

DISSERTAÇÃO

Novos paradigmas de visualização
de dados em larga escala

Daniela Sofia Trindade Bernardo

—

Mestrado em Design e Multimédia
Faculdade de Ciências e Tecnologia
Universidade de Coimbra

—

Setembro 2015

Orientador: Lício Roque
Co-orientador: Bruna Sousa
Orientador SAPO: Jorge Teixeira

—

Juri Arguente: Pedro Cruz
Juri Vogal: Maria Alice Geirinhas

—

Mestrado em Design e Multimédia
Faculdade de Ciências e Tecnologia
Universidade de Coimbra

—

Setembro 2015

Resumo

Esta dissertação foca-se na área da visualização de informação, com especial interesse na representação de dados em larga escala que se estendem por um longo período de tempo, bem como na interação como meio de explorar os dados representados. A investigação que suporta esta dissertação centra-se no estudo da evolução das formas gráficas, nos princípios e conceitos interativos aplicados à visualização de informação e no levantamento de princípios de bom design de gráficos. O resultado da investigação na prática de design é um protótipo *web* de uma visualização de informação interativa que aplica soluções derivadas dos conceitos interativos – seleção, exploração filtro, detalhar/abstrair e relacionar informação– e princípios de design estudados – como a utilização de números, palavras e desenhos, evitando conteúdos meramente decorativos; fazer uso de escalas relevantes e proporcionar à visualização uma qualidade narrativa.

Palavras-chave

Visualização de informação, Visualização de dados, Princípios interativos, Gráficos *time-series*, Densidade de dados, Dados em Larga Escala, IHC.

Abstract

This dissertation focuses on the field of information visualization with special interest in the representation of large-scale data over a big period of time, and also the interaction as a mean to explore the data. The research that supports this dissertation focuses on the study of the evolution of graphic forms, principles and concepts applied to interactive information visualization and collecting of good graphics design principles. The result of this practical design research is a web prototype of an interactive information visualization applying solutions derived from interactive concepts – selection, filter, detailing/abstracting and relate information – and studied design principles – such as the use of numbers, words and drawings, avoiding purely decorative content; make use of relevant scales and provide a narrative quality.

Keywords

Information visualization, Data visualization, Interactive principles, Time-series charts, Data density, Dados em Larga Escala, IHC.

Agradecimentos

Aos meus pais, pois sem a sua ajuda nunca poderia ter chegado a esta etapa acadêmica.

À minha mãe, a minha quarta orientadora, pelo apoio incondicional e por acreditar nas minhas capacidades.

Ao meu irmão, pelos favores prestados enquanto criava raízes sentada ao computador.

Aos meus orientadores pelos bons conselhos e a disponibilidade durante todo o processo.

À Eunice e à Dona Felizarda por terem sido a minha segunda família nos dias de semana, deste segundo semestre.

Aos colegas do SAPO – ao Bruno, ao Jorge, ao Fábio, ao Zé Pedo, ao Rogério, ao Chin e ao David – por tão bem saberem como acolher os novos colegas e integrá-los no ambiente SAPO. Por toda a ajuda, pelas pausas para o café, pela companhia ao almoço e, até, pelas conversas de ‘maionese’.

Ao David da usabilidade, pelos conselhos preciosos.

Aos amigos designers – à Leila, à Andreia, à Nini e ao Luís – pela ajuda e pelos conselhos nas horas mais precisas.

À Tita e à Adriana pela amizade duradora apesar da distância.

Aos meus avós maternos pelas palavras de incentivo.

À Filipa pela amizade, pela paciência e pelo ouvido amigo nas horas de maior desânimo.

A esta experiência fantástica, por vezes assustadora, desanimadora, por vezes inspiradora e, muitas vezes professora.

Ao André Lopes pela ajuda, pela compreensão da minha ausência em várias ocasiões.

Obrigada a todos!

Índice

Resumo & Palavras-Chave
Abstract & Keywords
Agradecimentos

11	INTRODUÇÃO
13	Contexto
13	Motivação
13	Objetivos
14	Metodologias
14	Estrutura do documento
17	ESTADO DA ARTE
19	Evolução das Formas Gráficas na Representação de Dados
31	IHC na Visualização de Informação
36	Design da Informação
46	Trabalhos Relacionados
58	Trabalhos Relacionados com Política
66	Conclusões
67	ABORDAGEM
69	Objetivos
69	Metodologias
70	Planeamento
73	PROPOSTA
75	Tema e Público-alvo
75	Definição dos Dados
77	Recolha de Dados
77	Dados Eleitorais
78	Máquina do Tempo
79	Fatores socioeconómicos
79	Problemática
80	Visualização
80	Narrativa
80	Metáfora Visual
83	Cenários de utilização
85	DESENVOLVIMENTO PRÁTICO
87	Design
87	Visualização das Eleições Legislativas
90	Visualização das Notícias
101	Visualização dos Fatores Socioeconómicos
102	Interação
108	Tecnologias
108	Implementação

133	AVALIAÇÃO
134	Metodologia
134	Requisitos
134	Sessão
135	Participantes
135	Perfis
137	Questões em pesquisa
137	Tarefas
143	Resultados
143	Resultados para as questões em pesquisa
146	Resultados e problemas das tarefas
158	Melhoramentos recomendados
161	CONCLUSÃO E PERSPECTIVAS FUTURAS
165	BIBLIOGRAFIA
167	BIBLIOGRAFIA DE IMAGENS
	ANEXO A
	ANEXO B
	ANEXO C
	ANEXO D
	ANEXO E

Introdução



O desenvolvimento das capacidades de processamento, armazenamento e interação dos computadores veio revolucionar a área da visualização de informação. Podemos dizer que, atualmente, estamos numa era onde o excesso de informação é uma realidade. A produção de conteúdos digitais é constante, dando origem a um armazenamento exaustivo dessas informações.

A recolha de dados é uma tarefa que nós, seres humanos, executamos cada vez mