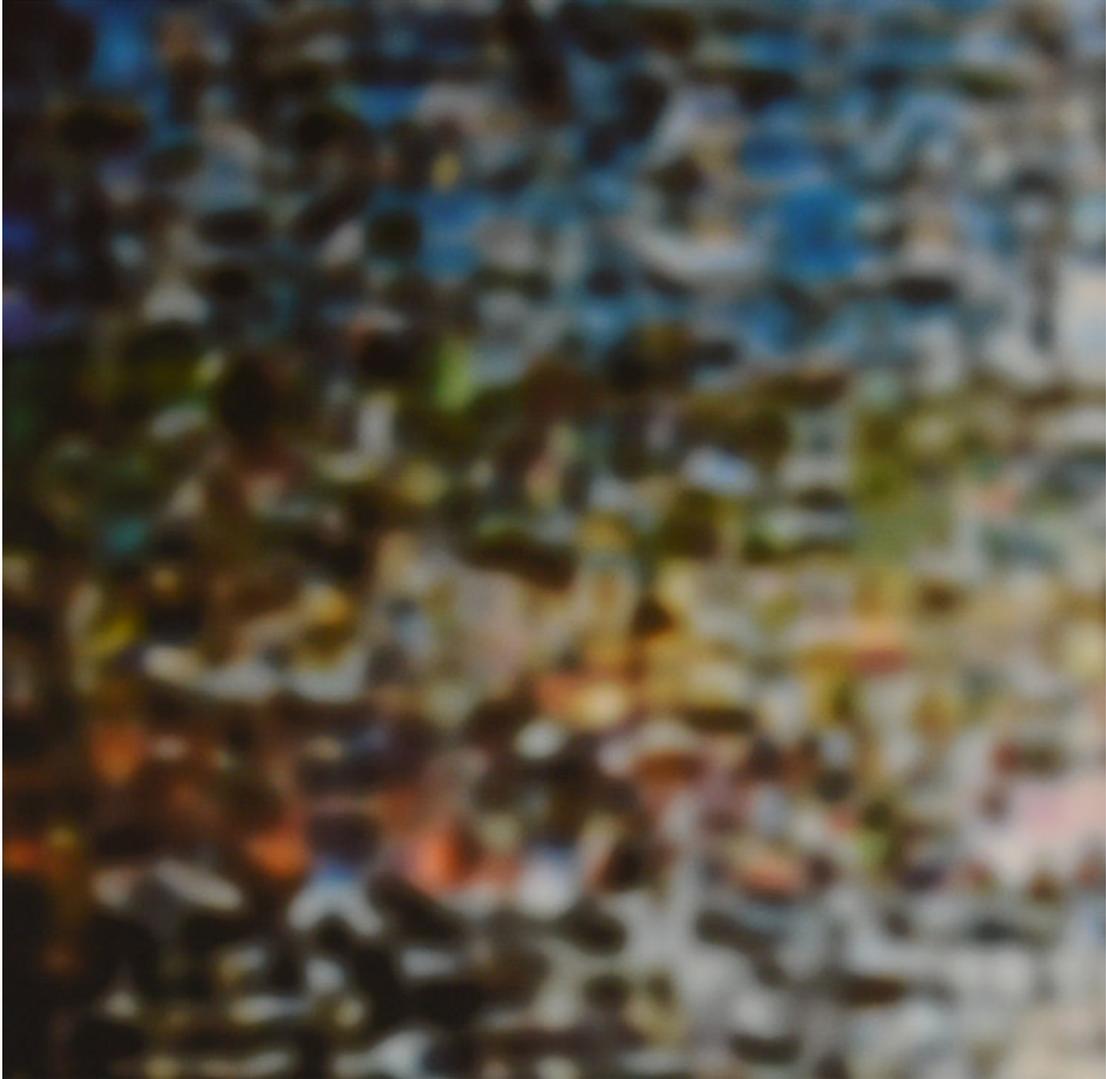


Figura 3.4 Gradiente de cores desenvolvido a partir de fotografias (Enslin & Enslin, 2016)

4 METODOLOGIA

A representação de dados é importante pois facilita a compreensão e análise dos mesmos (Fry, 2004). Neste capítulo é apresentado o plano de trabalho proposto e o plano de trabalho executado. Também é apresentado o processo de trabalho escolhido para o desenvolvimento da dissertação e como é que este se enquadrou com o processo de desenvolvimento real do projeto. É descrito como foi distribuído o tempo por cada tarefa ao longo do projeto. Cada fase do processo de trabalho é explicada de forma a compreender a sua aplicação no projeto.

4.1 PLANO DE TRABALHO

Inicialmente, foi proposto um plano de trabalho (Figura 4.1), onde foram definidas as tarefas gerais mais importantes bem como prazos para o desenvolvimento de cada um dos pontos propostos. Durante o desenvolvimento do projeto, o plano de trabalho inicial sofreu alterações na definição de algumas tarefas, como é possível observar na Figura 4.2, bem como nos períodos propostos para a concretização das tarefas, sendo que estas se mantiveram coerentes com o inicialmente apresentado.

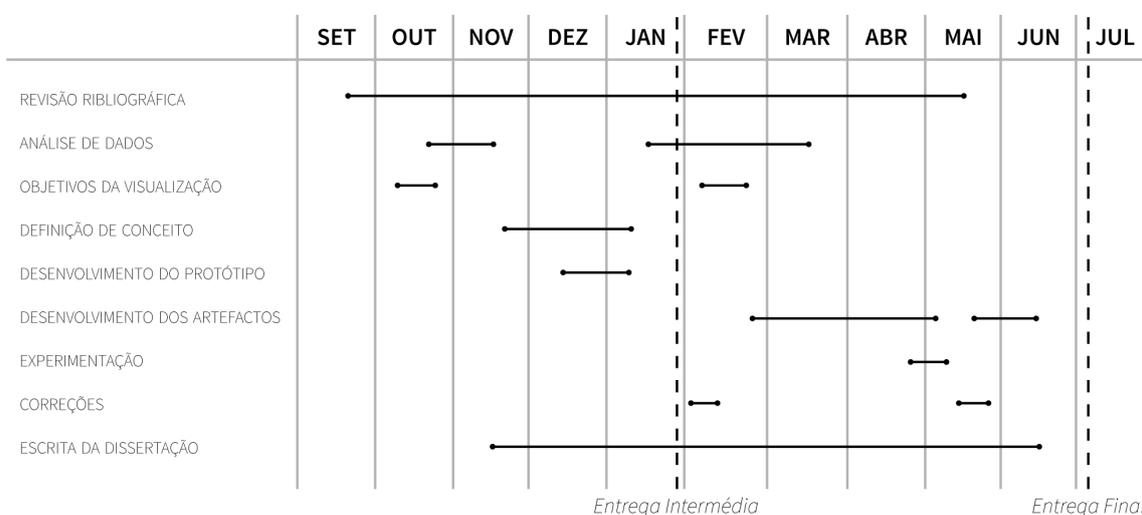


Figura 4.1 Plano de trabalho proposto.

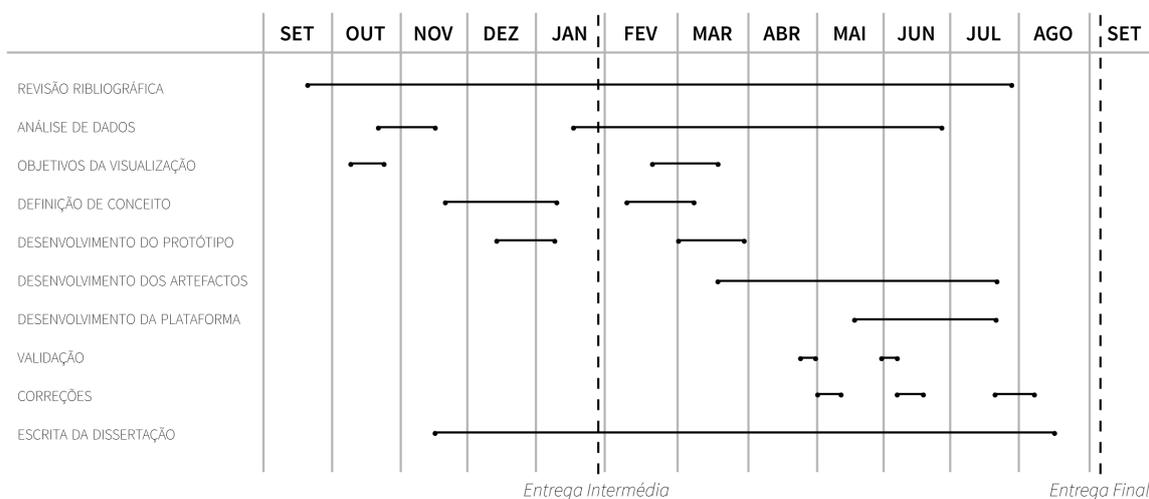


Figura 4.2 Plano de trabalho desenvolvido.

A tarefa de Revisão Bibliográfica engloba a recolha e análise de informação teórica e prática sobre Visualização de Informação e a análise e contextualização histórica das variações. Esta tarefa decorreu como foi inicialmente planeado, tendo ocorrido entre setembro e outubro. Para esta tarefa, foi feita uma pesquisa de poemas e textos que tivessem como inspiração o poema de Gonçalves Dias e uma pesquisa de referências visuais para o projeto. A investigação para o desenvolvimento do Estado da Arte foi feita de outubro a janeiro, tendo sido essencial para preparar o desenvolvimento das visualizações. A partir de fevereiro, esta tarefa serviu principalmente para procurar informação que tornasse os artefactos mais informativos, como por exemplo, a inclusão de momentos históricos e artísticos/literários importantes no Brasil, e eventos mundiais marcantes para os países mencionados nos textos. Esta informação tem como objetivo enriquecer e aproximar os utilizadores do projeto.

A tarefa de Análise de Dados no primeiro semestre esteve dividida em duas partes: a análise de dados de texto e a análise visual das hashtags do Instagram relacionadas com o poema. Ao longo do segundo semestre, focámo-nos apenas nos textos e na exploração dos dados cedidos por Joshua Enslin, tendo sido um processo progressivo com o objetivo de conseguir que os artefactos representassem informação que se complementem.

Para a tarefa Definição de Conceito, inicialmente foram elaborados esboços em papel do que poderiam ser as visualizações a desenvolver. No segundo semestre, após uma análise mais cuidada dos dados disponibilizados por Joshua Enslin e dos dados gerados, optou-se pelo desenvolvimento de uma plataforma que contivesse todas as visualizações. Foram elaborados novos esboços dos artefactos e foi definida a sua organização na plataforma.

Na Definição de Objetivos de Visualização foi feito o planeamento de cada visualização a partir da análise prévia de dados. Para cada artefacto foram determinados objetivos únicos, tendo estes vindo a ser ajustados com o progresso do projeto de forma a criar uma narrativa entre as visualizações.

A tarefa de Desenvolvimento de Protótipos, no primeiro semestre, foi desenvolvida em papel através de uma primeira análise aos dados cedidos. No segundo semestre, já com mais conhecimento sobre os dados disponíveis e as possibilidades de explorar, foram desenvolvidos novos esboços em papel, tendo seguidamente sido desenhados digitalmente para ver o funcionamento das formas e das cores que estavam a ser exploradas e para perceber qual seria a melhor forma de implementar e desenvolver as visualizações.

O Desenvolvimento dos Artefactos teve início após o planeamento das visualizações e da plataforma e após a escolha da ferramenta de progra-

mação. Esta tarefa teve início após a aquisição de dados suficientes para o desenvolvimento das visualizações pretendidas e decorreu simultaneamente com outras tarefas como: (i) a Análise de Dados, uma vez que as visualizações foram desenvolvidas à medida que surgiam mais resultados; (ii) a Validação, pois das reuniões com Joshua Enslen resultavam novas ideias para o desenvolvimento dos artefactos; (iii) o Desenvolvimento da Plataforma, pois era necessário criar algo que estivesse o mais ajustado possível para as visualizações; e por fim (iv) com as correções de pormenores que não estavam tão bem conseguidos como era o pretendido.

O Desenvolvimento da Plataforma é uma tarefa que não existia no primeiro plano de trabalho e surge após o desenvolvimento do conceito e da necessidade de criar um espaço no qual as visualizações pudessem ser facilmente acedidas e analisadas. A tarefa teve início quando as visualizações se encontravam num ponto de desenvolvimento mais avançado tendo ajudado a organizar e criar a harmonia necessária entre elas. Juntamente com as visualizações, na plataforma existe ainda informação sobre o poema, a exposição e sobre a presente dissertação.

A tarefa de Validação surge no novo plano de trabalhos em substituição da tarefa de Experimentação. Para a realização desta tarefa foram realizadas diversas reuniões com Joshua Enslen, resultando em diversas actualizações no desenvolvimento do trabalho. Destas reuniões resultaram ideias importantes para o desenvolvimento do projeto, como a aplicação de filtros nas visualizações, a aplicação de zoom, e a introdução de eventos mundiais e movimentos artísticos.

Na tarefa Correções foram realizados melhoramentos aos artefactos visuais tendo em conta o feedback recebido nas reuniões com Joshua Enslen, tendo sido também efectuados pequenos ajustes na finalização do projeto.

A Escrita da Dissertação teve início após a revisão da documentação cedida por Joshua Enslen e foi sendo desenvolvida ao longo do projeto, tendo o trabalho nesta tarefa ganho mais intensidade a partir de Junho.

4.2 PROCESSO DE TRABALHO

Para o desenvolvimento do trabalho iremos seguir o processo de visualização de dados de Ben Fry, adaptando-o ao projeto.

Aquisição dos Dados: Para o desenvolvimento de visualizações de informação é necessário obter dados credíveis e fidedignos (Fry, 2008). Para este projeto, os dados utilizados foram cedidos por Joshua Enslen. Joshua Enslen cedeu um conjunto de 500 textos relacionados com o poema de Gonçalves Dias, que foram posteriormente analisados, resultando em mais informação para as visualizações.

Análise e Estruturação dos Dados: Estruturar e categorizar os dados ajuda a perceber que mensagem é que podemos transmitir. Os dados utilizados foram organizados da forma mais conveniente à utilização dos mesmo para o desenvolvimento das visualizações, sendo que as tabelas geradas foram sendo atualizadas sempre que existiam novos dados relacionados com a respetiva tabela. Foram acrescentadas à tabela originalmente cedida por Joshua Enslen, as palavras do poema de Gonçalves Dias e a presença de cada palavra em cada uma das variações, bem como os períodos históricos a que as variações pertencem.

Filtragem dos Dados: Removem-se os dados que não têm interesse (Fry, 2008). Desde o início do projeto percebemos que era importante contar a história da evolução e adaptação do poema ao longo dos anos. A filtragem dos dados foi feita de forma a perceber o que as variações poderiam mostrar e como é que a história também estava presente nos textos através do realce de períodos históricos e políticos do Brasil: Império, Primeira República, Estado Novo, Quarta República, Ditadura Militar e Sexta República (ver secção 5.4.4).

Estatística: Aplicar métodos estatísticos ajuda a descobrir padrões nos dados ou mesmo a transformá-los em informação numérica (Fry, 2008). Fez-se a contagem de palavras que mais aparecem no poema “Canção do Exílio” e a sua presença tanto nos outros 500 textos como por período histórico.

Representação dos Dados: Nesta etapa é feita a escolha do modelo de representação visual, a partir de objetivos definidos consoante a mensagem que queremos transmitir. Com os dados disponíveis, no primeiro semestre, desenvolveram-se representações visuais em papel do que se pretendia vir a implementar. No segundo semestre, após uma análise mais cuidada, foi definido o objetivo final para o projeto e, por conseguinte, cada visualização teve que ser repensada/estruturada individualmente.

Refinamento do Modelo Visual: Pretende-se que sejam efetuadas melhorias nos modelos visuais desenvolvidos anteriormente, de forma a que os modelos se tornem mais claros e interessantes para os utilizadores

(Fry, 2008). A partir dos modelos visuais desenvolvidos em papel, foram desenvolvidos modelos visuais digitais, melhorando pormenores como escalas, cores, formas dos objetos. Os artefactos foram tendo melhorias ao longo de todo o processo de criação de forma a irem ao encontro dos objetivos idealizados.

Interação: Neste ponto, os artefactos são trabalhados ao nível da interação, permitindo ao utilizador a exploração das visualizações. A interação permitirá também que o utilizador controle os dados que são apresentados. No contexto da dissertação, foram aplicados filtros que ajudam a criar a ponte entre a história presente em cada variação e a história do Brasil e do Mundo para uma contextualização e compreensão mais completa.

Validação: Nesta etapa verificam-se os artefactos de forma a compreender se estes estão bem concebidos, e que correções são necessárias efetuar. Esta etapa foi efectuada através da apresentação dos artefactos a Joshua Enslin ao longo do desenvolvimento dos mesmos. Através do feedback recebido sobre cada artefacto, foi possível aperfeiçoar cada artefacto tanto ao nível visual bem como a ao nível da interacção.

5 PROJECTO

Neste capítulo são descritos os tipos de dados que vão ser explorados e é explicado o papel que estes vão ter no desenvolvimento dos artefactos. É abordada a evolução do poema de acordo com os momentos temporais que transmitem sentimentos mais positivos ou negativos em relação ao Brasil. É apresentada a proposta de projeto, com alguns protótipos desenvolvidos inicialmente em papel. Por fim, é descrita a forma como é efetuada a implementação dos artefactos visuais.

5.1 DADOS DO PROJECTO

Para o desenvolvimento da parte prática da dissertação foi necessário analisar os 500 textos fornecidos por Joshua Enslen e gerar novos dados a partir desta análise. O estudo dos textos foi feito a partir das diferentes categorias criadas por Joshua Enslen. Estas categorias servem para organizar e explorar os textos através das suas semelhanças e diferenças com o poema original. As categorias são: *Significant Words*, *String Similarity*, *Syntactic Templates*, *Word Tokens* e *Modal Analysis*. A categoria *Significant Words* guarda os elementos com substantivos mais frequentes. A categoria *String Similarity* contém a classificação de cada texto através de um coeficiente relacional. Cada texto foi comparado com o poema original em que foi calculado o número de alterações necessárias para este ser transformado no texto original. Os *Syntactic Templates* são modelos gramaticais existentes no poema de Gonçalves Dias que dão origem às novas versões com modificações. Na categoria *Word Tokens* são guardadas as palavras que alteram as do poema original. Por fim, na *Modal Analysis* é onde se classifica os sentimentos presentes nas variações que podem ser positivos, negativos ou outros. Os dados utilizados para o desenvolvimento das visualizações são referentes a: (i) *Modal Analysis* que engloba a classificação por sentimento dos textos sendo que estes variam entre positivo, negativo e outro; (ii) *String Similarity* que contém o coeficiente de semelhança calculado para cada texto; e (iii) a lista com o número total de vezes que cada palavra aparece nos textos (Apêndice A). A partir destes, conseguimos gerar a lista de palavras do poema original mais utilizadas e mais significativas, criando assim a categoria *Significant Words*.

Estes dados foram verificados e divididos por dois grupos diferentes de dados, Dados Simples e Dados Complexos. Os Dados Simples correspondem a toda a informação sobre cada texto, como o ano em que foi escrito/publicado, o autor, e o título do texto. Os Dados Complexos são todos os dados que foram posteriormente criados e que servem de base para o desenvolvimento das visualizações, como o coeficiente de semelhança (*String Similarity*), os sentimentos prevalentes nos textos (*Modal Analysis*) e o conteúdo dos textos, as palavras do poema original (*Significant Words*) e as palavras mais significativas das variações.

Como auxiliar à geração de dados foi utilizada a ferramenta *AntConc* que permite fazer a análise de conteúdo e concordância textual. Uma vez que não existiam dados suficientes para a exploração das categorias de *Syntactic Templates* e *Word Tokens*, criou-se, a partir dos dados cedidos pelo Joshua Enslen, a categoria de *Word Frequency* que engloba todas as palavras de todos os textos e a frequência com que estas surgem. Esta lista de palavras foi explorada de forma a saber-se também a frequência de palavras por período histórico do Brasil e em quantos textos estas apareciam. Para facilitar a exploração da frequência de palavras (*Word Frequency*), a partir da lista de palavras fornecida por Joshua Enslen, foi

feita uma seleção das mesmas, tornando a lista mais concisa. Para além disto, e de forma a fortalecer a ligação das variações com o poema original, foram excluídos advérbios e preposições e palavras cuja presença era inferior a 9 no total dos 500 textos.

O desenvolvimento da categoria *Significant Words* surge a partir da lista de frequência de palavras, fornecida por Joshua Enslin. Decidiu-se criar uma nova lista de palavras que contivesse apenas as palavras mais significativas do poema de Gonçalves Dias, com a contagem da presença de cada uma nos restantes textos. A partir desta lista verificou-se a existência de cada palavra em cada um dos textos e também foi feita uma matriz para verificar a combinação das palavras entre si nos textos (ex.: minha - terra, terra - minha).

A partir da definição destas categorias desenvolveram-se os artefactos que relacionam entre si mais do que um tipo de dados.

5.2 FERRAMENTAS DE DESENVOLVIMENTO

A escolha da ferramenta para o desenvolvimento das visualizações foi feita com o objetivo de obter resultados visuais o mais próximos do objetivo final, a criação de uma plataforma web que reúna todas as explorações visuais e permita a interação do utilizador com as mesmas.

Foi feita uma pesquisa para perceber a melhor forma de desenvolver os gráficos para a web. Foi escolhido o *JavaScript* para o desenvolvimento e foram consideradas as seguintes ferramentas para a criação dos artefactos: (i) o *Processing.js*, que permitiria fazer o desenvolvimento do trabalho em *processing* e depois adaptar para a web, mas que, embora seja uma boa opção, torna o processo de desenvolvimento da plataforma mais lento; (ii) o *Raphaël JS*, que é uma biblioteca do *JavaScript* que permite desenvolver os objectos pretendidos através de vectores mas que não tem já implementada uma forma de combinar os objectos com os dados, sendo a própria documentação mais centrada no desenvolvimento de objectos; e (iii) o *D3.js*, que também é uma biblioteca de *JavaScript* dedi-

cada ao desenvolvimento de gráficos e representação de dados, com um bom conjunto de referências e documentação e sendo assim a mais simples das três ferramentas para o desenvolvimento das ideias pretendidas.

Para cada ferramenta foram realizados pequenos testes de forma a ser possível escolher a mais indicada para o desenvolvimento dos artefactos. Dos testes realizados com o *Processing.js* não foram conseguidos quaisquer resultados visuais tendo sido tomada a decisão de testar as outras bibliotecas encontradas. Com o *Raphaël.js* começou-se por realizar testes de formas para perceber o funcionamento da biblioteca e progrediu-se até se conseguir desenvolver um gráfico de barras, mas sem dados reais (Figura 5.1). Foram ainda testadas as representações visuais, explicadas na secção 5.4.1, para os gráficos (Figura 5.2).

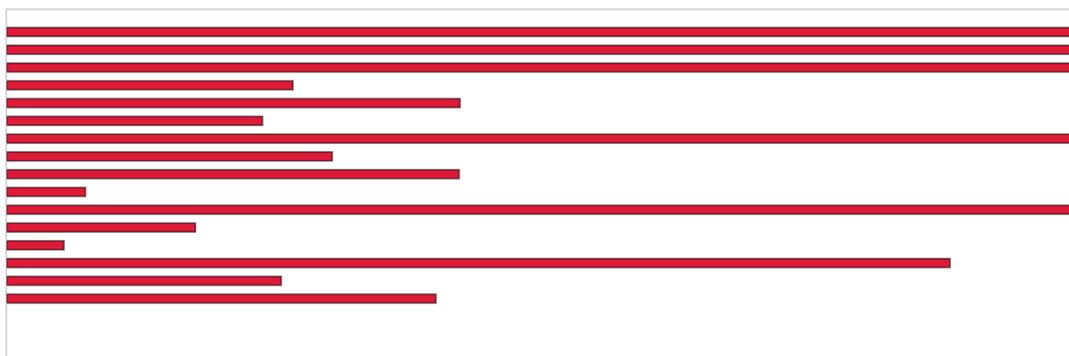


Figura 5.1 Experimentação da ferramenta Raphaël.js.



Figura 5.2 Desenho das formas exploradas na biblioteca Raphaël.js.

Após uma maior experimentação das ferramentas, o *D3.js* foi selecionado para o desenvolvimento dos artefactos, uma vez que esta biblioteca se revelou a mais simples e intuitiva para a obtenção dos resultados desejados e a que tem a documentação mais completa disponível online. Por exemplo, na Figura 5.3, podemos observar os primeiros testes da biblioteca utilizando já os dados disponibilizados. O resultado desta primeira experimentação aproxima-se do resultado final de uma das visualizações pretendidas.

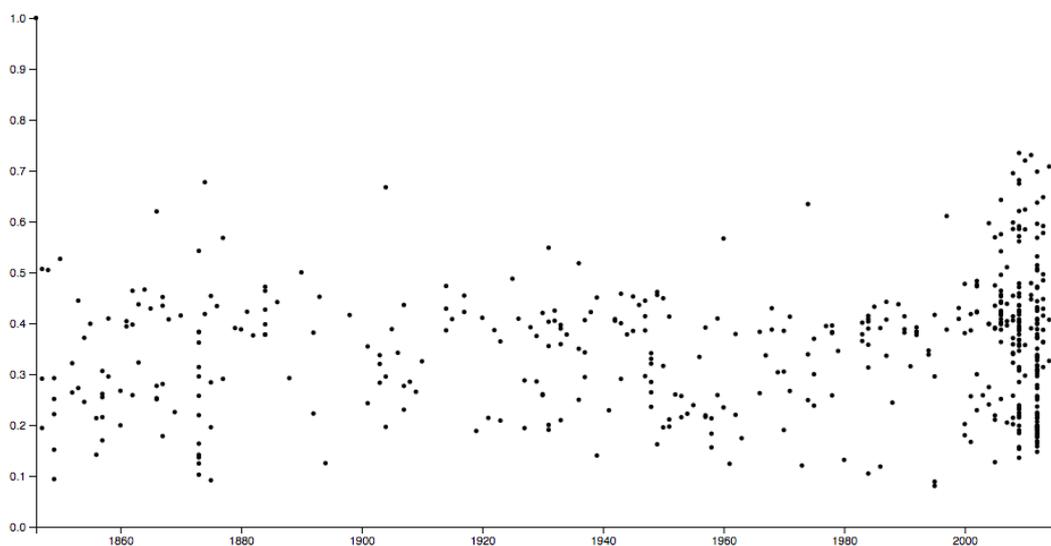


Figura 5.3 Primeiros resultados obtidos com a exploração da ferramenta *D3.js*.

5.3 PRIMEIROS ESBOÇOS

Como proposta de projeto, foi sugerido o desenvolvimento de uma aplicação que contenha várias visualizações sobre os dados. Pretende-se que os utilizadores desta aplicação tenham várias perspetivas sobre a história e evolução da “*Canção do Exílio*” e a oportunidade de explorar os objetos.

Nestes esboços iniciais, iriam ser desenvolvidos um conjunto de artefactos focados em dados distintos sobre o poema. Apesar disto, e embora os artefactos sejam vistos como objetos diferentes, pretende-se que estes se relacionem entre si. Um dos artefactos iria fazer a representação da origem das variações. O utilizador poderia ver que parte do poema de Gonçalves Dias originou e inspirou as variações. Neste objeto, as variações também se ligariam entre si devido à semelhança de frases, análise desenvolvida por Joshua Enslin. No centro do ecrã estaria localizada a representação do poema “*Canção do Exílio*” de Gonçalves Dias enquanto que à volta do ecrã estão posicionadas as variações. Estas variações iriam ter o identificador positivo/negativo/outro. O utilizador poderia seleccionar uma frase do poema e ver que poemas se relacionam com essa frase. Esta semelhança seria ainda destacada através do aparecimento da frase igual nos poemas. Podem ser frases parcialmente semelhantes como para a frase “*Minha terra tem palmeiras*” uma variação “*Minha terra tem loureiros*” de António Ferreira (1847). Também seria possível ao utilizador seleccionar uma variação, ver com que variações esta partilha semelhanças e ver as semelhanças com o poema original (Figura 5.4).

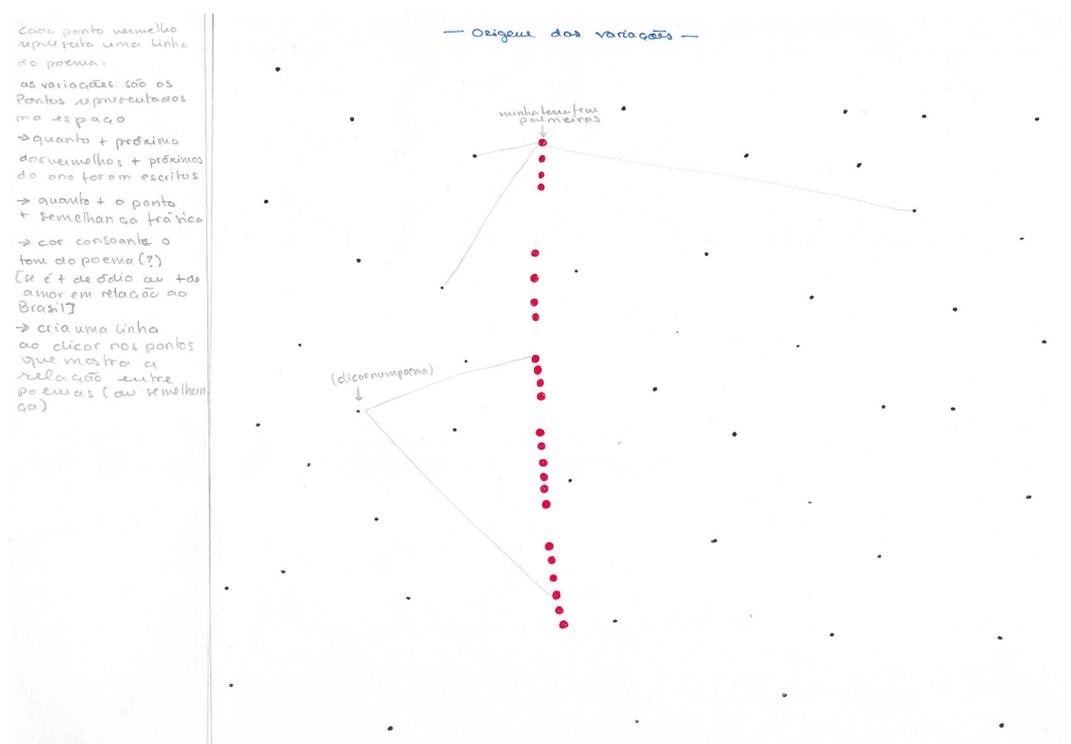


Figura 5.4 Protótipo em papel das origens das variações.

Outra das representações propostas no primeiro semestre permitiria que o utilizador visse o poema através de imagens. Neste artefacto, o utilizador veria as imagens do *Instagram* que foram identificadas com a *hashtag* correspondente à frase do poema original, mas também poderia ver o poema em si. Como referido anteriormente, existem frases que não são utilizadas como *hashtag*, sendo que, um dos desafios deste artefacto seria encontrar um equilíbrio entre texto e imagem.

O *Storyboard* representado na Figura 5.5 mostra como o poema seria apresentado ao utilizador, através de retângulos. O utilizador ao seleccionar um retângulo vê a imagem do *Instagram* que foi identificada com a *hashtag* da frase. Para fazer uma seleção das fotografias iria ser calculada uma média de fotografias por cada *hashtag*. O tamanho de cada retângulo seria calculado através dos gostos que cada fotografia tem, logo uma fotografia com mais gostos ocuparia mais espaço na linha.

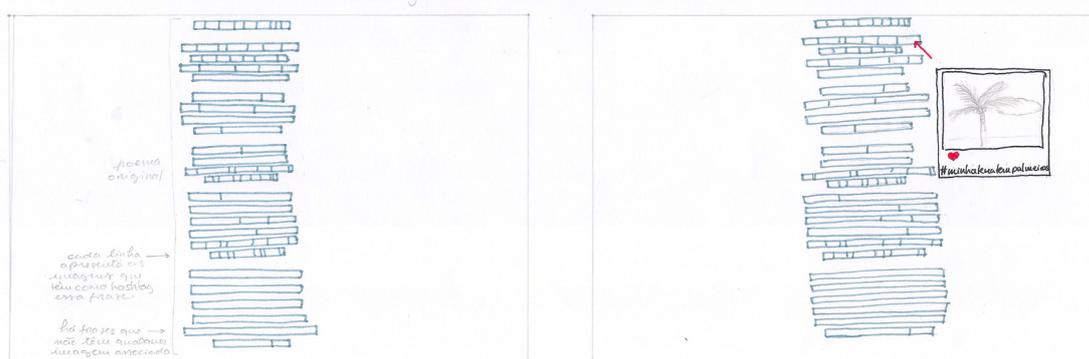


Figura 5.5 Storyboard em papel do funcionamento da visualização através de fotografias do *Instagram*.

Na Figura 5.6 está representado o poema do lado esquerdo e as linhas retangulares do lado direito. Cada linha da estrutura da direita corresponde a uma linha do poema, para realizar o preenchimento idealizado seriam procurados os versos do poema no *Instagram* como sendo *hashtags*. Após uma pesquisa e análise dos *hashtags*, concluiu-se que uma linha pode ter vários *hashtags* associados, ou nenhum. Por exemplo, foi possível verificar que na frase “Onde canta o Sabiá” existem quatro *hashtags* associadas: #ondecantaosabiá; #minhaterratempalmeirasondecantaosabiá; #minhaterratempalmeirasondecantaosabiá; #ondecantaosabiá. Enquanto que na frase “Sem qu’inda aviste as palmeiras”, não existe qualquer *hashtag* associada. Esta incoerência, levou ao abandono desta proposta de usar o *Instagram*.

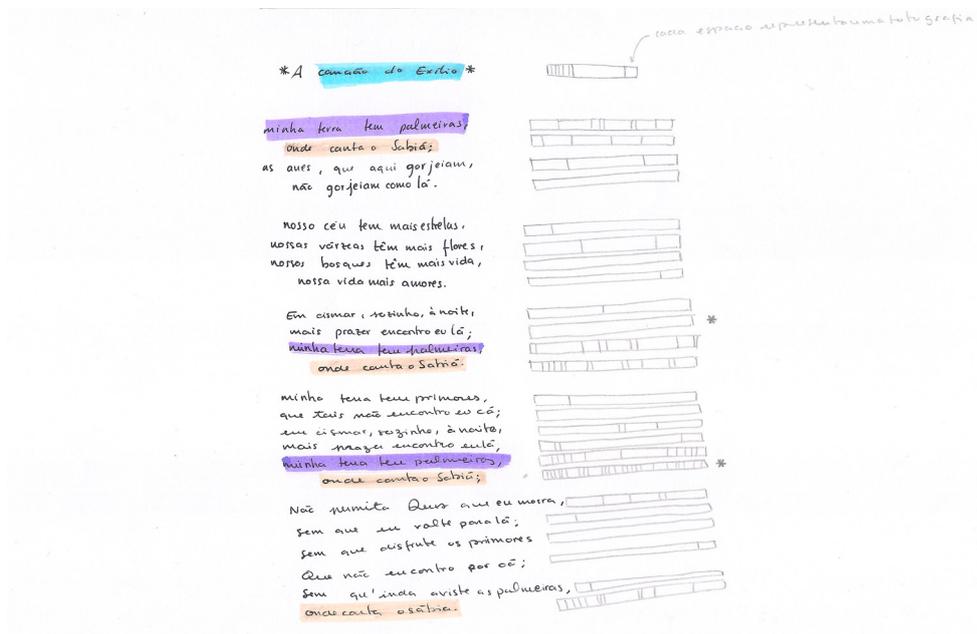


Figura 5.6 Explicação em papel da visualização através de fotografias do Instagram.

Noutra das representações propostas pretendia-se que seja atribuída uma cor às frases do poema mais citadas que serviram como inspiração para a criação de paródias. Desta forma, pretendemos criar uma mancha de cor que relacione as partes das variações com o poema original, evidenciando frases comuns. (Figura 5.7) Os poemas, através da semelhança iriam posicionar-se de forma a adquirirem as cores das frases que têm presentes nos seus textos. Numa breve análise foi possível verificar a frequência de várias frases em alguns textos como no poema “*Modinha do Exílio*” de Ribeiro Couto (1939) que tem a presença de três frases originais, também foi possível verificar que em alguns textos a presença limita-se a uma frase, como é possível ver no poema “*Cor Local*” de Glauco Mattoso (1978).

5.4 DESENVOLVIMENTO DO PROJETO

O desenvolvimento do projeto foi feito de forma progressiva e em simultâneo com a análise de dados. Inicialmente, foram definidos os objetivos para cada uma das visualizações e elaborados esboços em papel de forma a explorar o potencial dos objetivos idealizados. Após a conclusão dos esboços iniciou-se o desenvolvimento dos artefactos recorrendo à biblioteca *D3.js* para *JavaScript*.

Ao longo de todo o processo de desenvolvimento e sempre que era concluída uma etapa foi importante obter a aprovação e o feedback de Joshua Enslin, que desta forma ajudou a que o projeto fosse orientado para o tipo de informação mais relevante a apresentar aos utilizadores.

Para as visualizações desenvolvidas foram atribuídos nomes de forma a enfatizar a ideia de hierarquia e ligação entre as visualizações. A primeira visualização, que apresenta as variações numa timeline, chama-se “*O Poema ...ao longo do tempo*”. Para a segunda visualização foi escolhido o nome “*O Poema ...presente nos textos*”, esta visualização é um radar chart que apresenta a influência do poema de Gonçalves Dias nas variações. A última visualização apresenta o número de vezes que as palavras aparecem nos 500 textos através de um bubble chart com o nome “*O Poema ..em palavras*”.

5.4.1 ESTUDO DE SÍMBOLO E CORES

Uma das principais preocupações com o desenvolvimento das visualizações é a forma como é feita a representação e distinção dos dados que estão a ser apresentados em cada um dos artefactos. Surgiu assim um processo de criação e exploração de diferentes formas de representar os dados, bem como de diferentes cores que pudessem ser usadas de igual forma em cada uma das visualizações, criando uma maior coerência entre visualizações.

Para a escolha das cores foi feita uma pesquisa, com o intuito de criar um projeto visualmente acessível a todos os utilizadores. Deste modo, as cores foram escolhidas de forma a que fossem passíveis de serem distinguidas por uma pessoa que apresente qualquer tipo de daltonismo. Nessa pesquisa, foi possível perceber que combinar determinadas cores torna-

ria o projeto visualmente inacessível e, se necessário, a melhor alternativa seria o uso de texturas como auxiliar para a distinção de cores (Santana, 2017). Na Figura 5.9 está representado parte do primeiro artefacto com as cores escolhidas inicialmente à esquerda enquanto que do lado direito é possível ver a mesma parte do artefacto com a variável de daltonismo protanopia que consiste na incapacidade de ver a cor vermelha e as suas tonalidades (Santana, 2017).

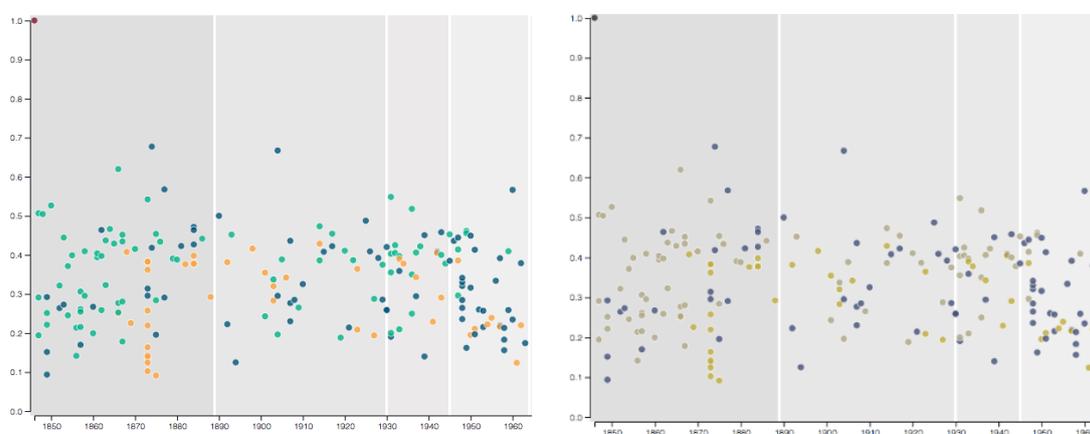


Figura 5.9 À esquerda está representada a visualização com as cores normais, à direita está a visualização com as cores do tipo de daltonismo protanopia.

No processo de desenvolvimento da forma para representar cada texto procurou-se, inicialmente, utilizar formas geométricas simples, usando a linha e explorando a espessura e descontinuidades (Figura 5.10), para representar diversas categorias dos dados. Em simultâneo foi criada uma hierarquia de formas que representasse a hierarquia dos dados: (i) texto ou poema completo, (ii) estrofe ou parágrafo de cada texto, (iii) verso ou uma linha; (iv) expressão ou palavra (Figura 5.10). Enquanto que a forma seria utilizada para representar as diferentes partes dos textos, a cor iria fazer a distinção entre o tipo de análise representada. Com esta ideia surgiu o problema de representação dos textos positivos, negativos e outros (*Modal Analysis*). Inicialmente, a solução encontrada foi a variação dos tons da saturação para as cores escolhidas para cada um dos outros temas (Figura 5.11). Foram também desenvolvidos outros estudos de cores com recurso a texturas (Figura 5.12) mas, com a diminuição do tamanho das formas, surgia o problema de interpretação dos dados. No Apêndice B é possível observar o teste das diferentes formas nos primeiros esboços das visualizações.

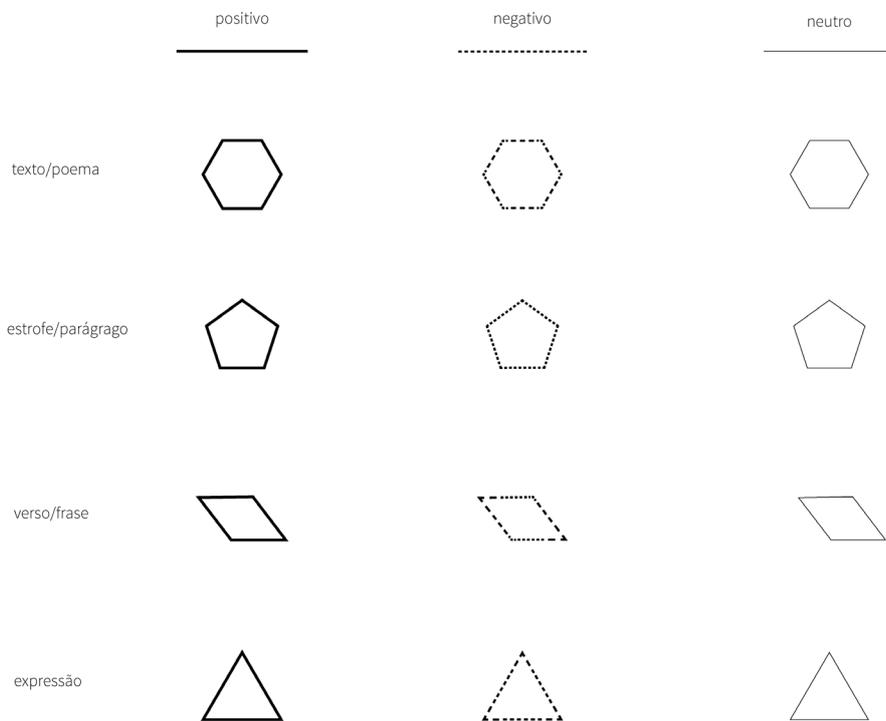


Figura 5.10 Estudo inicial de formas e tipo de linha



Figura 5.11 Estudo de cor por temas e variação dos sentimentos dos textos.

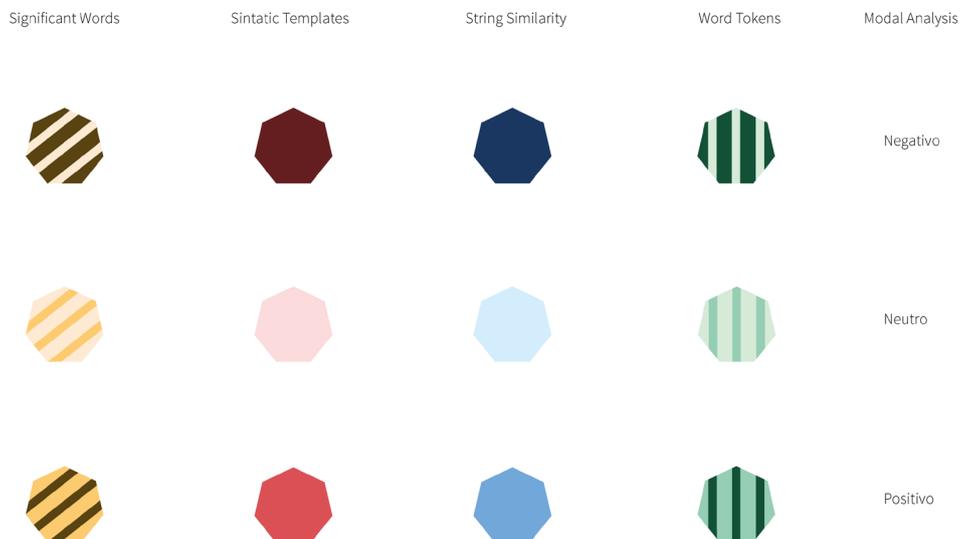


Figura 5.12 Estudo de variação de cores e textura.

Ao longo de todo o desenvolvimento do projeto prático procurou-se explorar continuamente as formas e as cores de forma a combinar estas mais eficientemente entre si. O desenvolvimento das formas mais complexas foi feito a partir da sobreposição de iniciais de cada grupo (ex.: poema/texto = P+T) tendo as formas, a partir daí, sido desenvolvidas de forma a torná-las mais complexas (Apêndice B). Mais tarde, foi idealizado um símbolo e uma cor que representassem o poema de Gonçalves Dias destacando-o dos outros poemas ao longo das diferentes visualizações. As restantes formas foram sendo aprimoradas de forma a que transmitissem a ideia de hierarquia e ordem, uma vez que era importante o utilizador perceber que a diferença de formas estava relacionada com o tipo de dados que estava a ser representado (Figura 5.13). Com as formas desenvolvidas foram realizados testes nos artefactos que se iriam concretizar nesta segunda fase do projeto, mais concretamente com a visualização “O Poema... através do tempo” para testar a legibilidade da utilização de símbolos (Figura 5.14). No Apêndice B é possível encontrar mais imagens relacionadas com o processo de estudo das formas e cor.

Após variados testes de aplicação das formas nos artefactos, optou-se pela utilização apenas da cor como forma de distinção dos dados, uma vez que a representação de dados através de símbolos só seria verdadeiramente utilizada num dos artefactos, quebrando um dos principais objetivos idealizados que era criar harmonia entre todas as visualizações, criando uma relação entre representações.



Figura 5.13 Novo estudo de Símbolos

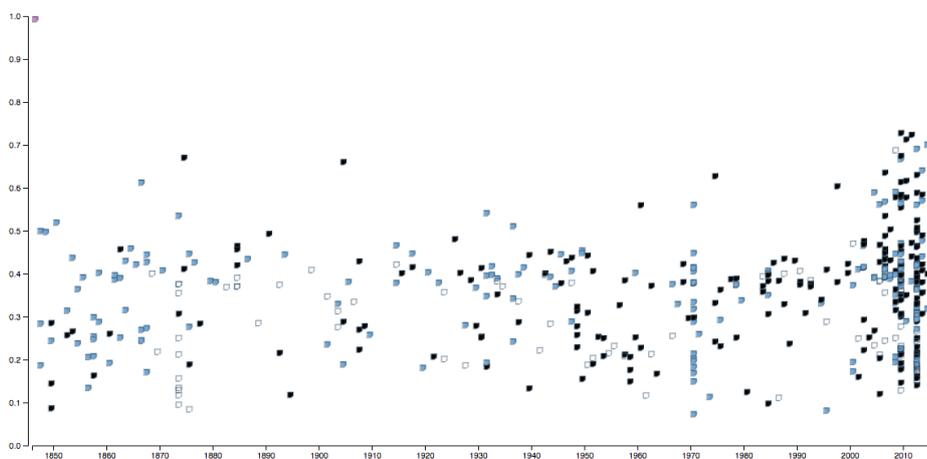


Figura 5.14 Estudo da utilização dos diferentes símbolos na primeira visualização

Como é possível ver na Figura 5.15, a cor foi então usada para distinguir os poemas nas seguintes categorias: verde para positivo, laranja para outros, azul escuro para negativo e vermelho para o poema original. Para a visualização “O Poema ...através do tempo” as cores referidas anteriormente são atribuídas a cada círculo, que representa uma variação, consoante o sentimento com que esta foi classificada, positivo, negativo, outro. Na visualização “O Poema ...presente nos textos” a representação é feita de forma semelhante ao artefacto anterior sendo que aqui é utilizada a linha ao invés do círculo para fazer a representação de cada variação. Na terceira visualização, “O Poema ...em palavras” é apenas utilizada a cor atribuída ao poema original para ser feita a distinção entre as palavras presentes no poema de Gonçalves Dias e as palavras que estão presentes nas variações.

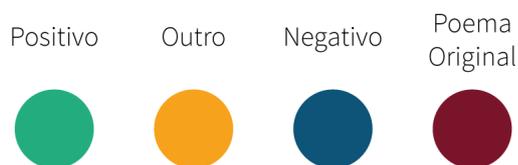


Figura 5.15 - Cores finais que diferenciam os tipos de dados a representar.

As cores escolhidas para os períodos governativos do Brasil foram um gradiente de cinzentos que vai do mais escuro para o mais claro (Figura 5.16). Esta escolha deveu-se ao facto de que esta informação aparecer por trás dos dados principais, no fundo, evitando criar um desequilíbrio entre o que é representado e diferenciando zonas temporais de dados pontuais (poemas).



Figura 5.16 Cores finais que diferenciam os períodos governativos do Brasil.

A representação dos movimentos artísticos é feita através de um gradiente de azuis (Figura 5.17). Uma vez que a representação das variações classificadas como negativas também é feita através da cor azul escura, utilizou-se a transparência de forma a realçar os dados e a mostrar os períodos artísticos (Figura 5.18).

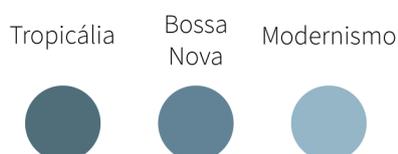


Figura 5.17 Cores finais que representam os movimentos artísticos e literários.

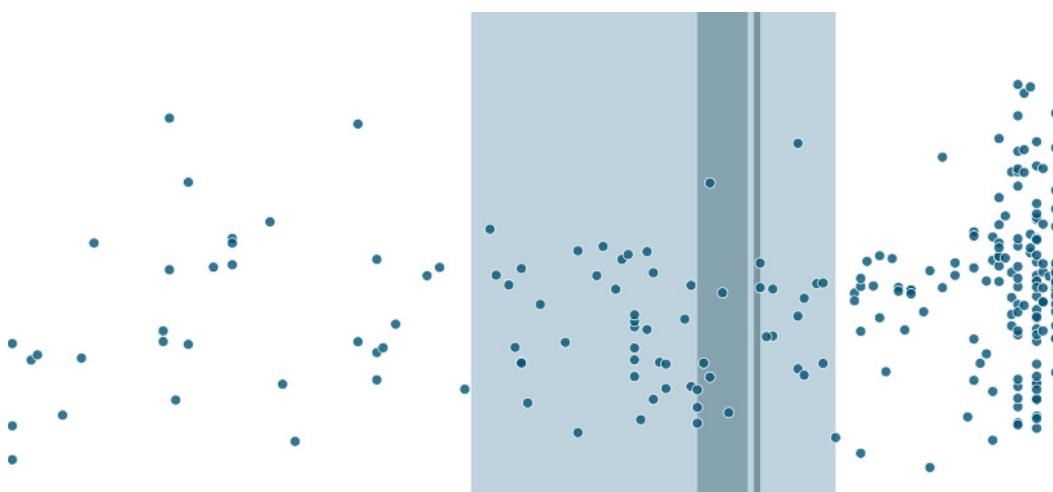


Figura 5.18 Combinação dos dados Negativos com os movimentos artísticos.

Para a representação dos eventos mundiais foi escolhida uma paleta de cores que variam entre castanho e bege (Figura 5.19). Como nos movimentos artísticos, estas cores quando aplicadas na visualização têm transparência, permitindo observar sobreposições de eventos e períodos (Figura 5.20). Assim, foi possível encontrar um equilíbrio entre cores, de forma a coexistirem num espaço limitado sem comprometerem a informação representada por outras cores.



Figura 5.19 Cores representativas dos eventos mundiais.

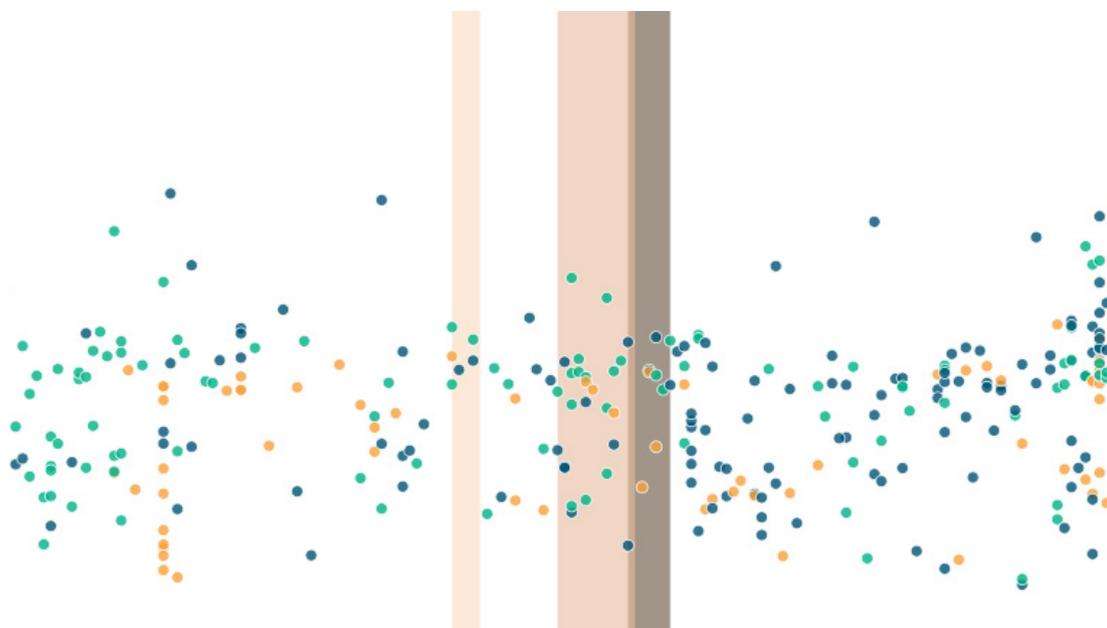


Figura 5.20 Cores representativas dos eventos mundiais aplicadas no artefacto.

5.4.2 EBOÇOS DAS VISUALIZAÇÕES FINAIS

As visualizações foram planeadas faseadamente, à medida que eram conseguidos resultados da análise dos dados que beneficiassem as mesmas. Este método iterativo para desenvolver o conteúdo dos artefactos fez com que surgisse um distanciamento visual entre os protótipos idealizados e sugeridos no subcapítulo 5.3. A partir dos primeiros estudos dos dados foram sendo desenvolvidos protótipos em papel, tendo sido essencial definir os objetivos de cada visualização, uma vez que facilitou a escolha e criação de dados necessários para cumprir cada um dos objetivos.

Na primeira visualização, os objetivos passam por representar períodos históricos, o índice de semelhança (*String Similarity*) e o sentimento (*Modal Analysis*) de cada texto, permitindo ao utilizador fazer uma contextualização histórica, e observar a evolução das variações desde o poema de Gonçalves Dias até aos dias de hoje. Como é possível observar na Figura 5.21 optou-se pela representação destes dados através de uma *Timeline* que no eixo do x fizesse a disposição dos textos por ano e no eixo do y a organização seria feita através do coeficiente de semelhança. Cada objeto representa uma variação e o respectivo sentimento, através de formas e cores (Figura 5.22).

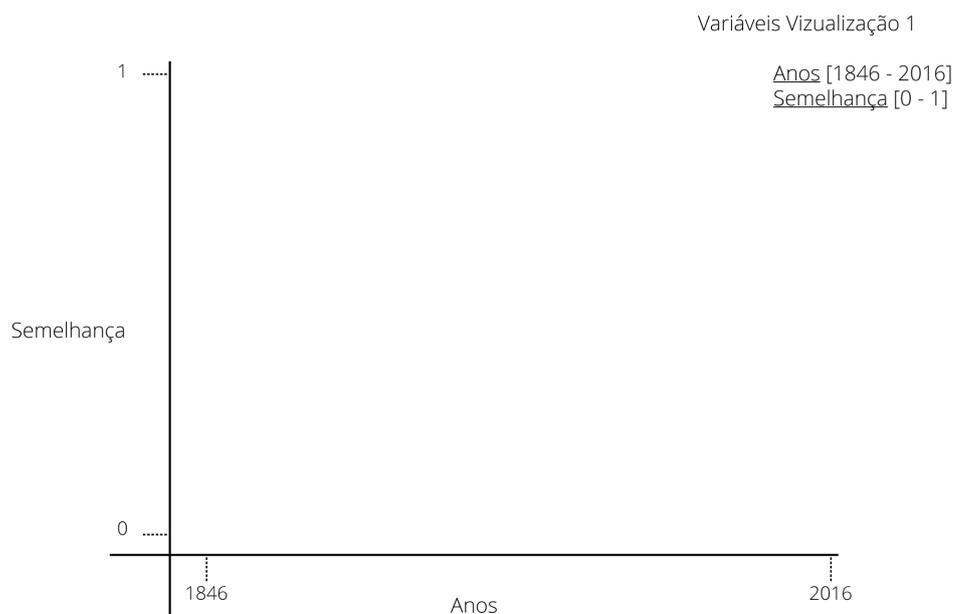


Figura 5.21 Esquema da visualização “O Poema...ao longo do tempo”.

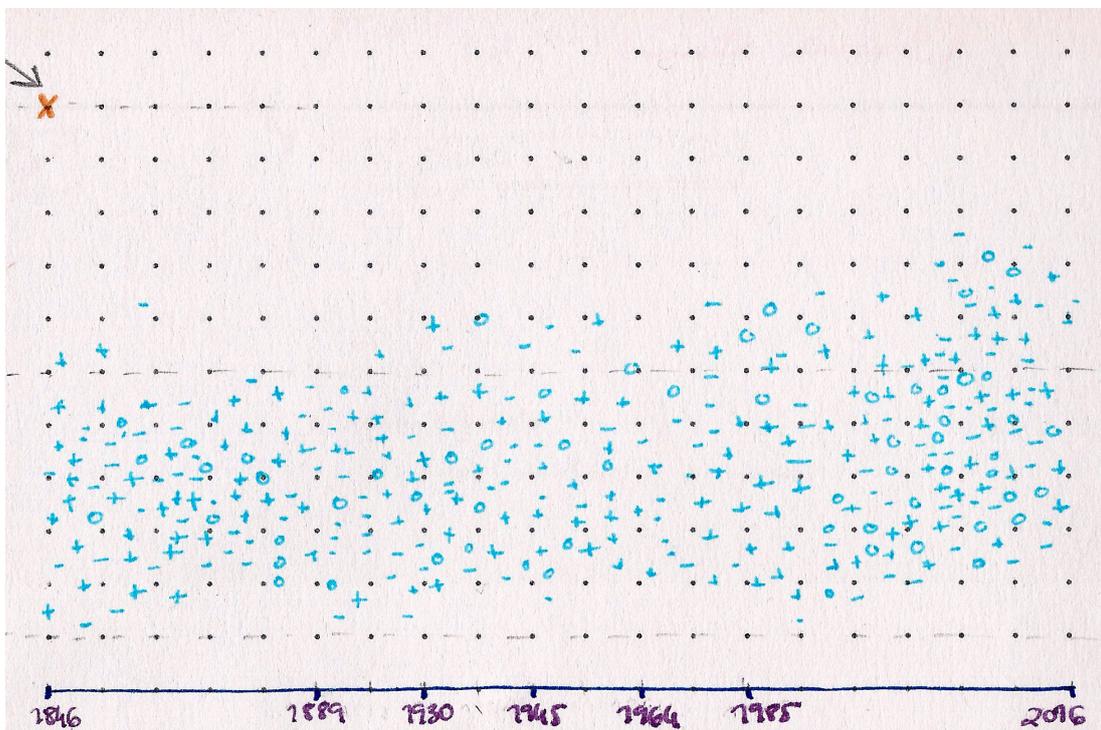


Figura 5.22 Esboço em papel da visualização “O Poema ...ao longo do tempo”.

Os objetivos iniciais para a segunda visualização, passam pela representação da frequência de todas as palavras, dos sentimentos das variações que contêm as palavras e do relacionamento das palavras entre si. O tipo de representação escolhida para esta visualização foi um *Bubble Chart* que permitiria representar a frequência das palavras através do tamanho dos círculos e a relação das palavras através da proximidade entre círculos (Figura 5.23). Neste artefacto, a interação permitiria ao utilizador seleccionar uma palavra, ver as combinações mais frequentes com as restantes palavras representadas e ver os textos que utilizavam essa mesma palavra. Como é possível observar no esquema da Figura 5.24, para cada palavra é representado o valor total da sua presença no conjunto dos 500 textos.

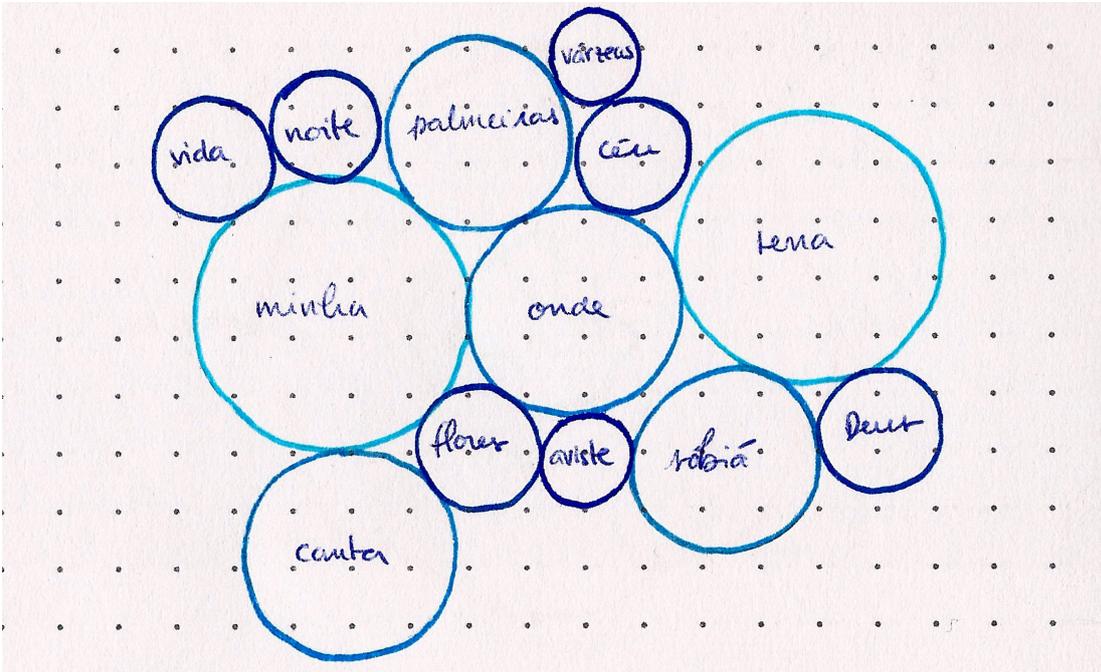


Figura 5.23 Esboço da visualização “O Poema ...em palavras”.

Variáveis Visualização 2

Palavras [Variações e Poema Original]
 Presença Total [0 - 1800]

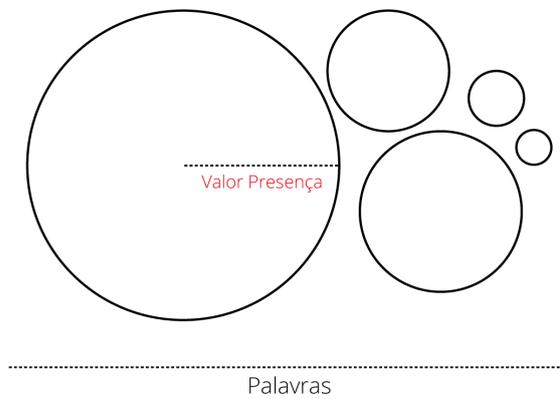


Figura 5.24 Esquema do planejamento da visualização “O Poema ...em palavras”.

Inicialmente a terceira visualização tinha como objetivos mostrar a presença do poema de Gonçalves Dias através das palavras mais frequentes da “*Canção do Exílio*” e mostrar para cada variação que palavras do poema original é que estão presentes. Na Figura 5.25 é possível perceber o funcionamento do gráfico que, através de um *Radar Chart*, teria as palavras da categoria *Significant Words* distribuídas à volta do círculo, sendo que o valor do centro do círculo seria 0 (zero) e o seu valor máximo seria de 1 (um) representando uma normalização dos valores da categoria de *String Similarity*. Com a utilização destas duas categorias, para cada variação é possível criar um identificador único uma vez que sempre que é identificada uma palavra do poema original nessa variação é atribuído o valor da semelhança, representando a influência de Gonçalves Dias para cada texto (Figura 5.26).

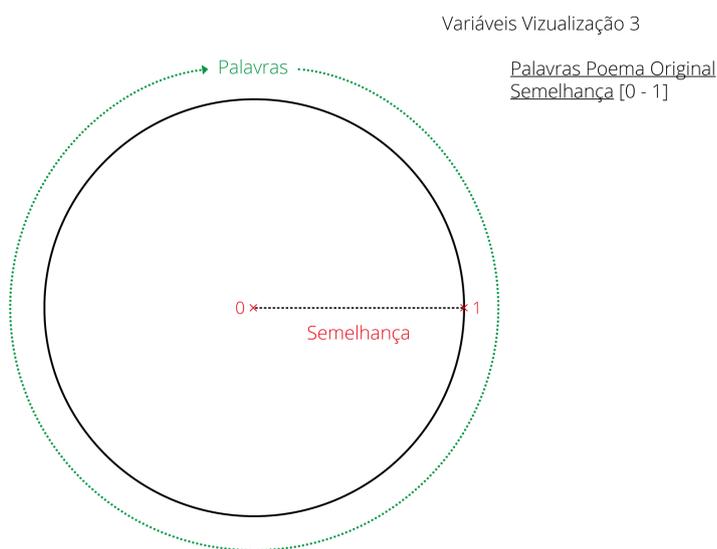


Figura 5.25 Esquema do funcionamento da visualização “O Poema ...presente nos textos”.

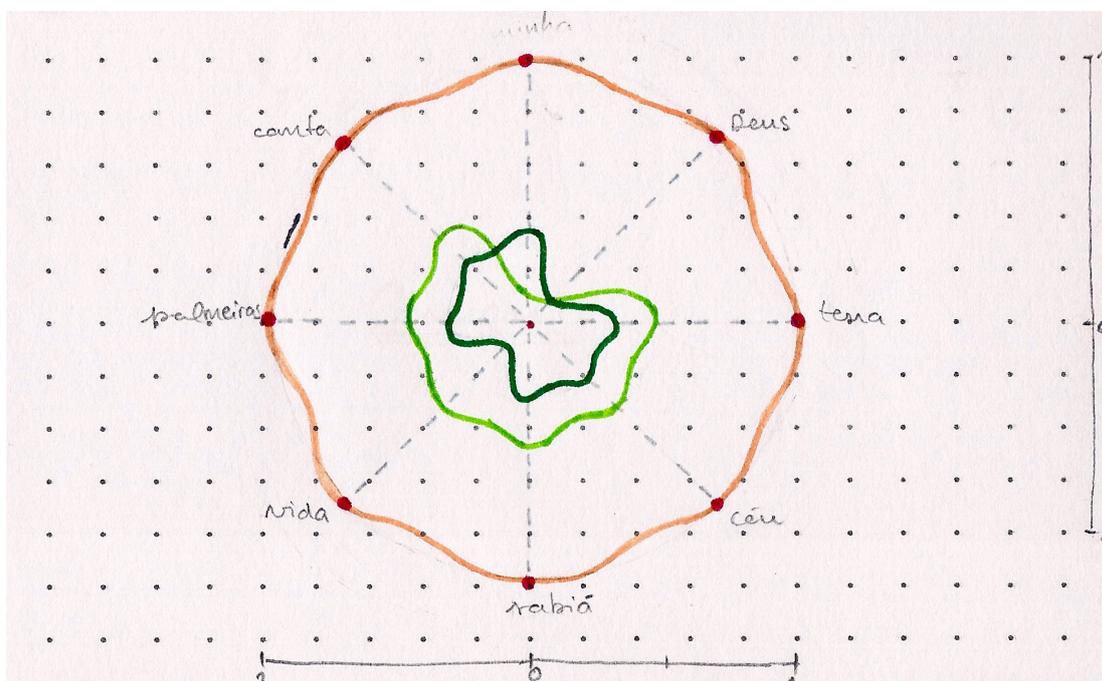


Figura 5.26 Esboço da visualização “O Poema ...presente nos textos” utilizando os textos de Carlos Drummond de Andrade.

5.4.3 DESENVOLVIMENTO DAS VISUALIZAÇÕES

Para o desenvolvimento de cada artefacto foi crucial o planeamento prévio, bem como a idealização do resultado final pretendido e a escolha da ferramenta de desenvolvimento. Após termos as principais decisões definidas, as visualizações foram sendo criadas à medida que obtivemos os tipos de dados necessários para cada uma delas.

A visualização “O Poema ... ao longo do tempo” foi planeada de forma a mostrar as variações no ano em que foram escritas e o seu valor de semelhança com o poema de Gonçalves Dias, acabando por se tornar uma versão interactiva da visualização desenvolvida por Joshua Enslin para a sua exposição (subcapítulo 3.3). Uma vez que os dados necessários para o desenvolvimento da primeira visualização já se encontravam disponíveis (*String Similarity* e *Modal Analysis*) rapidamente foi iniciado o desenvolvimento deste artefacto.

Inicialmente, representámos através de símbolos todas as variações criando a distinção através da categoria de *Modal Analysis*, como é possível observar na Figura 5.27. Com esta representação foi possível perceber que os vários elementos da visualização se poderiam sobrepor, existindo assim a ocultação de dados, principalmente a partir do ano 2000 em que o número de variações aumenta consideravelmente.

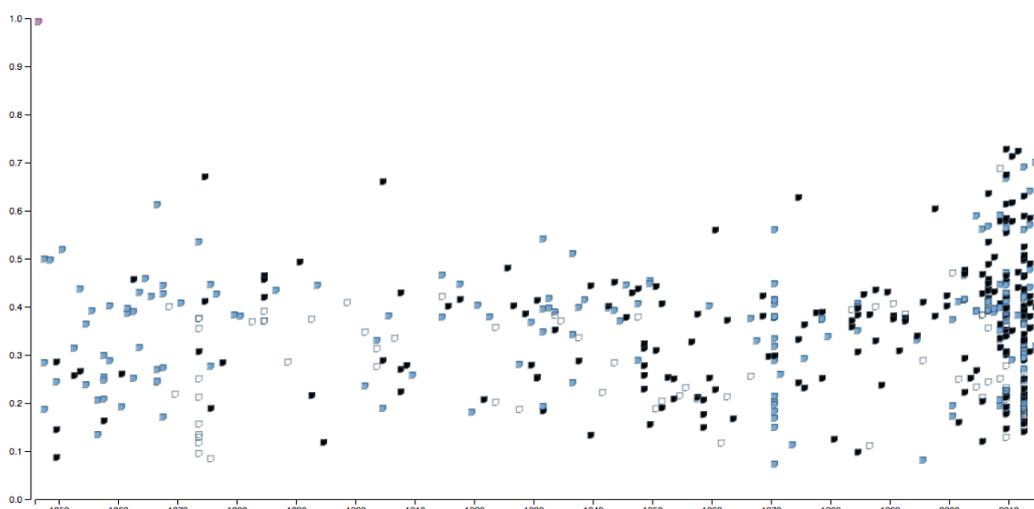


Figura 5.27 Etapa inicial do desenvolvimento da visualização “O Poema ...ao longo do tempo”.

A representação dos anos no eixo horizontal mostra aos utilizadores os anos, desde 1846 até 2016. Neste eixo pretendia-se ainda representar as alturas em que houve mudança de período governativo. Após ser efetuada uma pesquisa pela documentação do *D3.js* percebemos que não seria exequível representar estes períodos históricos no eixo horizontal, pelo que optou-se por fazer a representação dos anos por décadas. Para a representação dos períodos históricos optou-se assim inicialmente pela divisão através de linha (Figura 5.28) e mais tarde pelo sombreamento da área referente a cada período temporal.

O gráfico representado na Figura 5.28 encontrava-se numa fase de desenvolvimento em que já tínhamos optado pela exploração da cor como forma de distinção dos dados ao invés dos símbolos. Nesta imagem, especificamente, a verde estão representados as variações classificadas como sendo positivas, a azul escuro as variações negativas e a amarelo as variações que receberam a classificação de outras. A vermelho está representado o poema de Gonçalves Dias, que consideramos que deveria ser realçado através de uma cor diferente das cores classificativas do *Modal Analysis*.

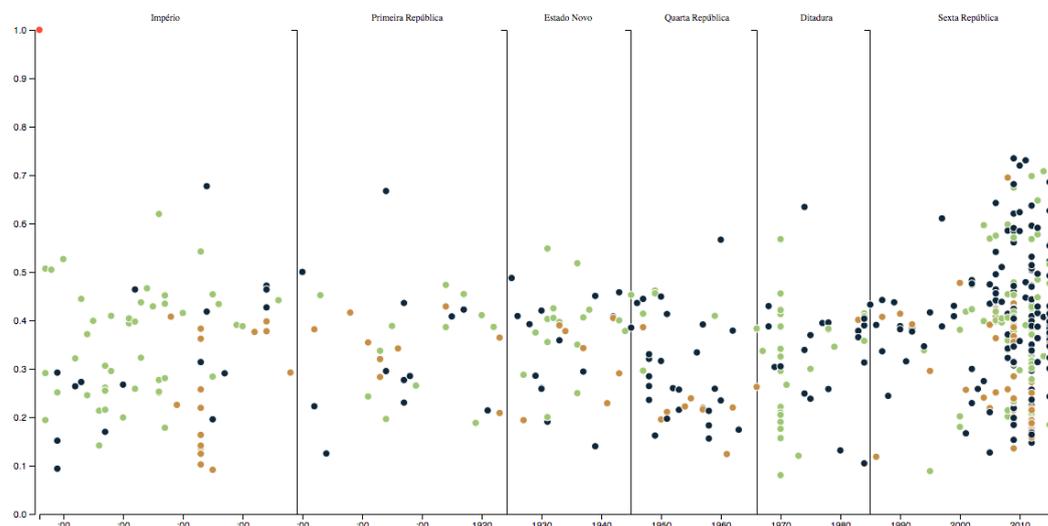


Figura 5.28 Criação da delimitação dos períodos históricos na visualização “O Poema ...ao longo do tempo”.

Na Figura 5.29 já foi aplicado o sombreado nas áreas referentes a cada período histórico permitindo que o utilizador compreenda melhor a duração de cada período. Mais uma vez as cores foram alteradas para tentar perceber quais funcionariam melhor visualmente.

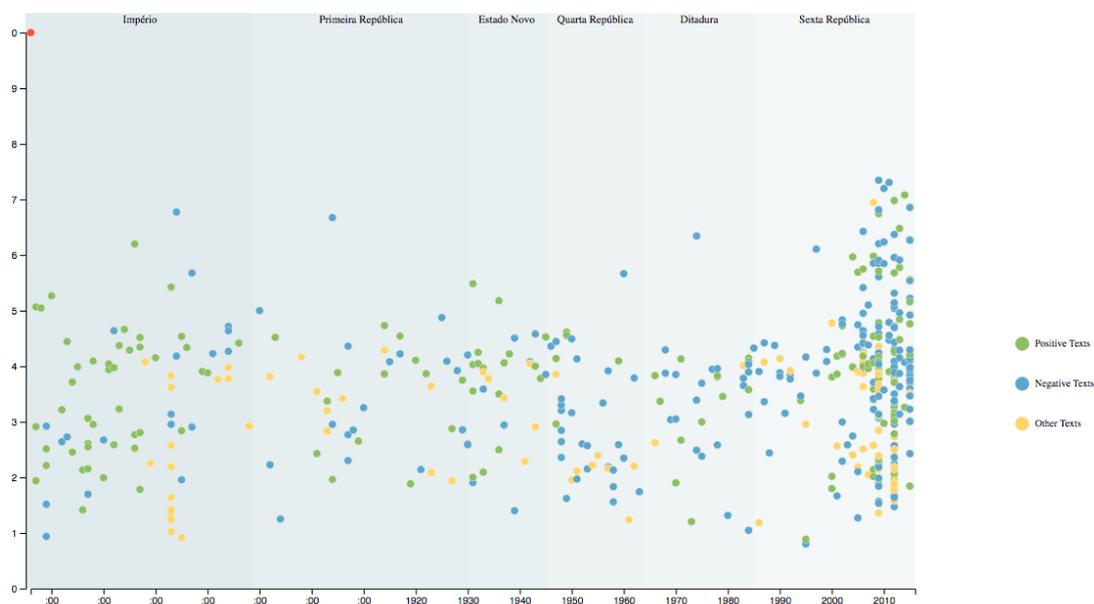


Figura 5.29 Aplicação das áreas sombreadas dos períodos históricos na visualização “O Poema ...ao longo do tempo”.

Na Figura 5.30 apresentamos o artefacto mais próximo ao objeto final. Aqui já estão presentes todos os elementos de interatividade finais: (i) a seleção de tipo de textos, ou seja, o utilizador tem a liberdade de escolher se apenas quer ver as variações positivas, negativas ou outras, ou uma combinação destas; (ii) a escolha de eventos históricos e a visualização dos movimentos artísticos literários; (iii) a seleção de período histórico, que ao ser selecionado já era possível ver a área correspondente a esse período; (iv) a funcionalidade de zoom que permite aos utilizadores verem os elementos gráficos que representam as variações do poema em mais detalhe, principalmente as que se encontram agrupadas ou sobrepostas; (v) um tooltip que permite ver que variação corresponde a cada ponto através do seu título, autor e ano; (vi) e a barra de pesquisa que estava a começar a ser desenvolvida para dar a possibilidade ao utilizador de procurar por um autor ou texto em específico. Em simultâneo eram feitos testes de cor para se tentar perceber como é que se poderiam ver todos os dados apresentados, tendo sido também testado o uso de transparências.

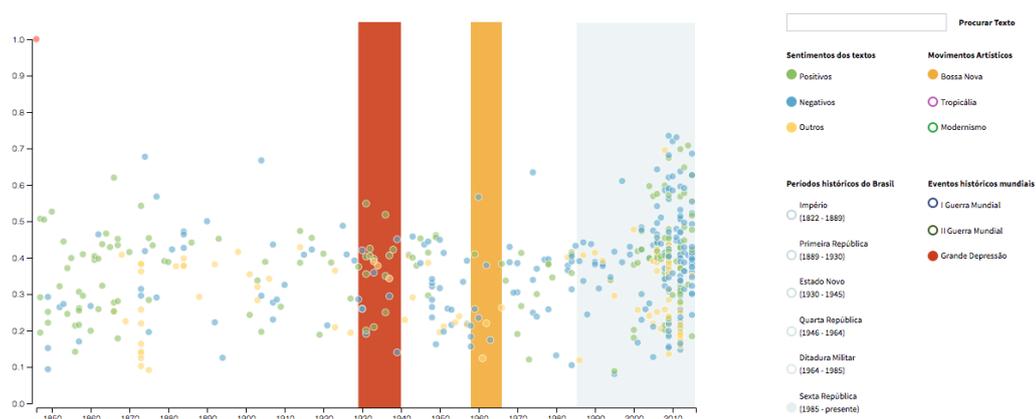


Figura 5.30 Artefacto e filtros correspondentes a “O Poema ...ao longo do tempo.”

Na versão final da primeira visualização (Figura 5.31), o utilizador pode seleccionar e observar as variações como mencionado anteriormente (Apêndice C), bem como fazer a seleção individual dos períodos temporais (Figura 5.32), dos movimentos artísticos (Figura 5.33) e dos eventos históricos mundiais (Figura 5.34). É possível o utilizador procurar por um autor ou título de texto na barra de pesquisa localizada por cima dos filtros (Figura 5.35). Para o desenvolvimento da barra de procura foi utilizado o plugin *EasyAutocomplete* de Łukasz Pawełczak (2015) que permite fazer a pesquisa dos elementos desejados, incorporando o resultado com o código desenvolvido no *D3.js*. Para cada visualização foi ainda desenvolvido um pequeno texto explicativo do que está a ser apresentado ao utilizador e como é que este pode interagir com o artefacto.

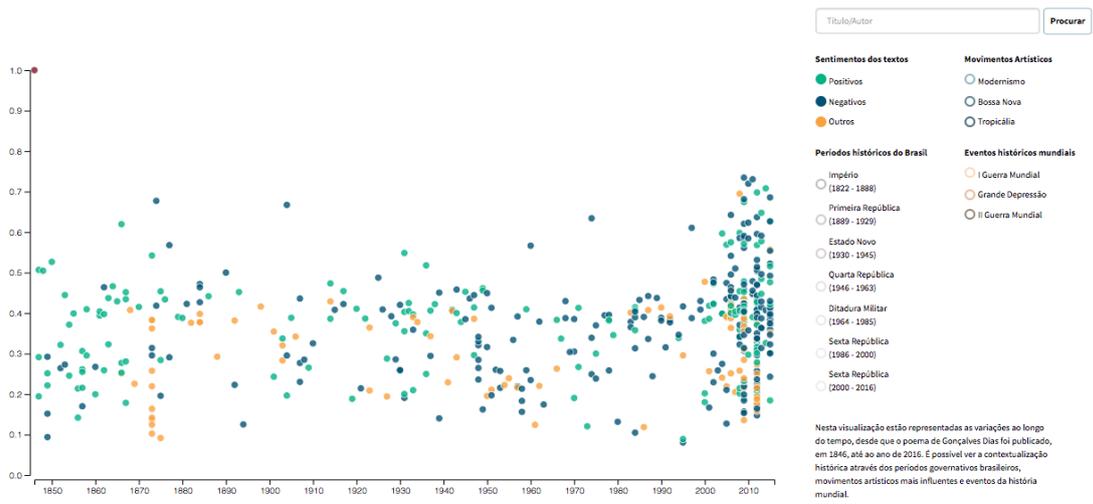


Figura 5.31 Visualização “O Poema ...ao longo do tempo” finalizada.

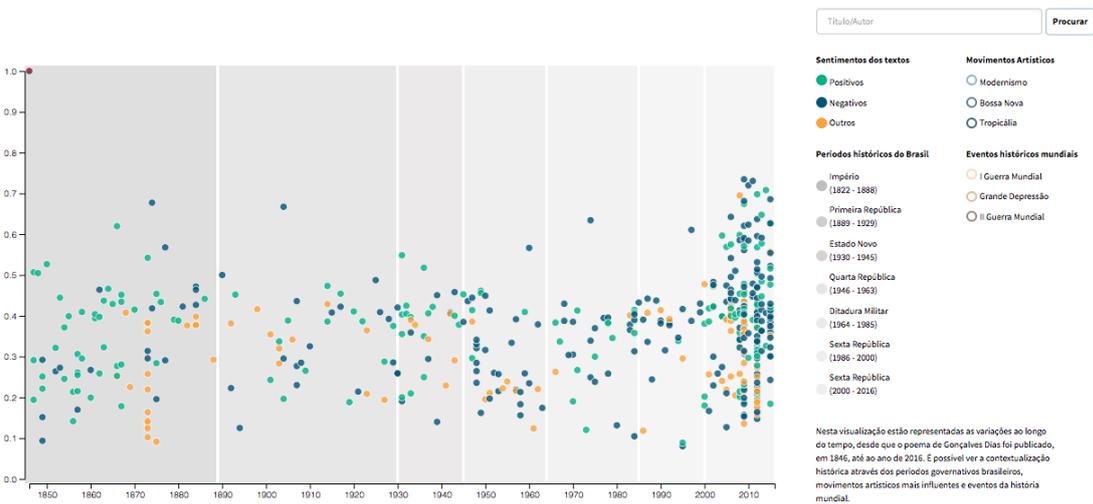


Figura 5.32 Representação dos períodos históricos do Brasil.

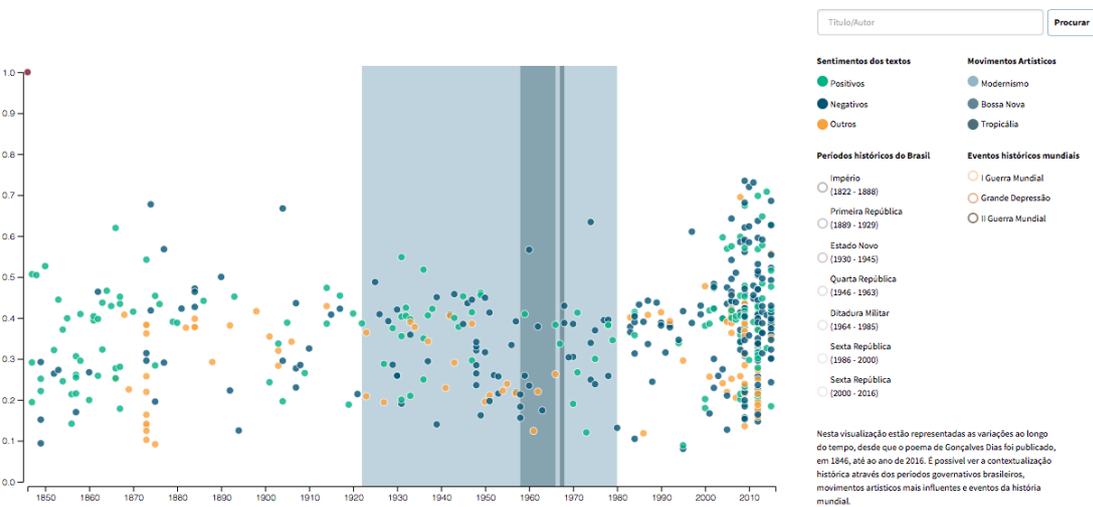


Figura 5.33 Representação dos Movimentos artísticos literários

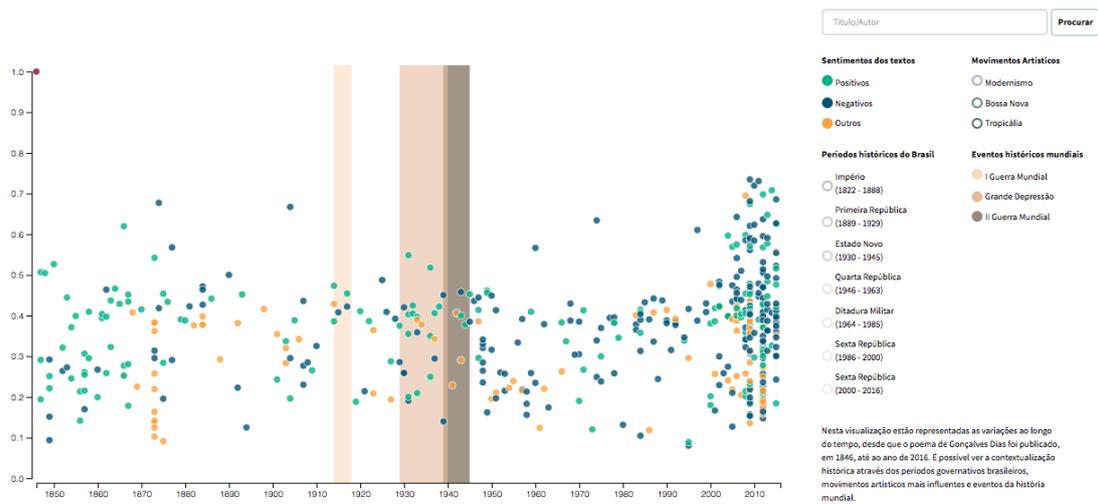


Figura 5.34 Representação dos eventos históricos mundiais.



Figura 5.35 Representação do resultado da pesquisa para o poema com o título “Fim de Festa” com mouseover sobre o poema para aparecer o tooltip no canto superior direito do gráfico.

O segundo artefacto a ser desenvolvido foi “O Poema ...presente nos textos”. Este artefacto é um *Radar Chart* que apresenta as palavras mais relevantes da “Canção do Exílio” (*Significant Words*) presentes em cada variação. As palavras selecionadas estão distribuídas à volta de um círculo em que o coeficiente de semelhança vai desde o centro do círculo até a curva do mesmo (secção 5.4.2). Para o desenvolvimento desta visualização foi seguido o modelo desenvolvido por Nadieh Bremer (2015). Na Figura 5.36 é possível observar o primeiro resultado do artefacto com os dados de todas as variações e as cores selecionadas na fase inicial. Nesta etapa foi possível perceber que seria necessário implementar uma forma de filtrar os dados para que estes se tornassem mais legíveis.

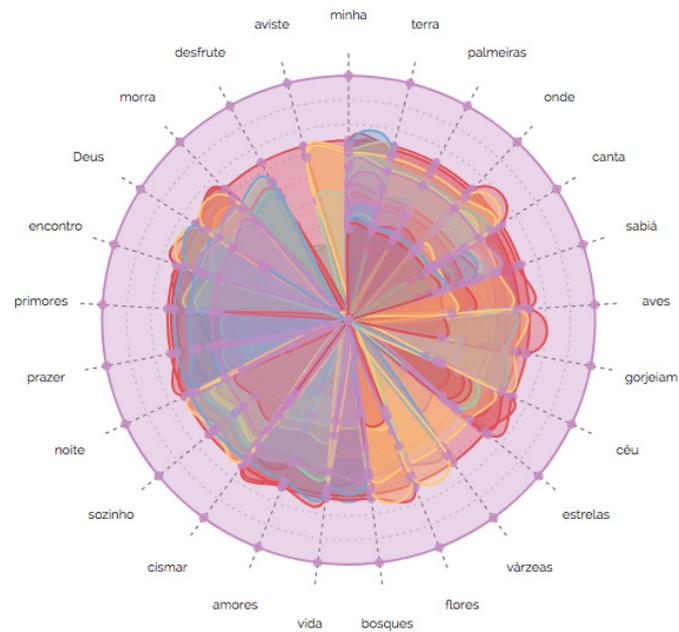


Figura 5.36 Representação inicial da visualização “O Poema ...presente nos textos.”

Na Figura 5.37 está representada a visualização já com os elementos visuais definidos, tendo sido escolhida a utilização de contorno nas formas representativas das variações de forma a facilitar a visualização. Os elementos de interação disponíveis aos utilizadores seriam observar pelos períodos históricos brasileiros, pela distinção entre variações positivas, negativas e outras (*Modal Analysis*) e também através de uma barra de procura.

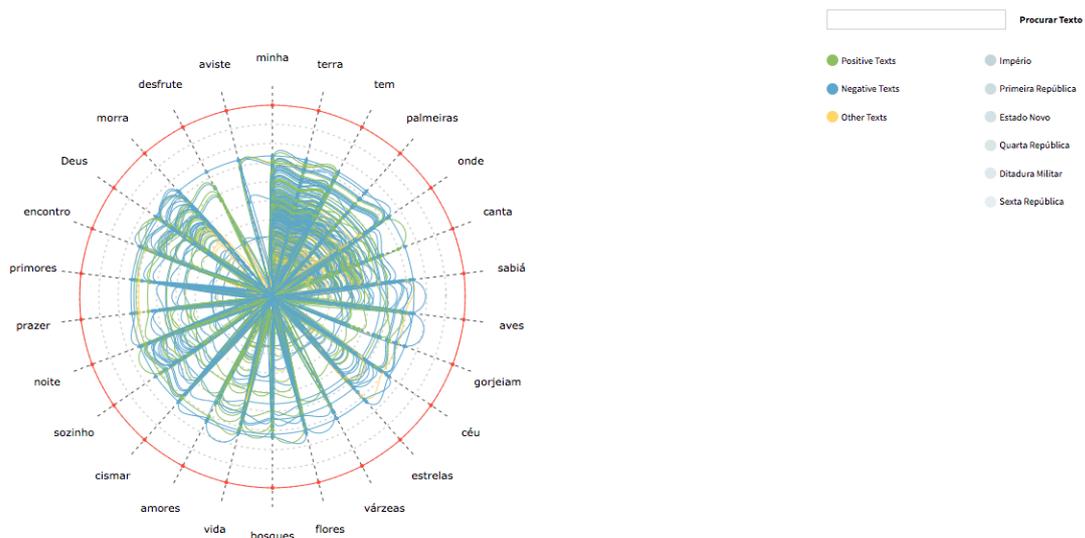


Figura 5.37 Artefacto já com os elementos visuais e de interação definidos.

O gráfico finalizado é apresentado na Figura 5.38 em que optámos pela utilização de linhas com espessura reduzida para o utilizador ter uma maior perceção da dimensão total dos dados que estão a ser representados. O utilizador pode fazer a seleção do período histórico que quer observar (Figura 5.39) e também pode selecionar as categorias do *Modal Analysis*. De forma a facilitar a perceção das palavras que cada variação contém, o utilizador ao fazer mouseover sobre um variações. Ao interagir com cada variação, a linha ganhará uma maior a espessura e aparecerá a informação sobre o título, autor e ano do texto selecionado. Como último elemento interativo do artefacto foi adicionada uma barra de procura, semelhante à visualização anterior, para que seja possível observar um só autor ou um só texto (Figura 5.40). Por fim foi adicionado um texto explicativo do que os utilizadores podem fazer e o que estão a ver na presente visualização. Mais imagens da visualização no Apêndice D.

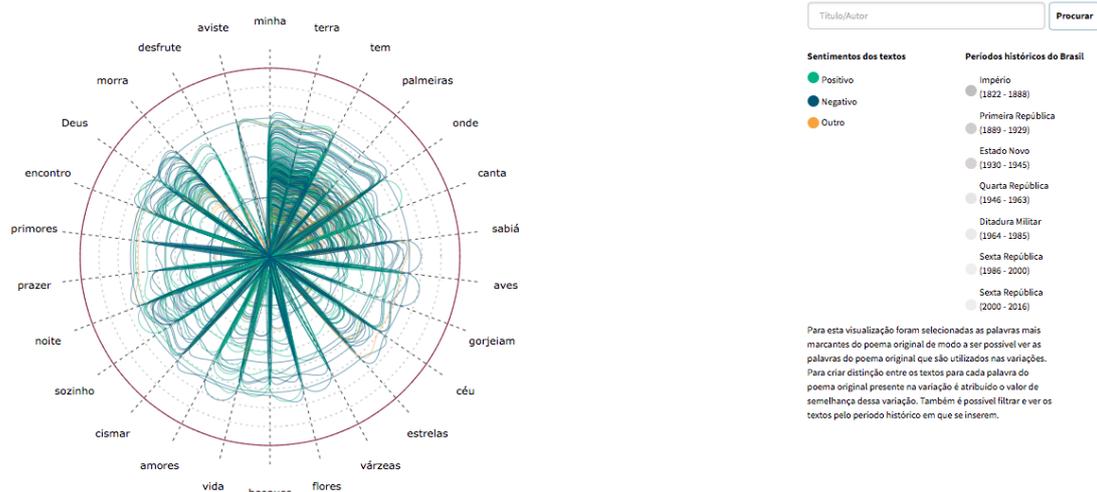


Figura 5.38 Versão final do artefacto “O Poema ... presente nos textos”.

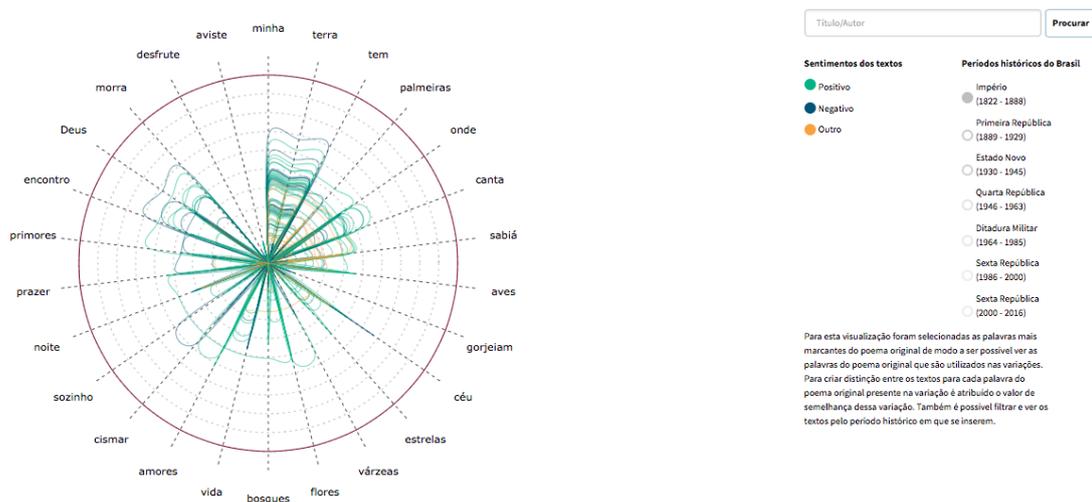


Figura 5.39 Representação das variações correspondentes ao período Império.

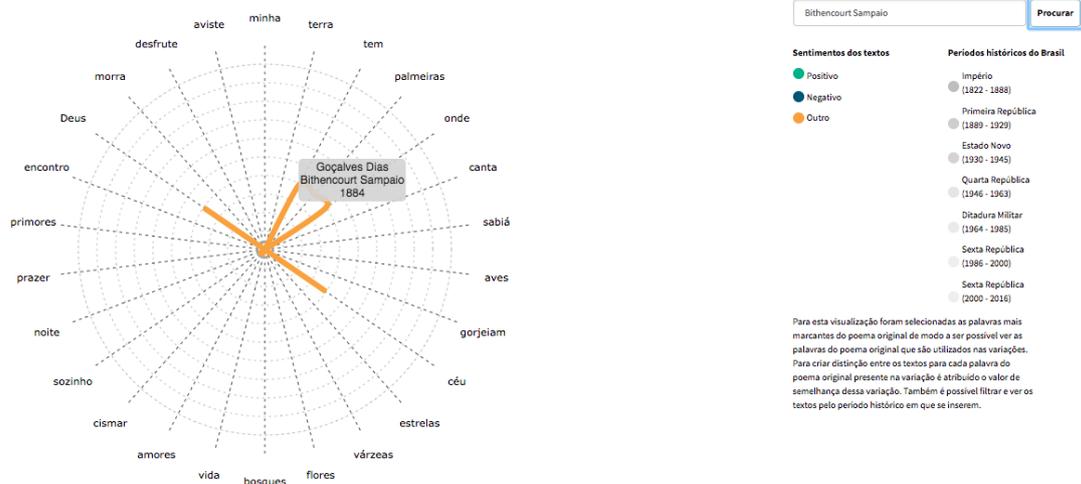


Figura 5.40 Realce do resultado da procura

O último artefacto a ser desenvolvido foi o *Bubble Chart*. Esta visualização, embora visualmente mais simples, foi a mais complexa a ser desenvolvida pois foi necessário gerar os dados para a mesma bem como o seleccionar modelo de *Bubble Chart* que seria mais adequado ao que queríamos mostrar. Foram necessárias tomar várias decisões de quantas palavras iriam ser representadas, e se seria visível alguma alteração considerável de palavras entre os períodos. Esta visualização mostra a presença de 223 palavras diferentes, que estão presentes nas variações e na “*Canção do Exílio*”. Estas palavras foram seleccionadas de uma lista que continham mais de 10800 palavras. De forma ser possível representar a variedade de presenças limitamos as palavras para o mínimo de 9 presenças no total dos 500 textos, que é o valor mínimo encontrado na “*Canção do Exílio*” com a palavra “*aviste*”. Após a limitação do número de presença mínimo que as palavras teriam que ter, procurámos eliminar, independentemente do seu valor, todas as palavras que fossem pronomes, conjunções, interjeições, alguns advérbios e também algumas preposições. Das restantes palavras foram também eliminadas as que teriam menos importância geral para a construção da visualização. Na Figura 5.41 mostramos a primeira representação, apenas com as palavras do poema de Gonçalves Dias. Esta versão serviu para perceber o como se teria que conjugar a diversidade de tamanhos.

Foi também explorada outra alternativa (Figura 5.42) na qual para as palavras do poema de Gonçalves Dias, foi criada uma matriz de pares de palavras que ocorrem no poema e foi feita essa contagem para todas as variações. Foi então representado o número de vezes que uma palavra está presente, quanto mais escura e maior fosse a espessura da linha maior era o valor dessa presença. A direção da ligação entre conjugação das palavras é representada pela linha, sendo que a linha preenchida representaria a primeira conjugação (ex.: *minha terra*) enquanto que a linha tracejada faria a representação da segunda conjugação possível para o mesmo conjunto de palavras (ex.: *terra minha*).

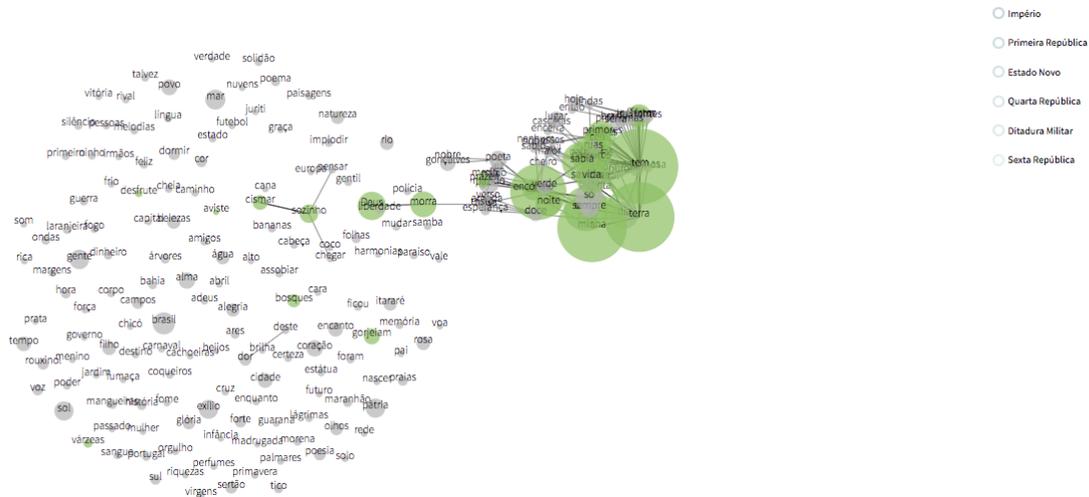


Figura 5.43 Terceiro modelo de representação da visualização “O Poema ...em palavras” já com as palavras das variações.

O modelo já mais próximo ao finalizado é apresentado na Figura 5.44. Uma vez que criava muito ruído foi criado um *tooltip* para o utilizador, ao fazer *mouseover*, saber que palavra é que está a ver. Nesta fase ainda foi considerado apresentar as ligações de cada palavra por período mas essa opção acabou por não ser implementada. Neste artefacto, as palavras a cor de rosa pertencem à “Canção do Exílio” enquanto que a cinzento estão as palavras das variações.

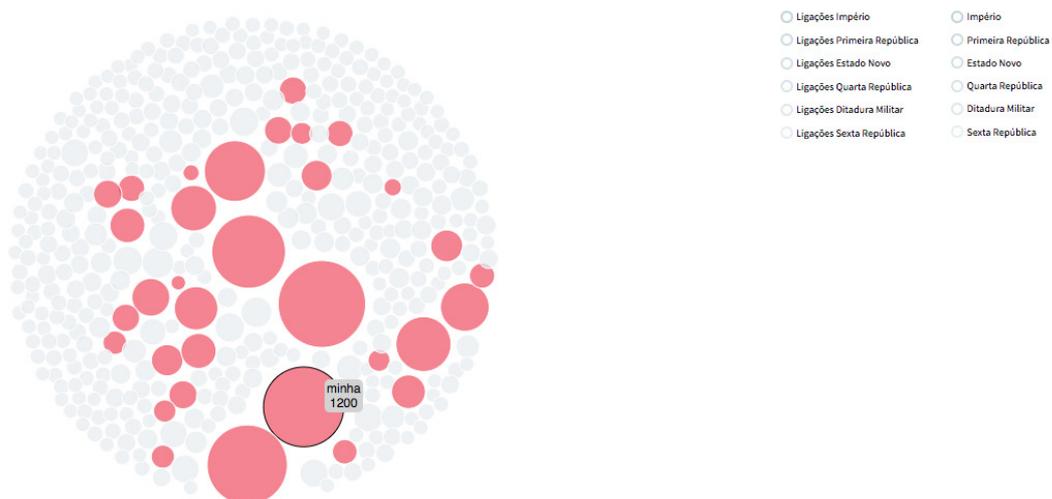


Figura 5.44 Quarto modelo de representação da visualização “O Poema ...em palavras”.

Na Figura 5.45 apresentamos o modelo final da visualização “O Poema ...em palavras”. Devido à diversidade de presenças na nossa lista de palavras (Apêndice A), o valor máximo é 1453 e o valor mínimo 9. Para controlar tamanho dos círculos foi criada uma forma de mapear esses tamanhos para que todos os círculos fossem representados e para que fosse visível a sua diferença sem que os valores maiores fossem demasiado grandes em relação às palavras de menor valor. Ainda para reforçar essa diferença, nas palavras com presença superior a 1000 foi atribuído menos valor de transparência realçando as palavras que se encontram nestas condições. Ao utilizador é permitido interagir com a visualização alternando entre os períodos históricos brasileiros (Figura 5.46), ao fazê-lo, para cada período é representado o mesmo gráfico com as palavras presentes em cada época. Quando uma palavra, que estava presente no conjunto total de anos, não tem presença naquela época não é desenhada, deixando o espaço da mesma em vazio. Nos vários períodos é possível ver essas mudanças, das palavras que estavam no total de todos os anos e deixam de aparecer em determinada época. O utilizador pode ainda fazer mouseover sobre qualquer círculo para ver que palavra é e qual é o seu valor (Figura 5.46). Mais imagens desta visualização estão disponíveis no Apêndice E.

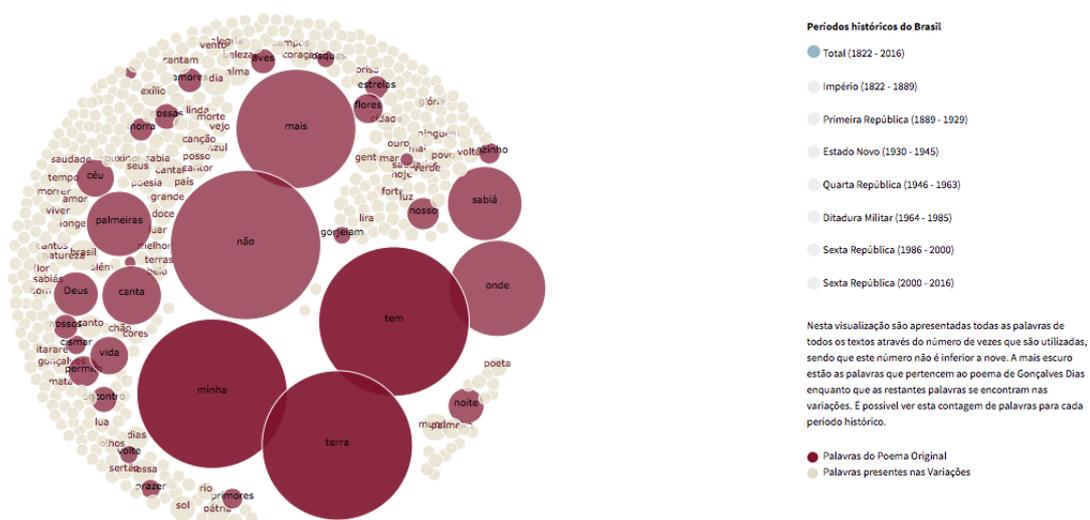


Figura 5.45 Representação final da visualização “O Poema ...em palavras”.

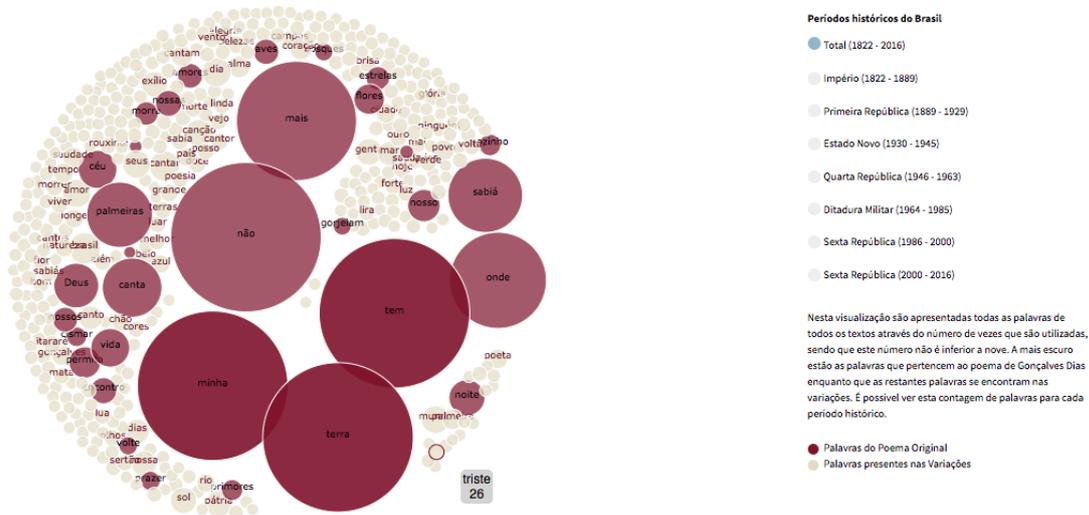


Figura 5.46 Representação da funcionalidade de tooltip com mouseover sobre a palavra triste.

5.4.4 FILTROS

Uma forma de despertar o interesse aproximando os utilizadores dos artefactos é tornar estes interactivos, dando a possibilidade ao observador de manipular os dados e contextualizá-los com temas com os quais estão familiarizados. Para permitir esta contextualização dos dados, foi feita a seleção de vários momentos da história do Brasil e do Mundo.

A primeira pesquisa efectuada incidiu sobre datas históricas relevantes para a história do Brasil, tendo sido reunida uma lista de acontecimentos que poderiam estar mencionados nas variações. Feita essa comparação, verificou-se que os acontecimentos não estavam explicitamente mencionados mas existiam muitas referências ao período político brasileiro. Assim, foi feita a recolha de períodos governativos, começando pelo Império que era o período governativo na altura em que Gonçalves Dias escreveu “*Canção do Exílio*”. O Império esteve em vigor desde 1822 até 1889. No período após a publicação do poema e até ser declarado o fim

do Império ocorreram eventos importantes como a promulgação da lei Eusébio de Queiroz, que proibiu o tráfico de pessoas negras para o Brasil em 1848, e a lei Áurea, que em 1888 aboliu a escravatura no Brasil (Virtuous, 2009-2018). Com o término deste período tem início a Primeira República, período também conhecido como República Velha, entre 1889 até 1930. Esta época ficou marcada pela greve dos operários, em São Paulo, em 1917, e pelas revoltas das armadas (Virtuous, 2009-2018). O período seguinte ficou conhecido por Estado Novo ou Era Vargas e esteve em vigor de 1930 a 1945. Durante este período foi promulgada a segunda constituição da república em 1934, e foram criadas a Consolidação das Leis do Trabalho em 1943 (Virtuous 2009-2018). A Quarta República tem início com a promulgação de uma nova constituição em 1946 e termina com um golpe militar em 1964. Desse golpe militar surge um novo período governativo na história do Brasil, a Ditadura Militar que esteve no poder até 1985 (Virtuous 2009-2018). Após o término do período ditatorial em 1985 tem início a Sexta República, também conhecida como Nova República, que está em vigor até aos dias de hoje. Este período tem sido marcado pelos mais diversos acontecimentos que têm vindo a abalar a economia e os governos com sucessivas suspeitas de corrupção.

Para além dos períodos políticos era importante permitir uma contextualização dos textos com os movimentos artísticos literários mais importantes. Pela influência no povo brasileiro, os movimentos escolhidos foram Modernismo, Bossa Nova e Tropicália. O Modernismo foi um movimento com diversas fases sendo que no geral esteve muito activo de 1922 até 1960. Alguns dos autores de variações do poema de Gonçalves Dias, como Oswald de Andrade, Carlos Drummond de Andrade, Manuel Bandeira, Cassiano Ricardo, Murilo Mendes, Ferreira Gullar, Guilherme de Almeida e Ribeiro Couto são reconhecidos no movimento modernista literário (InfoEscola, 2006-2018). A Tropicália foi um movimento que decorreu entre 1967 e 1968. O seu principal objetivo foi agitar o mundo da música da época. Dos seus participantes há registo de Torquato Neto, Gilberto Gil, Rita Lee e Capinam terem escrito variações do poema de Gonçalves Dias (Oliveira, 2007). O movimento Bossa Nova, sendo um movimento mais direcionado à música, foi considerado pelo aspecto da composição das letras de músicas pois compositores como, Chico Buarque e Torquato Neto tiveram como inspiração a *"Canção do Exílio"* para a escrita de letras de músicas. Este movimento decorreu entre 1958 e 1966 e em simultâneo com o Modernismo tendo tido o seu impacto e a sua própria expressão (TodaMatéria, 2017).

A escolha dos eventos da história mundial foi feita inicialmente com a recolha de um conjunto de eventos. Posteriormente, foram filtrados através da análise da influência e impacto que estes tiveram no Brasil e se era possível encontrar referências aos mesmos nas variações. Após essa análise, optou-se pela representação de três grandes eventos da história mundial: a Primeira e Segunda Guerra Mundial e a Grande Depressão.

A participação do Brasil na Primeira Guerra Mundial acontece após serem abatidos navios brasileiros pelos Alemães. A participação do Brasil envolveu o fornecimento de matéria-prima em falta aos países aliados mas também envolveu a participação das tropas na frente ocidental. Com esta intervenção, o Brasil viu a sua indústria desenvolver-se uma vez que as indústrias europeias ficaram debilitadas (Neves, n.d). O Brasil participa na Segunda Guerra Mundial a partir de 1944, após navios alemães bombardearem cinco navios brasileiros de mercadorias. Este ataque acontece após a cessação das relações entre o Brasil e a Alemanha. Em resposta ao ataque são enviados 25 mil soldados que combateram ao lado dos países aliados contra a Alemanha (Silva, n.d).

A Grande Depressão, ou Crise de 1929, teve início nos Estados Unidos e acabou por se alastrar a outros países do mundo. Um dos países afetados foi o Brasil, cujas exportações de café estavam bastante dependentes dos Estados Unidos. Com a crise, o país deixou de exportar café para o seu principal comprador, o que levou a uma acumulação de café nos armazéns. Era necessário encontrar uma solução para que o preço do produto não baixasse e a solução encontrada pelo governo da época foi a compra dos excedentes para ser queimado. A compra do café pelo governo, não fez abrandar a produção do produto, e as vendas continuaram a cair e a dívida do governo a aumentar. Esta situação agravou-se de tal forma que levou à revolução de 1930 (Vallone, 2009).

Com a selecção destes eventos é esperado que os utilizadores sejam capazes de, facilmente, localizar temporalmente as variações, sentindo-se instigados a interagir com os artefactos desenvolvidos e compreendendo a influência de determinados acontecimentos nas variações da *“Canção do Exílio”*.

5.5 PLATAFORMA FINAL

O desenvolvimento de uma plataforma surge da intenção de querer aproximar este projeto o mais possível à plataforma idealizada por Joshua Enslin. A partir desta ideia a plataforma web foi desenvolvida como meio de agrupar as visualizações entre si.

Inicialmente, foram desenvolvidos vários estudos em papel da estrutura geral da plataforma, sobre a forma como os artefactos iriam ser mostrados, e como iríamos colocar artefactos e filtros sem que os últimos retirassem o espaço necessário às visualizações. Também era importante decidir que informação extra iria estar presente na página de forma a que todo o projeto fizesse sentido.

Na Figura 5.47 está representado o último esboço em papel do planeamento da plataforma. Aqui existiria um menu que estaria escondido aparecendo apenas quando o utilizador assim escolhesse. O espaço para as visualizações seria central e os filtros apareceriam no fundo da página.



Figura 5.47 Esboço em papel da plataforma.

O passo seguinte foi desenvolver os esboços desenhados em papel para a web de forma a permitir explorar o espaço disponível, escolher fontes e cores a utilizar (Apêndice F). Esta exploração também permitiu ter novas ideias para o desenvolvimento da plataforma tendo ficado decidido nesta etapa que a página seria uma única página com scroll down. Com este modelo de página optámos por utilizar um menu superior de forma a criar menos perturbações nos elementos centrais da plataforma. A página seria composta por 3 secções diferentes, a primeira secção seria sobre o poema para permitir aos utilizadores conhecerem mais sobre o poema, a segunda seriam as visualizações e a terceira seriam questões que os utilizadores poderiam ter e que seriam respondidas naquela secção (Figura 5.48).

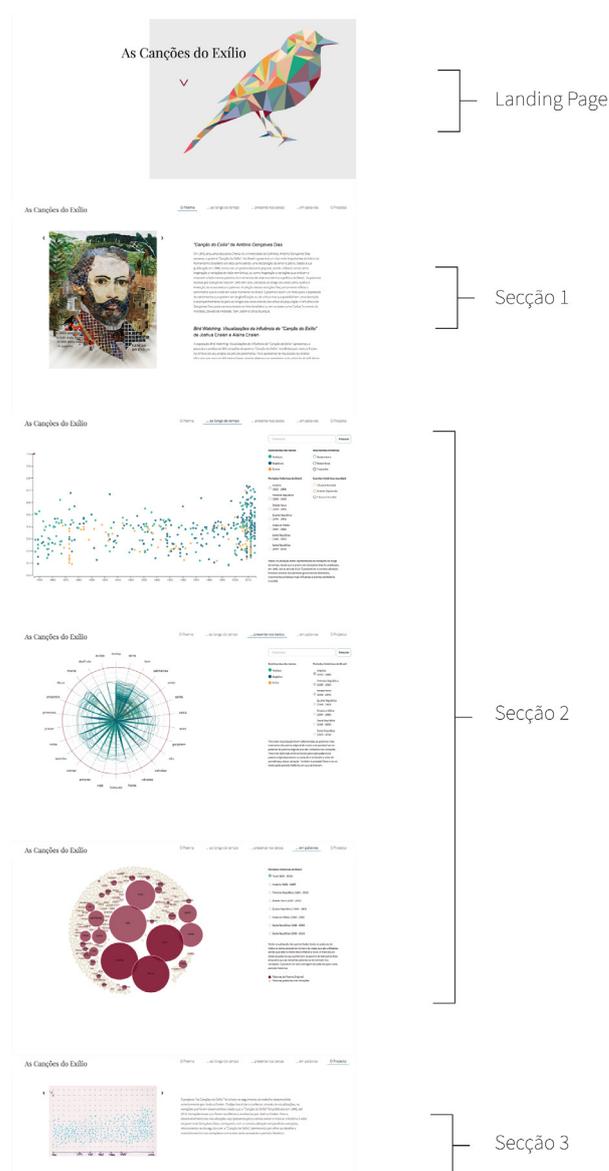


Figura 5.48 Estrutura da plataforma.

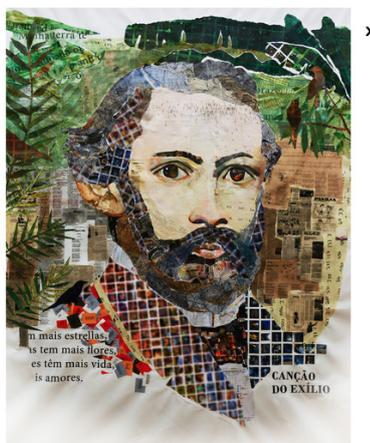
A partir desta ideia deu-se início ao desenvolvimento da plataforma, sendo que para dar destaque às visualizações optámos por minimizar a utilização de cor nos elementos não integrantes dos artefactos, sendo apenas utilizada na página inicial e para dar pequenos destaques necessários como aos títulos das páginas, optámos por fundo branco na página e no menu.

A página inicial da plataforma (Figura 5.49) dá ao utilizador a opção de seguir para a página seguinte ou de escolher o idioma que prefere, sendo esta escolha entre português e inglês (Apêndice G) de forma a fazer chegar o projeto a um público mais abrangente.



Figura 5.49 Página inicial da plataforma final.

Ao entrar no site, (Figura 5.50) é dada uma breve explicação sobre a “*Canção do Exílio*” e também sobre o projeto desenvolvido por Joshua Enslin. Para todo o site, como é possível observar na imagem, optámos pela utilização de um menu superior que facilita a transição entre páginas sem ter que percorrer o site todo.



"Canção do Exílio" de Antônio Gonçalves Dias

Em 1843, enquanto estudava Direito na Universidade de Coimbra, Antônio Gonçalves Dias escreveu o poema "Canção do Exílio". No Brasil o poema é um dos mais importantes do início do Romantismo brasileiro e é visto como sendo uma declaração de amor à pátria. Desde a sua publicação em 1846, tornou-se um poema bastante popular, sendo utilizado tanto como inspiração a variações de visão romântica, ou como inspiração a variações que criticam e mostram o lado menos positivo de momentos de crise económica e política do Brasil. As palavras escritas por Gonçalves Dias em 1843 têm sido utilizadas ao longo dos anos como auxílio à invenção de novos textos e poemas. A criação destas variações frequentemente reflete o sentimento que é vivido em cada momento no Brasil. O poema é assim um meio para a expressão de sentimentos que podem ser de glorificação ou de crítica mas que possibilitam uma descrição e acompanhamento do país ao longo dos anos através dos olhos da população. A influência de Gonçalves Dias pode ser encontrada no hino brasileiro ou em autores como Carlos Drummond de Andrade, Oswald de Andrade, Tom Jobim e Chico Buarque.

Bird Watching: Visualizações da influência de "Canção do Exílio" de Joshua Enslin e Alaina Enslin

A exposição *Bird Watching: Visualizações da influência de "Canção do Exílio"* apresentou a pesquisa e análise de 500 variações do poema "Canção do Exílio" recolhidas por Joshua Enslin no âmbito do seu projeto de pós-doutoramento. Para apresentar os resultados da análise efetuada aos mais de 500 textos foram criados objetos que permitem a visualização da influência

Figura 5.50 Página com explicação sobre o poema de Gonçalves Dias e sobre o projeto de Joshua Enslin.

Durante o planeamento do site concluímos que organizar as visualizações teria impacto na forma como estas seriam interpretadas. Assim, optámos por dispor os artefactos de forma a que o utilizador comece com uma leitura geral dos dados, prosseguindo para visualizações mais detalhadas não só sobre o poema original mas também sobre as variações. Assim a primeira visualização, a que foi dado o título "O Poema ...ao longo do tempo" (Figura 5.51) faz a contextualização histórica inicial e temporal, a visualização seguinte "O Poema ...presente nos textos" (Figura 5.52) mostra a relação que cada poema tem com a "Canção do Exílio", e por fim, a visualização "O Poema ...em palavras" (Figura 5.53) apresenta o conteúdo, ou seja as palavras, de todos os textos, poema original e variações, e como a presença das palavras se modifica ao longo do tempo.

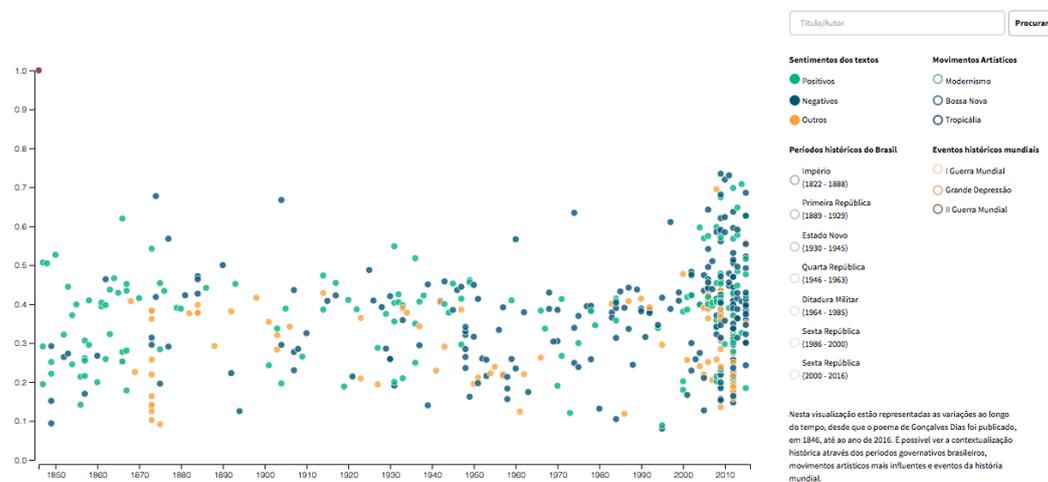


Figura 5.51 Página final da visualização "O Poema ...ao longo do tempo".

As Canções do Exílio

O Poema ... ao longo do tempo ... presente nos textos ... em palavras O Projecto

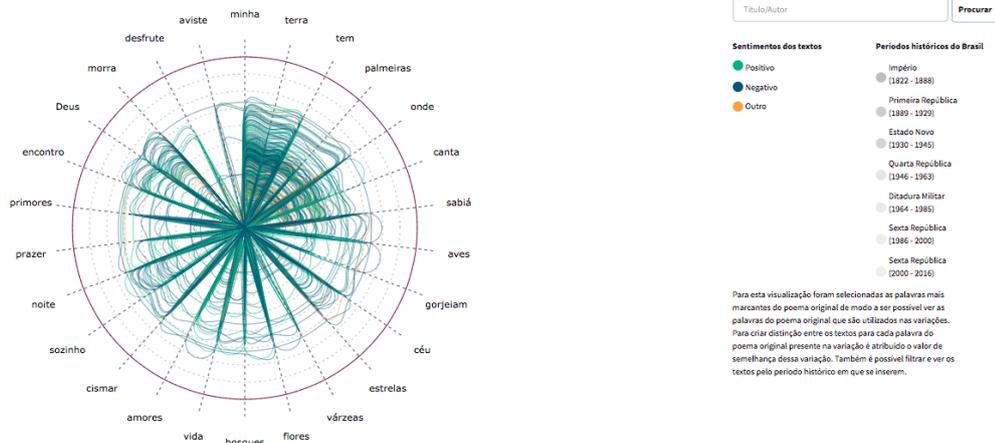


Figura 5.52 Página final da visualização “O Poema ...presente nos textos”.

As Canções do Exílio

O Poema ... ao longo do tempo ... presente nos textos ... em palavras O Projecto

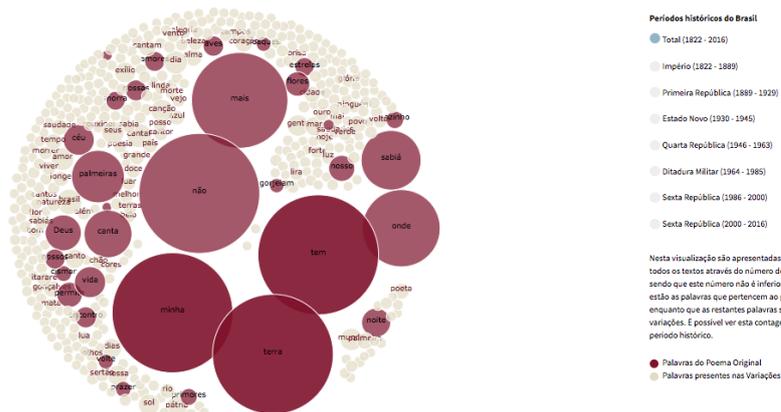
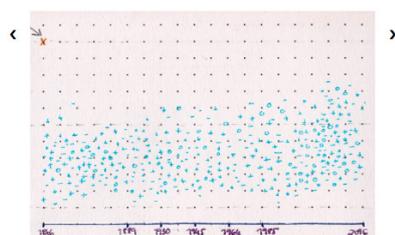


Figura 5.53 Página final da visualização “O Poema ...em palavras”.

A última secção da plataforma é a página que faz um enquadramento ao utilizador do propósito da plataforma, explicando também a organização escolhida para apresentar as visualizações (Figura 5.54).

As Canções do Exílio

O Poema ... ao longo do tempo ... presente nos textos ... em palavras O Projecto



O projecto “As Canções do Exílio” foi criado no seguimento do trabalho desenvolvido anteriormente por Joshua Enslin. O objectivo é dar a conhecer, através de visualizações, as variações que foram desenvolvidas desde que a “Canção do Exílio” foi publicada em 1846, até 2015. Variações essas que foram recolhidas e analisadas por Joshua Enslin. Para o desenvolvimento das visualizações aqui presentes procuramos contar e mostrar a história à volta do poema de Gonçalves Dias, começando com a contextualização temporal das variações, relacionando-as de seguida com a “Canção do Exílio”, terminando por olhar ao detalhe o conteúdo escrito nas variações e como este varia consoante o período histórico.

Figura 5.54 Página descritiva do projeto.

5.6 RESULTADO DA ANÁLISE DAS VISUALIZAÇÕES

Para cada visualização foi possível encontrar tendências e factos que importam ser apresentados. Na primeira visualização, “*O Poema ...através do tempo*”, é possível observar que, com o passar dos anos, a tendência é para que existam mais textos com classificação negativa. Nesta visualização, podemos concluir que no período histórico Império, dos 83 textos pertencentes a este período 45 são textos classificados como *positivos*, 19 são classificados como *negativos* e 19 *outros*. Na Primeira República é possível ver uma pequena inversão pois dos 44 textos 20 variações, estão classificadas como *negativas*, 16 são *positivas*, estão classificadas como *outras* 22 textos. No período Estado Novo, ou Era Vargas, a tendência da classificação das variações positivas volta a aumentar sendo que dos 30 textos, 16 variações são positivas, 8 textos são negativos e 6 textos são classificados como outros. A partir da Quarta República tem início a tendência para a o aumento de variações com classificação negativa. Neste período registámos 40 variações sendo que 27 textos são negativos, 5 variações são positivas e 8 textos, são classificados como outros. No período de Ditadura Militar contabilizamos 36 variações em que 13 dessas variações são positivas, com a classificação negativa temos 21 variações e 2 variações classificadas como outras. Na Sexta República durante o século 20 foram registamos 29 variações, destas 18 são negativas, 5 variações são positivas e 6 variações classificadas como outros. Ainda neste período mas já no século XXI foram recolhidos 238 variações. Destas variações 122 são negativas, 66 textos são positivos e 50 outros. Com esta análise é possível também observar que dos 500 textos analisados 238 variações foram escritas entre 2001 e 2015. Isto pode dever-se a vários fatores como não haver nos anos anteriores os meios de divulgação e propagação dos textos que começaram a existir a partir dos anos 2000.

A visualização seguinte, “*O Poema ...presente nos textos*”, permite-nos observar as tendências de utilização das palavras nas variações. Numa primeira observação, percebemos que as palavras mais utilizadas são: “*minha*”, “*terra*”, “*tem*”, “*onde*”, “*sabiá*”, e “*flores*” sendo que, para cada tipo de classificação, é possível perceber que há palavras menos utilizadas e palavras mais utilizadas. Para as variações com classificação positiva é possível observar, através da sobreposição dos dados são criadas zonas mais carregadas em determinadas palavras, as mais utilizadas (Figura 5.55). O oposto acontece para as palavras menos utilizadas. Concluímos assim que as palavras mais utilizadas são: “*minha*”, “*terra*”, “*tem*”, “*onde*” e “*Deus*”. Entre as palavras menos utilizadas podemos verificar que são: “*várzeas*”, “*cismar*”, “*desfrute*” e “*aviste*”. Nas variações com a classificação negativa podemos concluir que as palavras mais utilizadas são: “*minha*”, “*terra*”, “*tem*”, “*palmeiras*”, “*sabiá*”, “*céu*”, “*noite*”, “*Deus*” e “*morra*”. As palavras visivelmente menos utilizadas são: “*aviste*”, “*desfrute*”, “*várzeas*” e “*gorjeiam*”. Neste grupo classificativo do *Modal Analysis* podemos verificar que existe maior utilização das palavras do poema original. As variações classificadas como outras contêm mais as palavras “*minha*”, “*tem*”, “*onde*”

e “Deus” enquanto que as palavras “aviste”, “desfrute”, “várzeas”, “gorjeiam”, “cismar”, “prazer”, “sozinho” e “bosques” têm pouca ou nenhuma presença no conjunto total das variações assim classificadas. É possível ainda concluir que palavras são mais utilizadas por época uma vez que essa observação engloba as três categorias do *Modal Analysis*. No primeiro período histórico, o Império, concluímos que as palavras mais utilizadas são: “minha”, “tem”, “onde”, “canta”, “sabiá”, “flores”, “amores” e “Deus” enquanto que as palavras menos utilizadas são: “aviste”, “desfrute” e “gorjeiam”. No período da Primeira República as palavras mais utilizadas são: “minha”, “tem”, “onde” e “sabiá” enquanto que as menos utilizadas são “gorjeiam”, “várzeas”, “aviste”, “desfrute”, “sozinho”, “prazer” e “encontro”. No período do Estado Novo concluímos que as palavras mais utilizadas nas variações são: “minha”, “tem”, “onde” e “flores” e as palavras menos utilizadas são: “terra”, “gorjeiam”, “várzeas”, “bosques”, “cismar”, “prazer”, “primores”, “encontro”, “desfrute” e “aviste”. No período histórico seguinte, a Quarta República, as palavras que aparecem mais vezes nas variações são: “minha”, “tem”, “palmeiras”, “onde”, “canta” e “sabiá”. Já as palavras que têm menos frequência, ou nenhuma, nas variações são: “gorjeiam”, “bosques”, “prazer” e “aviste”. No período de Ditadura Militar apesar de encontrarmos dados mais dispersos podemos concluir que algumas das palavras mais frequentes nas variações são: “minha”, “tem”, “onde” e “sabiá”, enquanto que as palavras menos frequentes deste período são “várzeas”, “sozinho”, “prazer” e “desfrute”. Na Sexta República, durante o século XX, as palavras mais utilizadas são: “minha”, “tem” e “onde” enquanto que as com menor frequência são: “várzeas” e “desfrute”. No mesmo período histórico mas já no século XXI as palavras mais frequentemente utilizadas nas variações são: “minha”, “terra”, “tem” e “onde”, as palavras menos frequente neste período são: “gorjeiam”, “desfrute” e “aviste”.

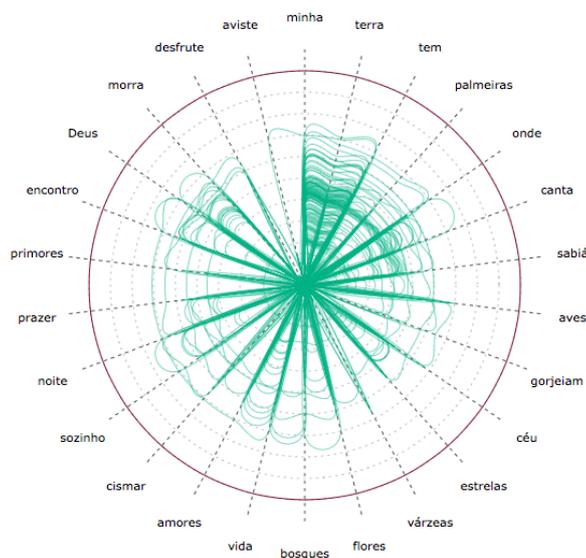


Figura 5.55 Representação das variações positivas.

No último artefacto “*O Poema ...em palavras*” após analisar os dados que estão a ser representados percebemos que mais do que a frequência de palavras podemos visualizar picos de utilização de determinadas palavras nos vários períodos históricos apresentados. Um dos fenómenos que nos foi possível observar é a utilização da palavra “*tanques*” que é mencionada 9 vezes. Estas menções acontecem apenas no período da Quarta República. Outro fenómeno semelhante ocorre com a palavra “*assobiar*”, que é mencionada apenas no período de Ditadura Militar 16 vezes. Na Sexta República, durante o século XX, mencionada 42 vezes podemos encontrar a palavra “*itararé*” que também apenas é mencionada nesse período. Outro fenómeno que existe é a utilização de palavras que não foram referenciadas durante vários períodos. Isto acontece com a palavra “*problemas*” que aparece no Império e só volta a reaparecer nas variações durante o período da Sexta República já no século XXI. Depois também existem as palavras que, não pertencendo ao poema de Gonçalves Dias e não tendo um valor presencial elevado, são continuamente utilizadas nas variações ao longo dos anos, como a palavra “*bom*” e a palavra “*gente*”.

Os dados apresentados são apenas uma pequena parte do que os dados em si podem mostrar. Com cada visualização tentámos ir de uma leitura mais geral das visualizações para os detalhes das mesmas, iniciando com a sua posição temporal criando a contextualização histórica, passando para a ligação que cada variação tem com a “*Canção do Exílio*” através das palavras do poema e terminando com a representação das palavras das variações que também nos contam como é que cada momento da história influenciou os autores das variações do poema de Gonçalves Dias.

6 CONCLUSÃO E TRABALHO FUTURO

6 CONCLUSÃO E TRABALHO FUTURO

Esta dissertação tem como objetivo mostrar a influência do poema de Gonçalves Dias por meio da Visualização de Informação. O desenvolvimento deste projeto resultou numa plataforma que reúne três artefactos. As visualizações desenvolvidas são uma das muitas explorações possíveis dos dados, são o ponto de partida para criação de novos artefactos e para a criação de novos dados.

Na área de visualização de texto, os documentos são vistos como objetos de relacionamentos complexos através das palavras que contém (Lima, 2011). Os documentos utilizados nesta dissertação mostram que através do tipo de palavras, e dos sentimentos textuais conseguimos perceber: o ambiente que não está descrito nos próprios textos, conseguimos perceber se o autor de uma variação tem influências de movimentos artísticos, ou se pelo contrário, apenas sente orgulho pelo seu país. O tipo de informação que se pode retirar destes documentos é vasta e permite conhecer não só a influência da “*Canção do Exílio*” mas também a história do Brasil. Os artefactos desenvolvidos no âmbito desta dissertação apresentam não só a influência de Gonçalves Dias mas também fazem uma contextualização histórica e cultural para as próprias variações, permitindo compreender melhor as mesmas. As visualizações foram organizadas de forma a que exista um fio condutor entre cada, conduzindo os observadores de uma leitura geral e distante dos dados para uma leitura próxima e detalhada. Apesar da existência deste fio condutor, as visualizações são objetos que funcionam de forma independente não estando dependentes umas das outras para que os dados que representam sejam compreendidos.

Desta forma, pensamos que os objetivos desta dissertação foram alcançados tendo sido criado um conjunto de visualizações que mostram e que dão a conhecer aos utilizadores a influência da “*Canção do Exílio*” e de Gonçalves Dias ao longo de 170 anos.

Futuramente, e pelo tipo de dados que contém, este projeto poderá continuar em desenvolvimento, podendo ser o trabalho desenvolvido até agora ampliado e melhorado. Aos dados existentes poderiam ser adicionados novos dados e poderiam ser procuradas novas formas de exploração dos mesmos recorrendo a diversas ferramentas de análise textual. Através da implementação de uma base de dados os utilizadores poderiam facilmente ver os textos das variações e poderia ser dada a permissão aos utilizadores de submeterem as suas próprias variações. Ao ser implementada esta funcionalidade seria importante também criar um sistema autónomo que atualizasse os dados em tempo real que através da utilização do D3.js seria exequível. Com mais resultados das análises aos textos poderiam ser desenvolvidas mais visualizações, e tornar as existentes mais dinâmicas. A plataforma poderia ser flexibilizada para englobar e apresentar os diversos projetos desenvolvidos e a desenvolver em torno da “*Canção do Exílio*” podendo fazer com que os diferentes públicos encontrassem tudo relacionado com o tema num só espaço.

BIBLIOGRAFIA

BIBLIOGRAFIA

- Accurat. (2012). *The real Montalbano*. Consultado a 17 de Janeiro de 2018, em <https://www.flickr.com/photos/accurat/8250026842/in/album-72157632185046466/>
- AnyChart. (2015). *First Ever Bar Charts Created in 14th Century [Interesting Fact]*. Consultado a 17 de Janeiro de 2018, em <https://medium.com/@AnyChart/first-ever-bar-charts-created-in-14th-century-interesting-fact-4fb9ae576509>
- Anthony, L. (2018). *AntConc* (Version 3.5.7) [Computer Software]. Tokyo, Japan: Waseda University. Available em <http://www.laurenceanthony.net/software>
- Ao Pé da Letra. (2011). *Espaço ao Pé da Letra: Romantismo e Modernismo no Brasil*. Consultado a 3 de Outubro de 2017, em <http://espacoaope-daletra.blogspot.pt/2011/07/romantismo-e-modernismo-no-brasil.html>
- Becker, A., Burch, M., & Gasparro, A. (2014). *Axis of Additives - WSJ.com*. Consultado a 17 de Janeiro de 2018, em <http://graphics.wsj.com/food-additives-ingredients/>
- Behrens, C. (2008). *The Form of Facts and Figures*. Consultado em <http://datavis.kducharme.com/wp-content/uploads/2011/08/design-patterns-interactive-infovis-Behrens.pdf>
- Bertin, J. (1999). *Sémiologie Graphique: Les Diagrammes - Les Réseaux, Les Cartes*. École des Hautes Études en Sciences Sociales.
- Card, S. K., Mackinlay, J., & Shneiderman, B. (1999). *Information Visualization: Using Vision to Think*. San Francisco, CA: Morgan Kaufman.
- Chen, M., Floridi, L., & Borgo, R. (2013). *What is Visualization Really for?*, 11. Consultado em <https://arxiv.org/pdf/1305.5670.pdf>
- de Oliveira, A., & Coelho, F. (2007). *Movimento – Tropicália*. Consultado a 27 de Julho de 2018, em <http://tropicalia.com.br/identifisignificados/movimento>
- Diana, D. (2017). *Canção do Exílio, de Gonçalves Dias - Toda Matéria*. Consultado a 15 de Janeiro 2018, em <https://www.todamateria.com.br/cancao-do-exilio-de-goncalves-dias/>
- Dias, A. G. (1846). “*Canção do Exílio*.” *Primeiros Cantos: Poesias*. Consultado a 18 de Janeiro de 2018, em https://books.google.pt/books?id=ma5O-AQAAIAAJ&pg=PA3&redir_esc=y#v=onepage&q&f=false

- Dias, A. G., Bueno, A., & Bandeira, M. (1998). *Poesia e prosa completas*. Editora Nova Aguilar. Consultado em <http://www.infoescola.com/livros/cancao-do-exilio/>
- Enslén, A., & Enslén, J. (2016). *Bird Watching — Alaina Enslén*. Consultado a 22 de Janeiro de 2018, em <https://www.alainaenslen.com/bird-watching/>
- Enslén, J., & Enslén, A. (2017). *Bird-Watching: Visualizações da Influência de “Canção do Exílio” | Materialidades da Literatura*. Consultado a 23 de Janeiro de 2018, em <https://matlit.wordpress.com/2017/07/12/bird-watching-visualizacoes-da-influencia-de-cancao-do-exilio/>
- Enslén, J., & Enslén, A. (2016). *Visita à exposição “Bird-Watching: Uma História Intertextual de “Canção do Exílio” - YouTube*. Faculdade de Letras da Universidade de Coimbra. Consultado em <https://www.youtube.com/watch?v=DcBlpBysnwE>
- Enslén, J., & Enslén, A. (2016). *Museu da Ciência da Universidade de Coimbra - Fotos*. Consultado a 23 de Janeiro de 2018, em <https://www.facebook.com/museudaciencia/photos/a.1207565232596376.1073741916.188225504530359/1207566642596235/?type=3&theater>
- Enslén, J., & Enslén, A. (2016). *Bird-Watching: conferência e exposição | Materialidades da Literatura*. Consultado a 23 de Dezembro de 2017, em <https://matlit.wordpress.com/2016/05/07/bird-watching-conferencia-e-exposicao/>
- Escritas.org. (1968). *Marginália II - Poema de Torquato Neto*. Consultado a 3 de Outubro de 2017, em <https://www.escritas.org/pt/t/13128/marginalia-ii>
- Faria, C. (n.d.). *Modernismo na literatura - história e movimentos literários*. Consultado a 27 de Julho de 2018, em <https://www.infoescola.com/literatura/modernismo/>
- Friendly, M. (2008). *A Brief History of Data Visualization. Handbook of Data Visualization*, 15–56 (1–43). https://doi.org/10.1007/978-3-540-33037-0_2
- Fry, B. (2004). *Computational Information Design*. Massachusetts Institute of Technology.
- Fry, B. (2008). *Visualizing Data - Exploring and Explaining Data with Processing Environment*. Sebastopol, CA: O'Reilly Media, Inc.

- Fry, B. J. (2004). *Computational Information Design*. SciencesNew York, Ph. D(May 1997), 170. <https://doi.org/10.1111/j.1468-3083.2010.03837.x>
- Gonzalez, J. (2012). *Maybe the best infographic ever*. - Jose Gonzalez. Consultado a 17 de Janeiro de 2018, em <http://www.jose-gonzalez.org/maybe-the-best-infographic-ever/>
- Gould, S. J. (1998). *Time's Arrow, Time's Cycle: Myth and Metaphor in the Discovery of Geological Time*. Cambridge, MA: Harvard University Press.
- Gruver, A., & Dutton, J. A. (n.d.). *Graduated and Proportional Symbol Maps | GEOG 486: Cartography and Visualization*. Consultado a 17 de Janeiro de 2018, em <https://www.e-education.psu.edu/geog486/node/1869>
- Hearst, M. (2009). *Information Visualization for Text Analysis (Ch 11) | Search User Interfaces | Marti Hearst | Cambridge University Press 2009*. Consultado a 2 de Janeiro de 2018, em http://searchuserinterfaces.com/book/sui_ch11_text_analysis_visualization.html
- Helena, M., & Telles, N. (2011). *No rastro da memória (Santa Rita do Sapucaí): As várias Canções do Exílio - 2 - Casimiro de Abreu*. Consultado a 3 de Outubro de 2017, em <http://ninitelles.blogspot.pt/2011/08/as-varias-cancoes-do-exilio-2-casimiro.html>
- Hennig, B., & Dorling, D. (2016). *The EU referendum. Political Insight*, 7(3), 20–21. Consultado em <http://www.viewsoftheworld.net/?p=4848>
- Hubacek, G. (2011). *Transparency: Are the Richest Americans Also the Best Educated? America's Richest Counties and Best Educated Counties*. Consultado a 17 de Janeiro de 2018, em <https://www.good.is/infographics/america-s-richest-counties-and-best-educated-counties>
- ims25. (2008). *Nightingale's "Coxcombs" | Understanding Uncertainty*. Consultado a 17 de Janeiro de 2018, em <https://understandinguncertainty.org/coxcombs>
- Landsteiner, N. (2013). *Minard: Napoleon's Retreat Em Moscow (Russian Campaign 1812-1813)*. Consultado a 17 de Janeiro de 2018, em <http://www.masswerk.at/minard/>
- Lima, M. (2011). *Visual Complexity - Mapping Patterns of Information*. Nova Iorque: Princeton Architectural Press.
- Lítera. (2015). *Lítera Brasil | Romantismo*. Consultado a 23 de Janeiro de 2018, em <https://www.literabr.com.br/copia-arcadismo>

- Mackinlay, J. (1986). *Automating the design of graphical presentations of relational information*. ACM Transactions on Graphics, 5(2), 110–141. <https://doi.org/10.1145/22949.22950>
- Mason, B. (2017). *The Underappreciated Man Behind the “Best Graphic Ever Produced.”* Consultado a 17 de Janeiro de 2018, em <https://news.nationalgeographic.com/2017/03/charles-minard-cartography-info-graphics-history/>
- Meirelles, I. (2013). *Design for Information*. Massachusetts: Rockport Publishers.
- Mislove, A. (n.d.). *Mapping the World, one Tweet at a Time*. Consultado a 17 de Janeiro de 2018, em <https://mislove.org/tweetmap/>
- Morfeus. (n.d.). *Gonçalves Dias – Canção do Exílio | Genius*. Consultado a 18 de Janeiro de 2018, em <https://genius.com/Goncalves-dias-cancao-do-exilio-annotated>
- Müller, B. (n.d.). *esono.com - Poetry on the Road 2006*. Consultado a 7 de Janeiro de 2018, em <http://www.esono.com/boris/projects/poetry06/>
- Murray, S. (n.d.). *Interactive Data Visualisation for the Web*.
- Paley, W. B. (2002). *TextArc: Still Screen Images | Alice’s Adventures in Wonderland*. Consultado a 7 de Janeiro de 2018, em <http://www.textarc.org/Stills.html>
- Palsky, G. (1996). *Des chiffres et des cartes. La cartographie quantitative au XIXe siècle. Mémoire de la section de géographie (Vol. 19). Ministère de l’enseignement supérieur et de la recherche, Comité des travaux historiques et scientifiques*. Consultado em <https://commons.wikimedia.org/wiki/File:Minard-carte-viande-1858.png>
- Piccolomini, S. (2015). *Il Corriere della Sera, LaLettura |Freedom in countries on Behance*. Consultado a 17 de Janeiro de 2018, em <https://www.behance.net/gallery/23964523/Il-Corriere-della-Sera-LaLettura-Freedom-in-countries>
- Posavec, S., & Lupi, G. (2015). *Dear Data*. Consultado a 17 de Janeiro de 2018, em <http://www.dear-data.com/theproject/>
- Posavec, S., & MCInerny, G. (n.d.). *(En)tangled Word Bank — Stefanie Posavec*. Consultado a 7 de Janeiro de 2018, em <http://www.stefanieposavec.com/entangled-word-bank/>

- Quarto Creates. (2014). Design Dad a Family Tree for Father's Day - Quarto Creates. Consultado a Janeiro 17, 2018, em <https://www.quartoknows.com/blog/quartocreates/2014/06/09/design-dad-a-family-tree-for-fathers-day/>
- Quarto Creates. (2014). Design Dad a Family Tree for Father's Day - Quarto Creates. Consultado a Janeiro 20, 2018, em <https://www.quartoknows.com/blog/quartocreates/2014/06/09/design-dad-a-family-tree-for-fathers-day/>
- Recantos Poéticos. (2012). Canção do Exílio As Avelãs - Jô Soares. Consultado a Outubro 3, 2017, em <http://recantopoeticolettig3.blogspot.pt/2012/04/cancao-do-exilio-as-avelas-jo-soares.html>
- Ribeca, S. (n.d.). Circle Packing - Learn about this chart and tools to create it. Consultado a Dezembro 2, 2017, em https://datavizcatalogue.com/methods/circle_packing.html
- Ribeca, S. (2015). The Data Visualisation Catalogue, 1–4. Consultado a em <https://datavizcatalogue.com/index.html>
- Robinson, A. H. (1982). *Early Thematic Mapping in the History of Cartography*. Chicago, IL: University of Chicago Press.
- Santana, F. (2017). Daltonismo e UX: a experiência para todos – Coletivo UX. Consultado a February 15, 2018, em <https://coletivoux.com/daltonicos-tambem-sao-usuarios-4f03ca40c30d>
- Schaffors, A., Müller, B., & Pfeffer, F. (2006). Poetry on the road. Consultado a Janeiro 7, 2018, em <http://www.esono.com/boris/projects/poetry06/>
- Shneiderman, B. (1999). The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings 1996 IEEE Symposium on Visual Languages* (pp. 336–343). <https://doi.org/10.1109/VL.1996.545307>
- Silva, D. N. (n.d.). Brasil na Segunda Guerra Mundial - Brasil Escola. Consultado a De Julho de 27, 2018, em <https://brasilecola.uol.com.br/historiab/brasil-segunda-guerra.htm>
- Simoes, A. J. G., & Hidalgo, C. A. (2011). The Economic Complexity Observatory: An Analytical Tool for Understanding the Dynamics of Economic Development. Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence. Consultado a Janeiro 17, 2018, em <https://atlas.media.mit.edu/pt/>

- Small, H. (1998). *Florence Nightingale's Statistical Diagrams. ... Conference Organised by the Florence Nightingale ...*. Consultado em <http://www.florence-nightingale-avenging-angel.co.uk/GraphicsPaper/Graphics.htm>
- Sousa, R. G. (n.d.). *Brasil na Primeira Guerra - Brasil Escola*. Consultado a 27 de Julho de 2018, em <https://brasilecola.uol.com.br/historiab/brasil-primeira-guerra.htm>
- Swain, F., & Jones, D. (n.d.). *Phantom Terrains*. Consultado a 18 de Janeiro de 2018, em <http://phantomterrains.com/>
- TodaMatéria. (2017). *Bossa Nova - Toda Matéria*. Consultado a 27 de Julho de 2018, em <https://www.todamateria.com.br/bossa-nova/>
- Tufte, E. R. (2006). *Beautiful Evidence*. Cheshire, CT: Graphic Press.
- Vallone, G. (2009). *Crise de 1929 atingiu economia e mudou a ordem política no Brasil - 24/10/2009 - Mercado - Folha de S.Paulo*. Consultado a 27 de Julho de 2018, em <https://m.folha.uol.com.br/mercado/2009/10/642391- crise-de-1929-atingiu-economia-e-mudou-a-ordem-politica-no-brasil.shtml>
- Virtuous. (n.d.). *Principais fatos que marcaram a História do Brasil (continuação) - Só História*. Consultado a 27 de Julho de 2018, em <https://www.sohistoria.com.br/ef2/lista/p2.php>
- Ward, M., Grinstein, G., & Keim, D. (2015). *Interactive Data Visualization - Foundations, Techniques, and Applications*. Boca Raton, FL: CRC Press, Taylor & Francis Group.
- Wattenberg, M. (2001). *Shape of Song--Method*. Consultado a 17 de Janeiro de 2018, em http://webcache.googleusercontent.com/search?q=cache:http://turbulence.org/Works/song/method/method.html&gws_rd=cr&dcr=0&ei=e5JjWqf0D4rkUclgpugO
- Wetzel, K. (n.d.). *Circular Treemaps*. Consultado a 20 de Janeiro de 2018, em <http://lip.sourceforge.net/ctreemap.html>
- Wikskär, M. (2015). *The Swedish Emigrants*. Consultado a 17 de Janeiro de 2018, em <https://www.behance.net/gallery/26165895/Infographic-Swedish-emigrants>
- Zakovich, R. (n.d.). *Shakespeare: A Data Visualization*. Consultado a 18 de Janeiro de 2018, em https://public.tableau.com/profile/rody.zakovich#!/vizhome/Shakespeare_/ShakespeareADataVisualization

Zoss, A. (n.d.). *LibGuides: Introduction to Text Analysis: Text Visualization*. Consultado em https://guides.library.duke.edu/text_analysis/text_vis

Minard-carte-viande-1858.png. (n.d.). Consultado a 17 de Janeiro de 2018, em <https://upload.wikimedia.org/wikipedia/commons/1/1c/Minard-carte-viande-1858.png>

Cartogram Maps: Data Visualization with Exaggeration - GIS Geography. (2018). Consultado a 3 de Dezembro de 2017, em <http://gisgeography.com/cartogram-maps/>

Flow Maps: Linking Movement with Direction Lines - GIS Geography. (n.d.). Consultado a 3 de Dezembro de 2017, em <http://gisgeography.com/flow-maps/>

Gonçalves Dias e a Canção do Exílio: Leituras Modernas e Intertextualidade | Zé Moleza. (n.d.). Consultado a 18 de Outubro de 2017, em <http://www.zemoleza.com.br/trabalho-academico/sociais-aplicadas/letras/goncalves-dias-e-a-cancao-do-exilio-leituras-modernas-e-intertextualidade/>

[REVISTA CONTEMPORÂNEA – DOSSIÊ HISTÓRIA & LITERATURA]. (n.d.). Consultado em http://www.historia.uff.br/nec/sites/default/files/10_Novas_cancoes_do_exilio_3.pdf

É Complexamente Azul: Variações do Exílio. (n.d.). Consultado a 26 de Setembro de 2017, em <http://complexamenteazul.blogspot.pt/2012/12/variacoes-do-exilio.html>

ANÁLISE ESTILÍSTICA DO POEMA DE GONÇALVES DIAS CANÇÃO DO EXÍLIO” (n.d.). Consultado a 25 de Setembro de 2017, em http://artigos.netsaber.com.br/resumo_artigo_40246/artigo_sobre_analise-estilistica-do-poema-de-gonalves-dias-canao-do-exilio

Paródia da Canção do Exílio - Pensador. (n.d.). Consultado a 25 de Setembro de 2017, em https://www.pensador.com/parodia_da_cancao_do_exilio/

Paródias – Canção do exílio | Mar de Poesia. (n.d.). Consultado a 25 de Setembro de 2017, em <https://mardepoesia.wordpress.com/2017/02/15/parodias-cancao-do-exilio/>

APÊNDICES

A - LISTA DE PALAVRAS

Link para pdf com a lista de palavras:

<https://drive.google.com/file/d/1woQOnEDvK9-AYn2c0UXD-WQLfBoBWaiq/view?usp=sharing>

B - ESTUDOS DE FORMAS E CORES

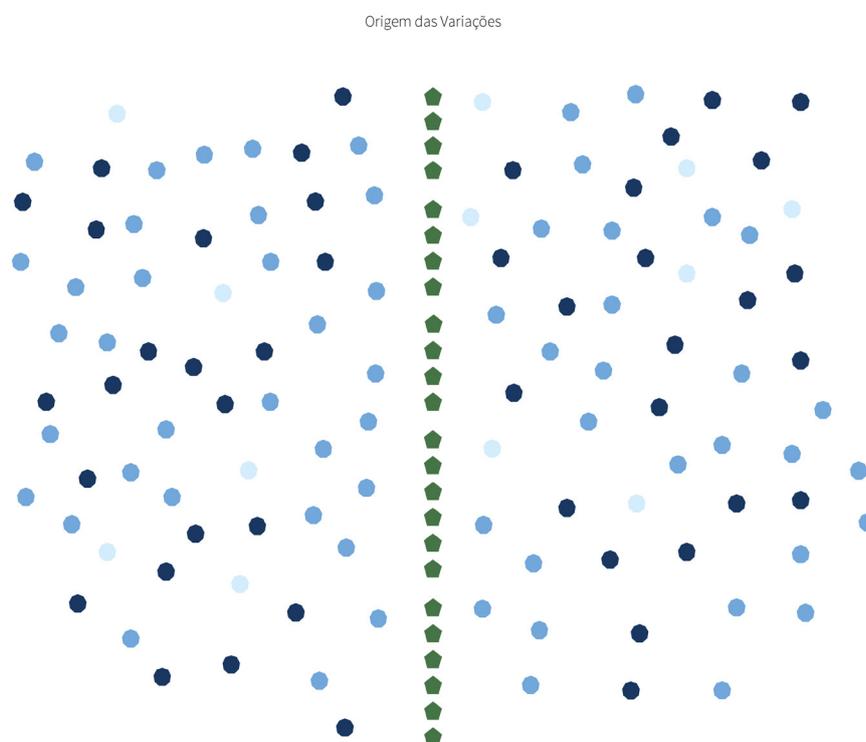


Figura B1 Aplicação das formas desenvolvidas.

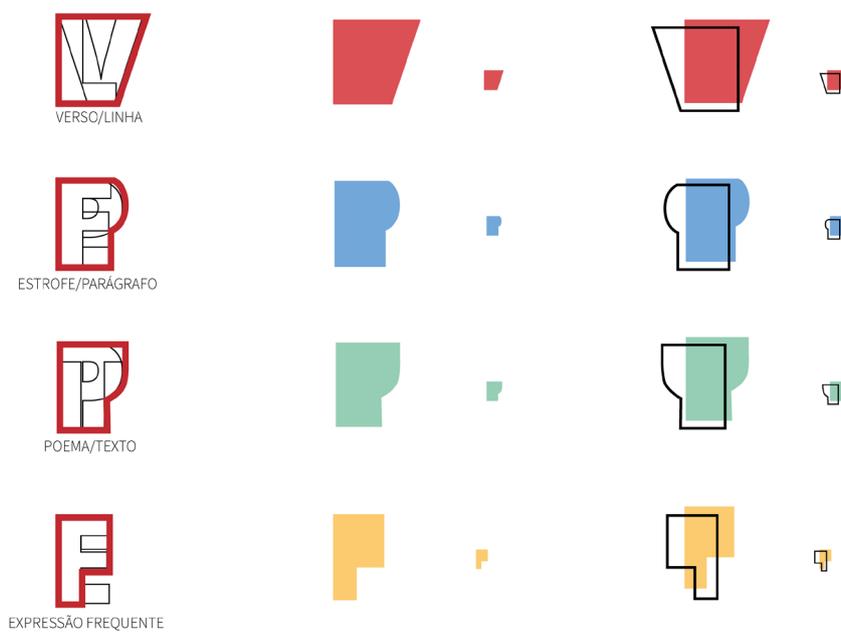


Figura B2 Exploração de novas formas para os diferentes tipos de opção visual.

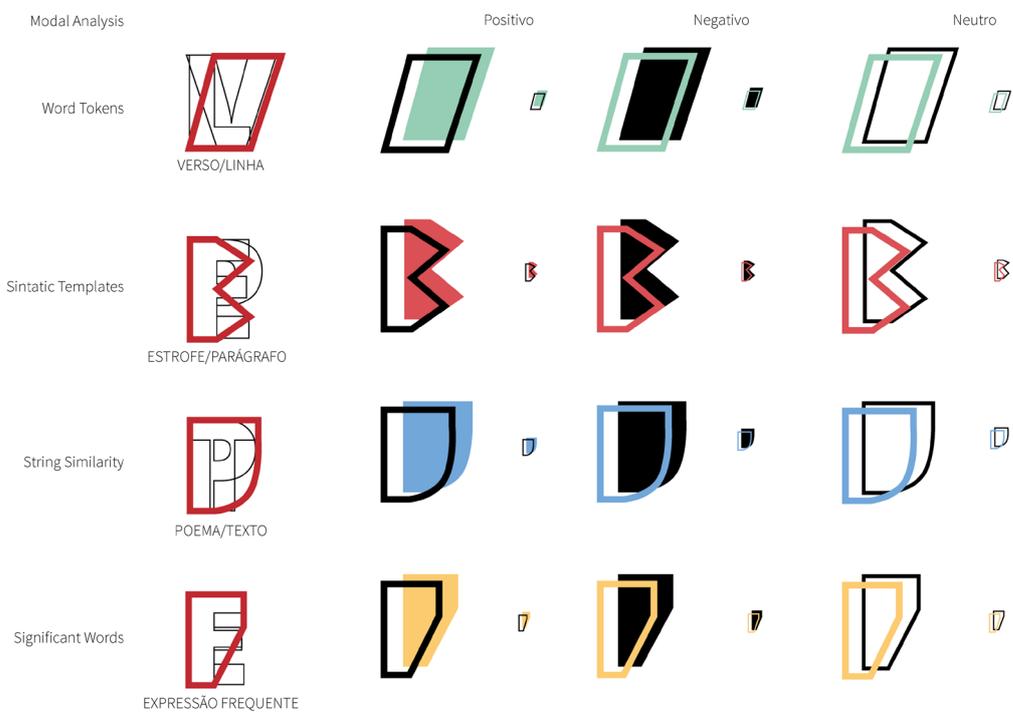


Figura B3 Novo estudo de formas mais complexas com distinção por sentimento.

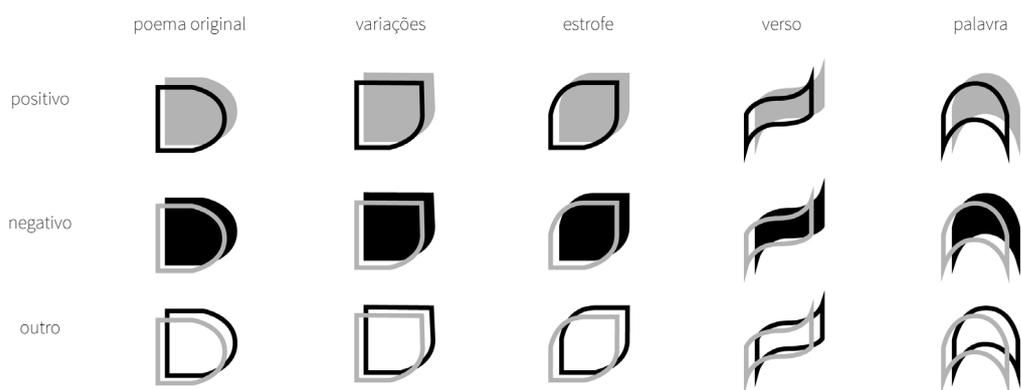


Figura B4 Estudo da sobreposição de símbolos consoante o seu tipo e a sua classificação

C - INTERAÇÃO DA VISUALIZAÇÃO “O POEMA ...AO LONGO DO TEMPO”

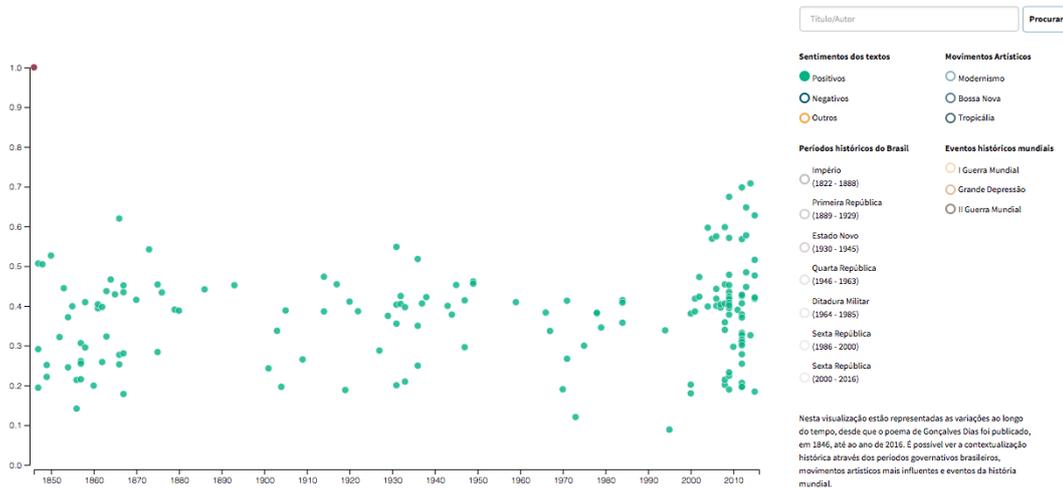


Figura C1 “O Poema ...ao longo do tempo” com a representação das variações “positivas.”

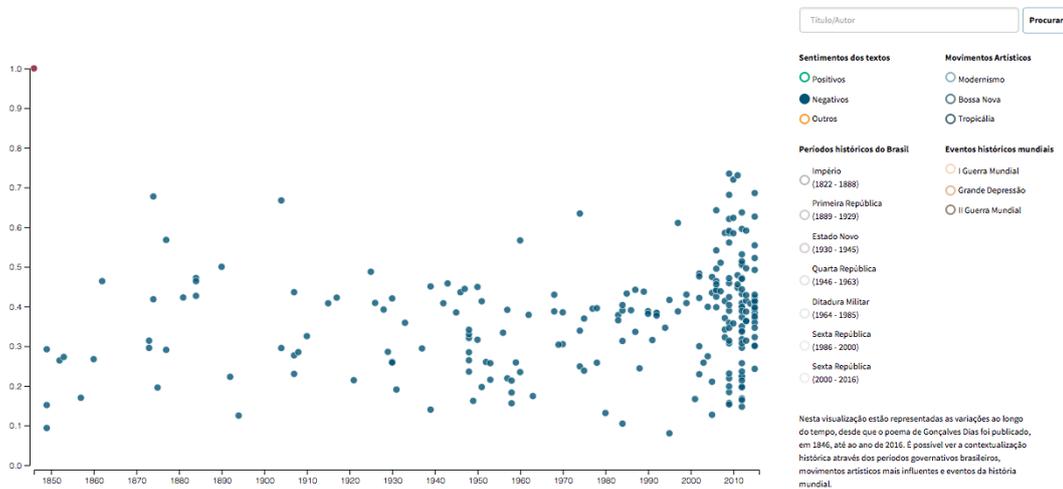


Figura C2 Representação das variações “negativas” no artefacto “O Poema ...ao longo do tempo”.

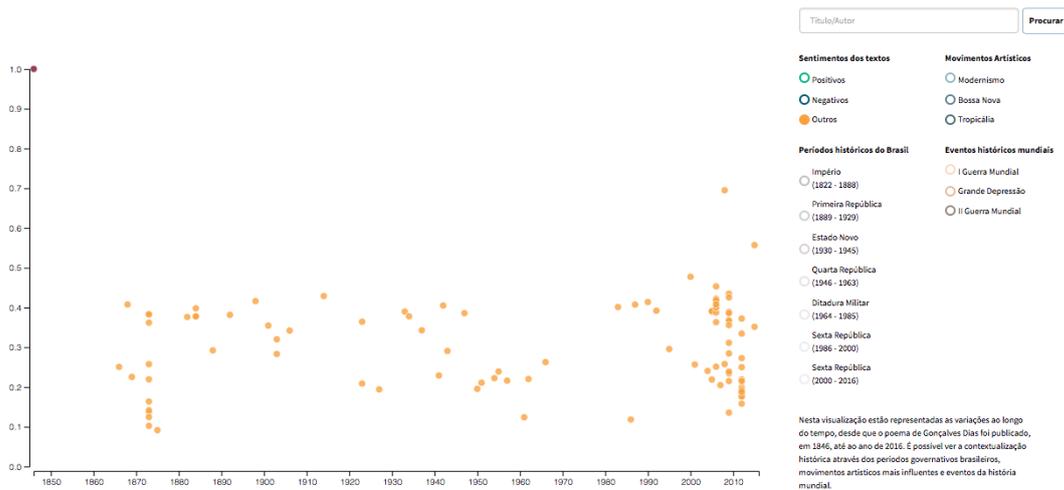


Figura C3 Representação das variações “outras” no artefacto “O Poema...ao longo do tempo”.

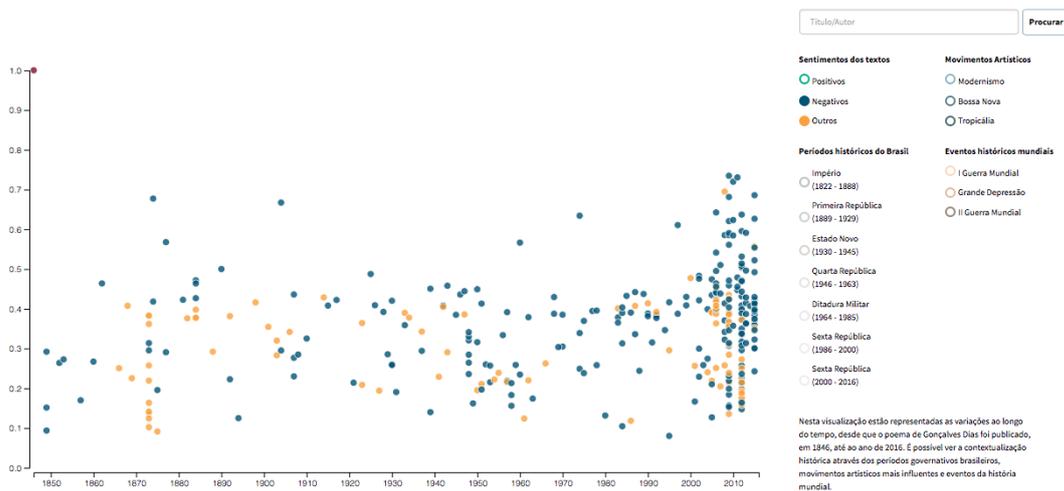


Figura C4 Representação das variações “negativas” e “outras”.

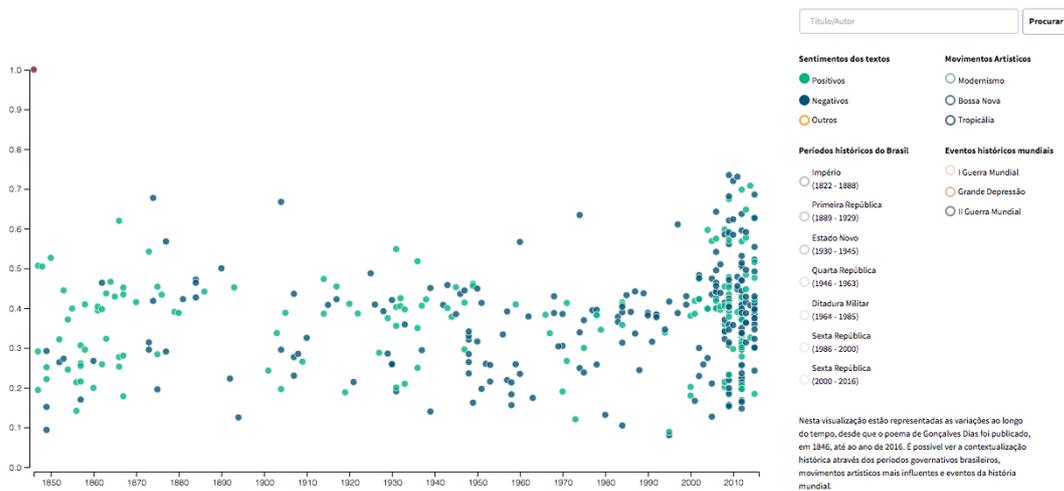


Figura C5 Representação das variações “positivas” e “negativas”.

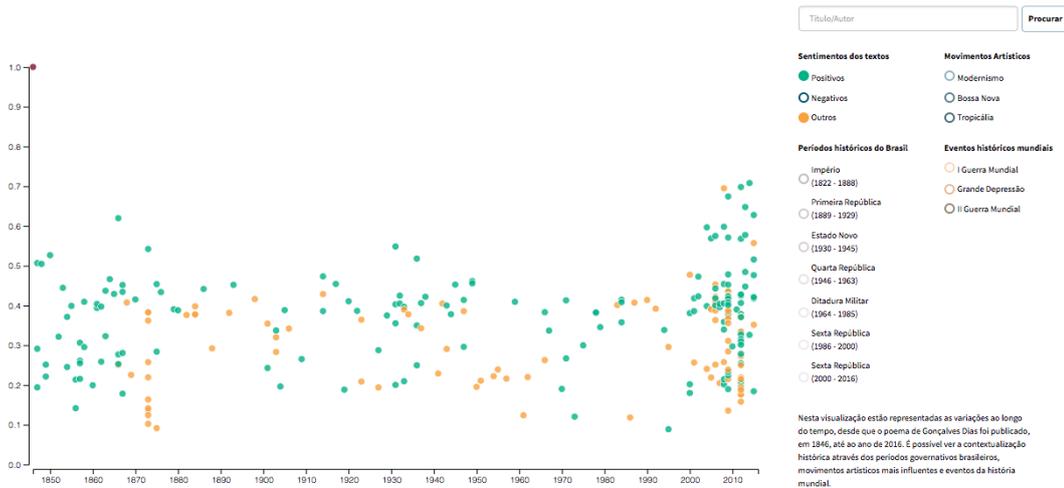


Figura C6 Representação das variações “positivas” e “outras”.

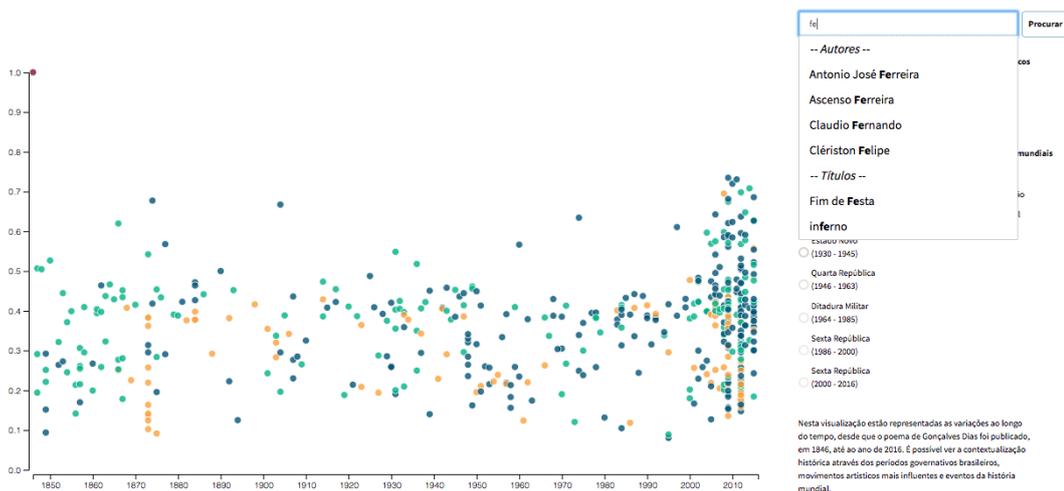


Figura C7 Procura com sugestões a partir do texto escrito pelo utilizador.

D - INTERAÇÃO DA VISUALIZAÇÃO “O POEMA ...PRESENTE NOS TEXTOS”

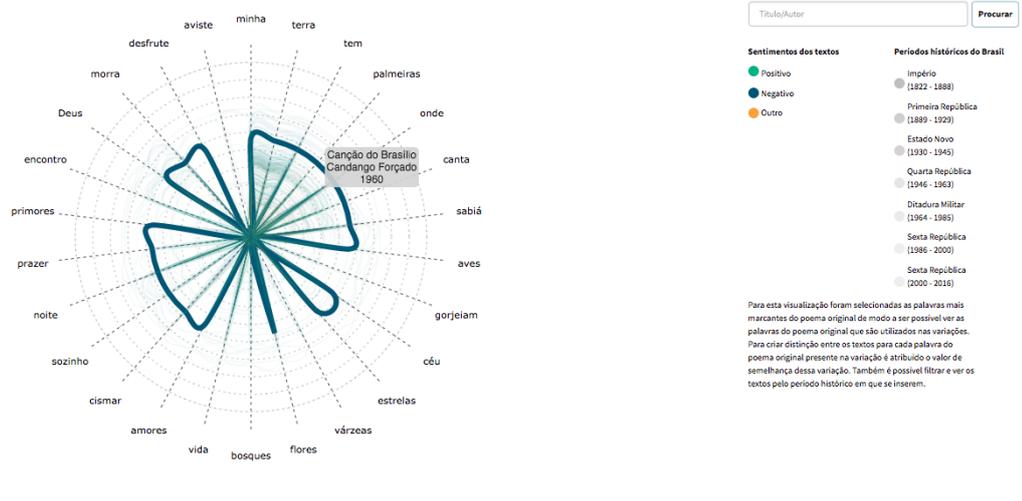


Figura D1 Realce com mouseover de uma variação.

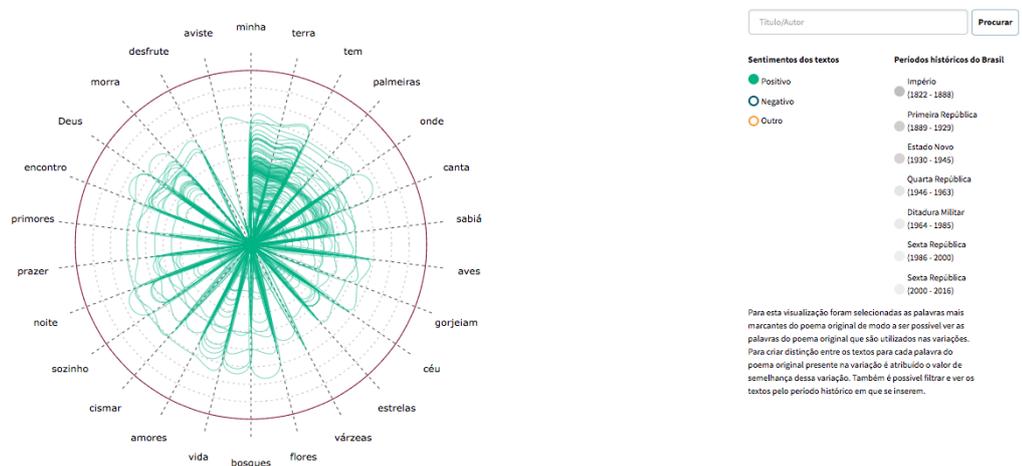


Figura D2 Representação dos poemas classificados como “positivos”.

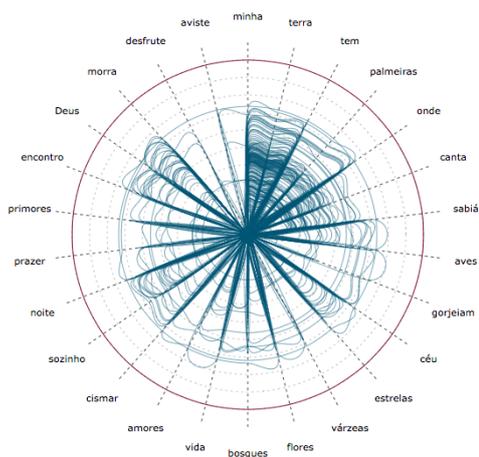


Figura D3 Representação dos poemas classificados como “negativos”.

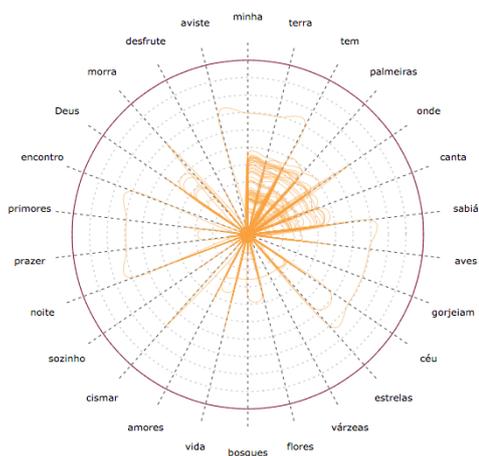


Figura D4 Representação dos poemas classificados como “outros”.

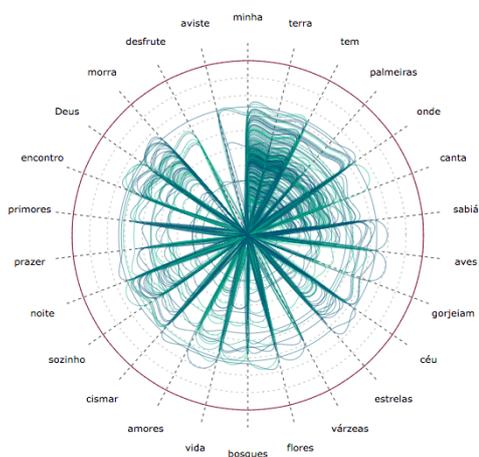
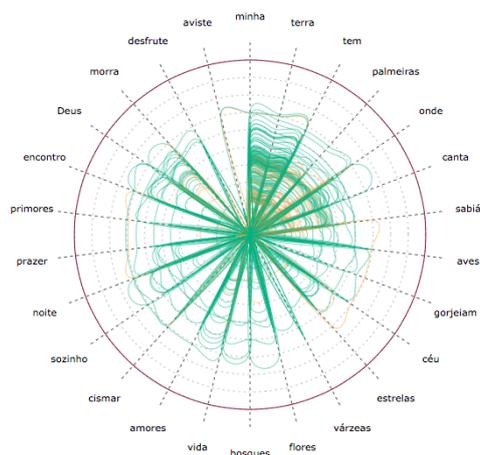


Figura D5 Representação da combinação dos poemas classificados como “positivos” e “negativos”.





Título/Autor

Sentimentos dos textos

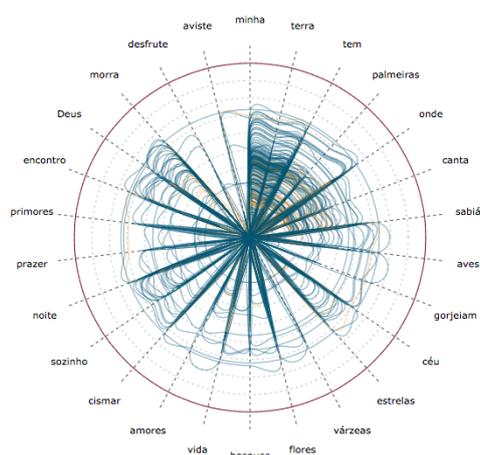
- Positivo
- Negativo
- Outro

Períodos históricos do Brasil

- Império (1822 - 1888)
- Primeira República (1889 - 1929)
- Estado Novo (1930 - 1945)
- Quarta República (1946 - 1963)
- Ditadura Militar (1964 - 1985)
- Sexta República (1986 - 2000)
- Sexta República (2000 - 2016)

Para esta visualização foram selecionadas as palavras mais marcantes do poema original de modo a ser possível ver as palavras do poema original que são utilizadas nas variações. Para criar distinção entre os textos para cada palavra do poema original presente na variação é atribuído o valor de semelhança dessa variação. Também é possível filtrar e ver os textos pelo período histórico em que se inserem.

Figura D6 Representação da combinação dos poemas classificados como “positivos” e “outros”.



Título/Autor

Sentimentos dos textos

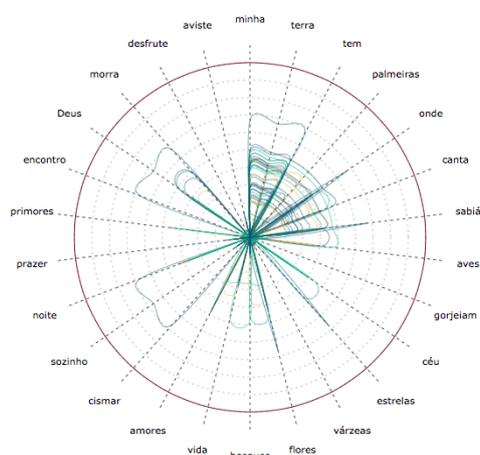
- Positivo
- Negativo
- Outro

Períodos históricos do Brasil

- Império (1822 - 1888)
- Primeira República (1889 - 1929)
- Estado Novo (1930 - 1945)
- Quarta República (1946 - 1963)
- Ditadura Militar (1964 - 1985)
- Sexta República (1986 - 2000)
- Sexta República (2000 - 2016)

Para esta visualização foram selecionadas as palavras mais marcantes do poema original de modo a ser possível ver as palavras do poema original que são utilizadas nas variações. Para criar distinção entre os textos para cada palavra do poema original presente na variação é atribuído o valor de semelhança dessa variação. Também é possível filtrar e ver os textos pelo período histórico em que se inserem.

Figura D7 Representação da combinação dos poemas classificados como “outros” e “negativos”.



Título/Autor

Sentimentos dos textos

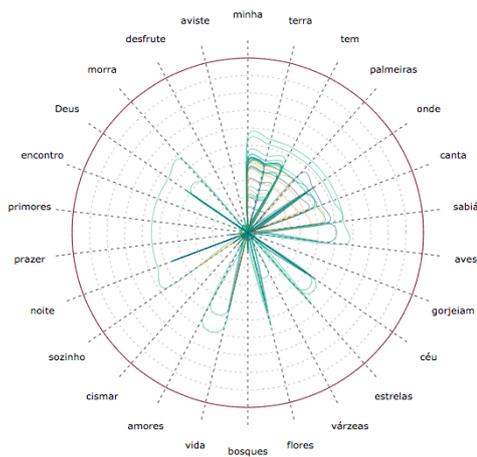
- Positivo
- Negativo
- Outro

Períodos históricos do Brasil

- Império (1822 - 1888)
- Primeira República (1889 - 1929)
- Estado Novo (1930 - 1945)
- Quarta República (1946 - 1963)
- Ditadura Militar (1964 - 1985)
- Sexta República (1986 - 2000)
- Sexta República (2000 - 2016)

Para esta visualização foram selecionadas as palavras mais marcantes do poema original de modo a ser possível ver as palavras do poema original que são utilizadas nas variações. Para criar distinção entre os textos para cada palavra do poema original presente na variação é atribuído o valor de semelhança dessa variação. Também é possível filtrar e ver os textos pelo período histórico em que se inserem.

Figura D8 Representação das variações correspondentes ao período Primeira República.



Título/Autor

Sentimentos dos textos

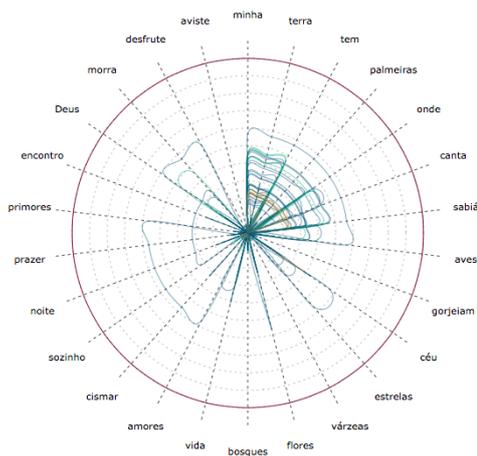
- Positivo
- Negativo
- Outro

Períodos históricos do Brasil

- Império (1822 - 1888)
- Primeira República (1889 - 1929)
- Estado Novo (1930 - 1945)
- Quarta República (1946 - 1963)
- Ditadura Militar (1964 - 1985)
- Seta República (1986 - 2000)
- Seta República (2000 - 2016)

Para esta visualização foram selecionadas as palavras mais marcantes do poema original de modo a ser possível ver as palavras do poema original que são utilizadas nas variações. Para criar distinção entre os textos para cada palavra do poema original presente na variação é atribuído o valor de semelhança dessa variação. Também é possível filtrar e ver os textos pelo período histórico em que se inserem.

Figura D9 Representação das variações correspondentes ao período Estado Novo.



Título/Autor

Sentimentos dos textos

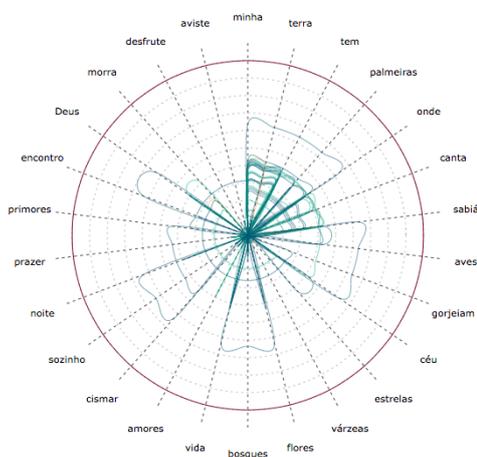
- Positivo
- Negativo
- Outro

Períodos históricos do Brasil

- Império (1822 - 1888)
- Primeira República (1889 - 1929)
- Estado Novo (1930 - 1945)
- Quarta República (1946 - 1963)
- Ditadura Militar (1964 - 1985)
- Seta República (1986 - 2000)
- Seta República (2000 - 2016)

Para esta visualização foram selecionadas as palavras mais marcantes do poema original de modo a ser possível ver as palavras do poema original que são utilizadas nas variações. Para criar distinção entre os textos para cada palavra do poema original presente na variação é atribuído o valor de semelhança dessa variação. Também é possível filtrar e ver os textos pelo período histórico em que se inserem.

Figura D10 Representação das variações correspondentes ao período Quarta República.



Título/Autor

Sentimentos dos textos

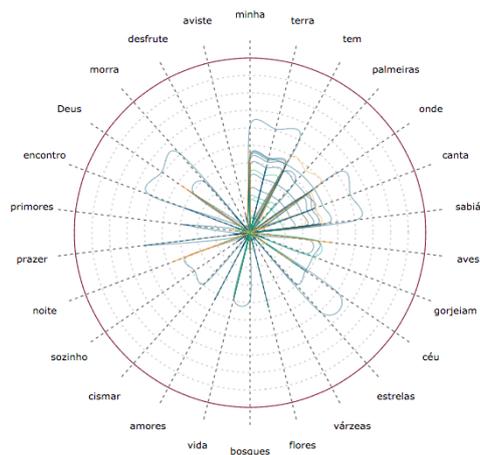
- Positivo
- Negativo
- Outro

Períodos históricos do Brasil

- Império (1822 - 1888)
- Primeira República (1889 - 1929)
- Estado Novo (1930 - 1945)
- Quarta República (1946 - 1963)
- Ditadura Militar (1964 - 1985)
- Seta República (1986 - 2000)
- Seta República (2000 - 2016)

Para esta visualização foram selecionadas as palavras mais marcantes do poema original de modo a ser possível ver as palavras do poema original que são utilizadas nas variações. Para criar distinção entre os textos para cada palavra do poema original presente na variação é atribuído o valor de semelhança dessa variação. Também é possível filtrar e ver os textos pelo período histórico em que se inserem.

Figura D11 Representação das variações correspondentes ao período Ditadura Militar.



Título/Autor

Sentimentos dos textos

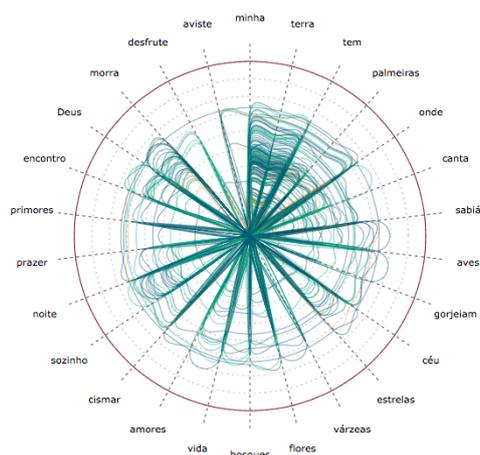
- Positivo
- Negativo
- Outro

Períodos históricos do Brasil

- Império (1822 - 1888)
- Primeira República (1889 - 1929)
- Estado Novo (1930 - 1945)
- Quarta República (1946 - 1963)
- Ditadura Militar (1964 - 1985)
- Sexta República (1986 - 2000)
- Sexta República (2000 - 2016)

Para esta visualização foram selecionadas as palavras mais marcantes do poema original de modo a ser possível ver as palavras do poema original que são utilizadas nas variações. Para criar distinção entre os textos para cada palavra do poema original presente na variação é atribuído o valor de semelhança dessa variação. Também é possível filtrar e ver os textos pelo período histórico em que se inserem.

Figura D12 Representação das variações correspondentes ao período Sexta República no século XX.



Título/Autor

Sentimentos dos textos

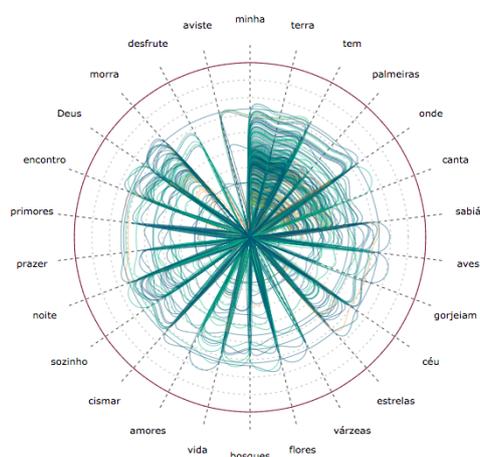
- Positivo
- Negativo
- Outro

Períodos históricos do Brasil

- Império (1822 - 1888)
- Primeira República (1889 - 1929)
- Estado Novo (1930 - 1945)
- Quarta República (1946 - 1963)
- Ditadura Militar (1964 - 1985)
- Sexta República (1986 - 2000)
- Sexta República (2000 - 2016)

Para esta visualização foram selecionadas as palavras mais marcantes do poema original de modo a ser possível ver as palavras do poema original que são utilizadas nas variações. Para criar distinção entre os textos para cada palavra do poema original presente na variação é atribuído o valor de semelhança dessa variação. Também é possível filtrar e ver os textos pelo período histórico em que se inserem.

Figura D13 Representação das variações correspondentes ao período Sexta República no século XXI.



-- Autores --

Afonso Romano Sant'anna
 Armindo Trevisan
 Bithencourt Sampaio
 Breno Lobato e Samara Cristina

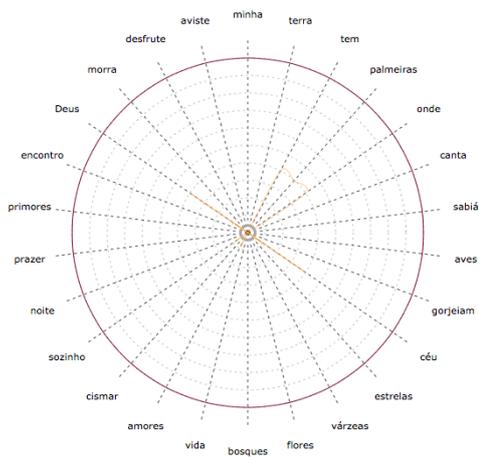
-- Títulos --

Exílio dos Sabiás
 E o sabiá?

Sexta República (1986 - 2000)
 Sexta República (2000 - 2016)

Para esta visualização foram selecionadas as palavras mais marcantes do poema original de modo a ser possível ver as palavras do poema original que são utilizadas nas variações. Para criar distinção entre os textos para cada palavra do poema original presente na variação é atribuído o valor de semelhança dessa variação. Também é possível filtrar e ver os textos pelo período histórico em que se inserem.

Figura D14 Representação da função de pesquisa no artefacto.



Bithencourt Sampaio

Sentimentos dos textos

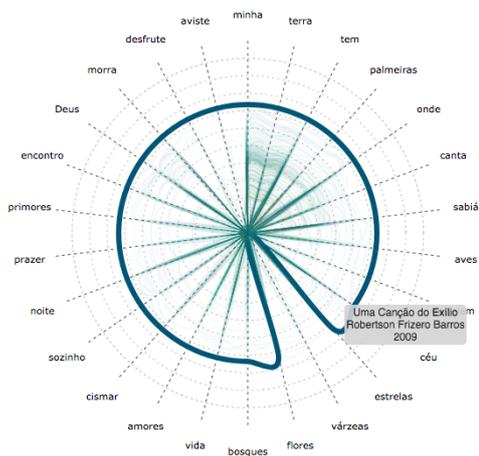
- Positivo
- Negativo
- Outro

Períodos históricos do Brasil

- Império (1822 - 1888)
- Primeira República (1889 - 1929)
- Estado Novo (1930 - 1945)
- Quarta República (1946 - 1963)
- Ditadura Militar (1964 - 1985)
- Sexta República (1986 - 2000)
- Sexta República (2000 - 2016)

Para esta visualização foram selecionadas as palavras mais marcantes do poema original de modo a ser possível ver as palavras do poema original que são utilizadas nas variações. Para criar distinção entre os textos para cada palavra do poema original presente na variação é atribuído o valor de semelhança dessa variação. Também é possível filtrar e ver os textos pelo período histórico em que se inserem.

Figura D15 Resultado da procura efectuada.



Título/Autor

Sentimentos dos textos

- Positivo
- Negativo
- Outro

Períodos históricos do Brasil

- Império (1822 - 1888)
- Primeira República (1889 - 1929)
- Estado Novo (1930 - 1945)
- Quarta República (1946 - 1963)
- Ditadura Militar (1964 - 1985)
- Sexta República (1986 - 2000)
- Sexta República (2000 - 2016)

Para esta visualização foram selecionadas as palavras mais marcantes do poema original de modo a ser possível ver as palavras do poema original que são utilizadas nas variações. Para criar distinção entre os textos para cada palavra do poema original presente na variação é atribuído o valor de semelhança dessa variação. Também é possível filtrar e ver os textos pelo período histórico em que se inserem.

Figura D16 Realce de texto com informação correspondente

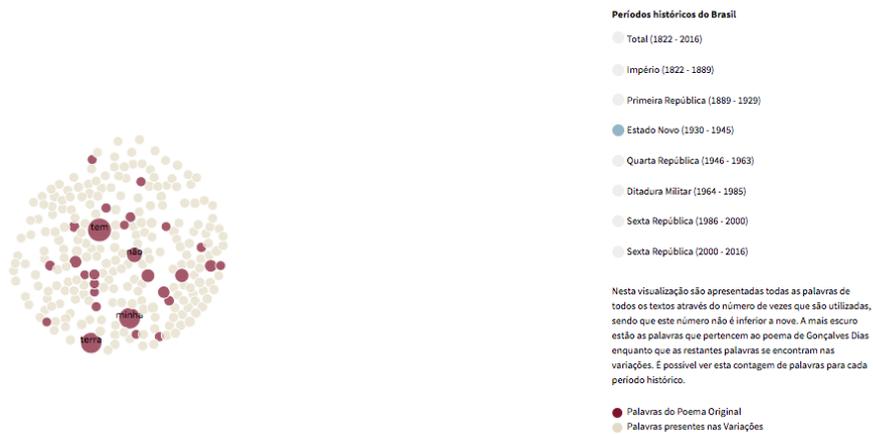


Figura E3 Representação da visualização “O Poema ...em palavras” no período Estado Novo.

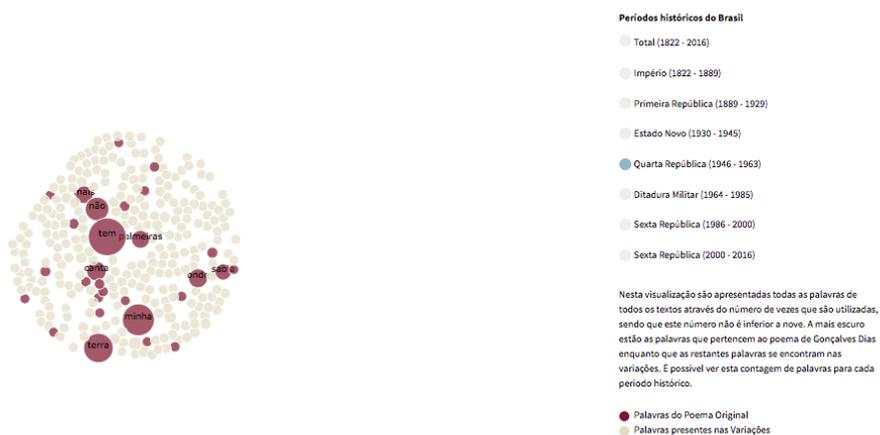


Figura E4 Representação da visualização “O Poema ...em palavras” no período Quarta República.

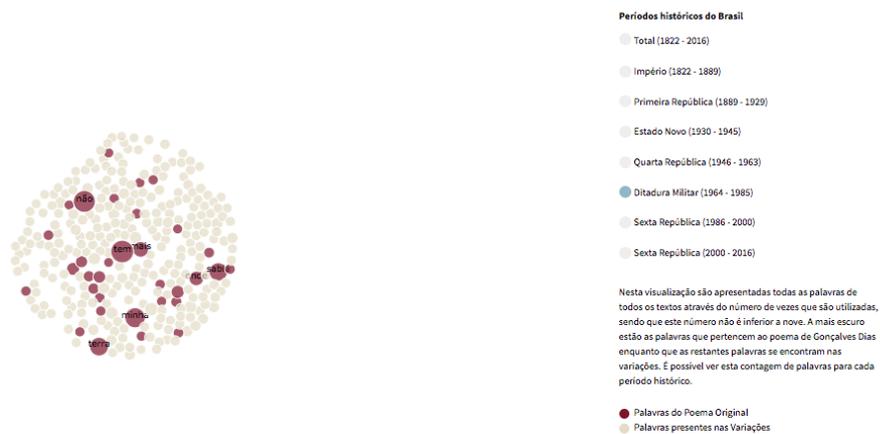


Figura E5 Representação da visualização “O Poema ...em palavras” no período Ditadura Militar.

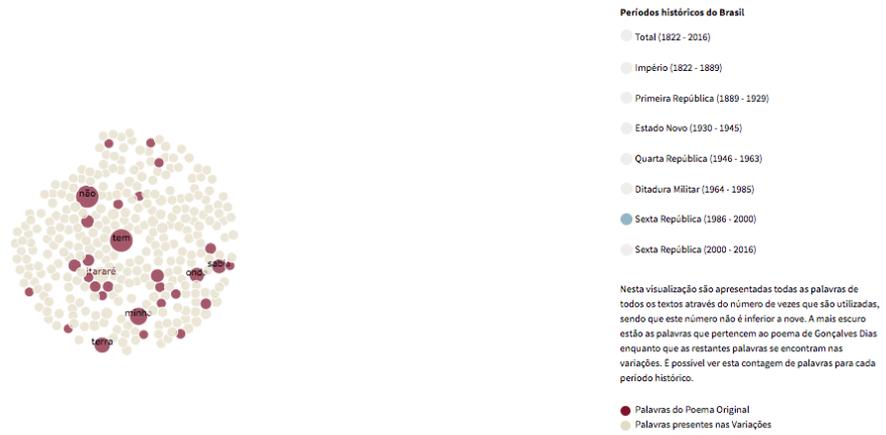


Figura E6 Representação da visualização “O Poema ...em palavras” no período Sexta República do séc.XX.

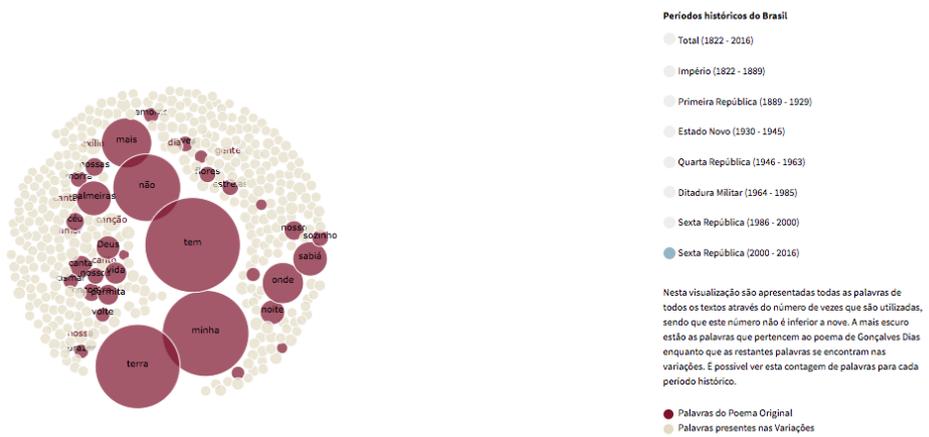
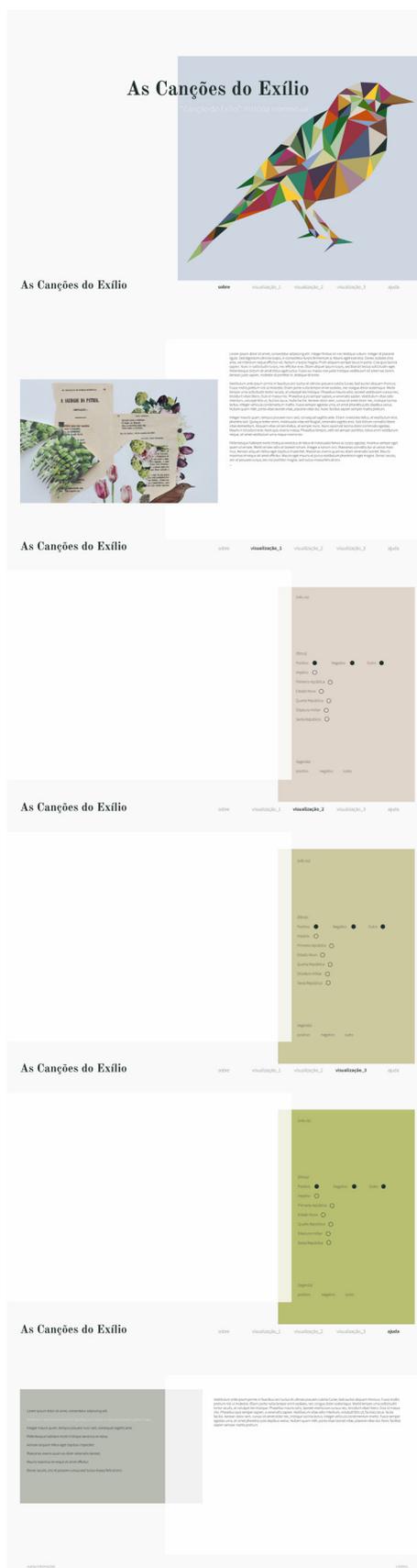


Figura E7 Representação da visualização “O Poema ...em palavras” no período Sexta República do séc.XXI.

F - ESBOÇO DIGITAL DA PLATAFORMA



G - PLATAFORMA EM INGLÊS

PT | EN

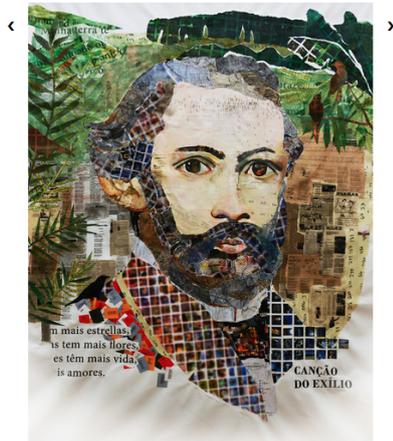
Songs of Exile



Figura G1 Página Inicial

Songs of Exile

[The Poem](#) ... over time ... present in the texts ... In words The Project



"Canção do Exílio" of Antônio Gonçalves Dias

In 1843, while Law student at University of Coimbra, Antônio Gonçalves Dias wrote the poem "Canção do Exílio". In Brazil the poem is one of the most important of the beginning of the Brazilian Romanticism and is seen as being a declaration of love to the motherland. Since its publication in 1846, it has become one of the most popular poems in Brazil, being used as much as inspiration for romantic or critical variations. The words written by Gonçalves Dias have been used over the years as an aid to the invention of new texts and poems. The creation of these variations often reflects the feeling that is lived in each moment in Brazil. The poem is thus a means for the expression of feelings that can be glorification or criticism but that allow a description and follow-up of the country over the years through the eyes of the population. The influence of Gonçalves Dias can be found in the Brazilian anthem or in authors like Carlos Drummond de Andrade, Oswald de Andrade, Tom Jobim and Chico Buarque.

Bird Watching: Visualizing the influence of "Canção do Exílio" by Joshua Enslin and Alaina Enslin

The exhibition *Bird Watching: Visualizing the influence of "Canção do Exílio"* presented the research and analysis of 500 variations of the poem "Canção do Exílio" collected by Joshua Enslin under his postdoctoral project. In order to present the results of the analysis of the texts, objects were created to allow the visualization of the influence of the poem of Gonçalves Dias. Joshua Enslin and Alaina Enslin have developed illustrations of texts, and author, and graphic representation of the similarity of variations with the original poem. The exhibition was present at

Figura G2 Página sobre a "Canção do Exílio" e a exposição

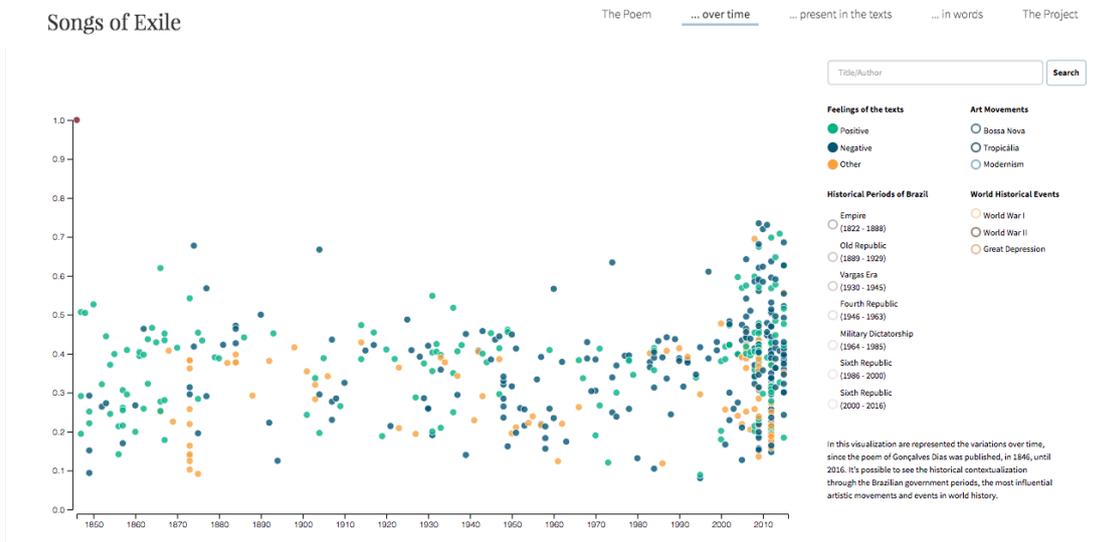


Figura G3 Página da primeira visualização.

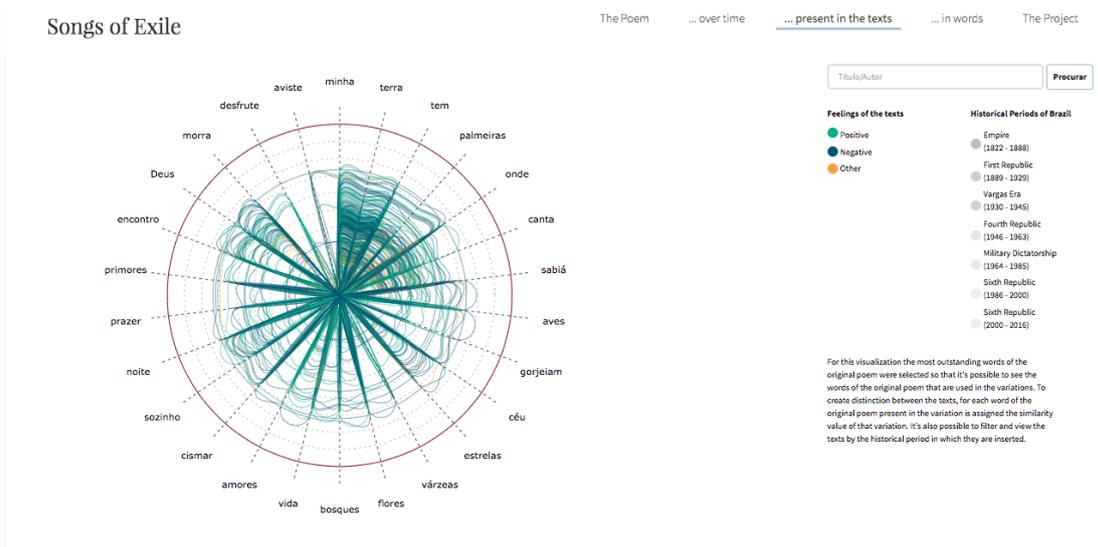


Figura G4 Página da segunda visualização.

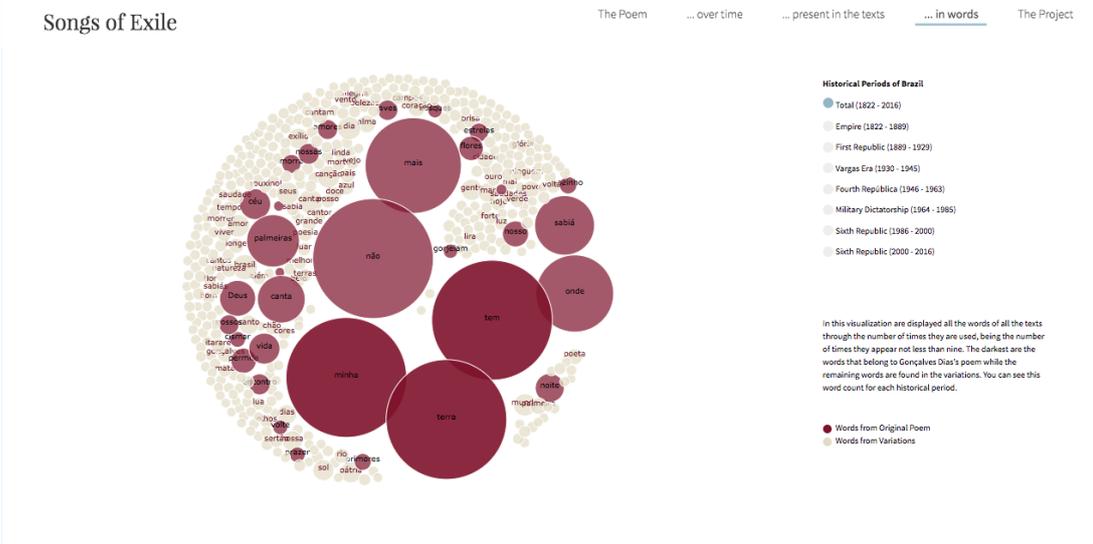


Figura G5 Página da terceira visualização.

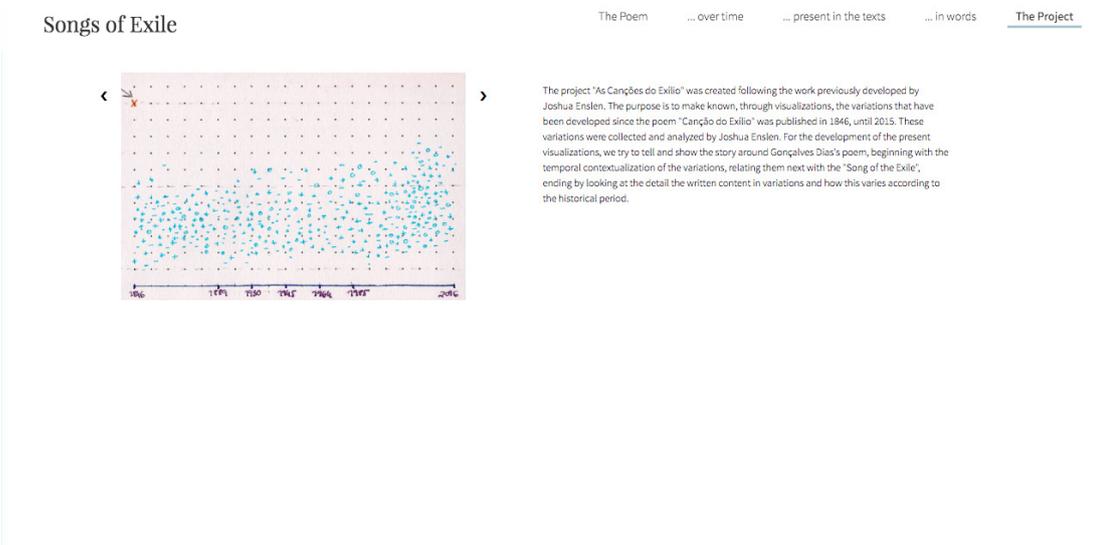


Figura G6 Página descritiva do projecto.

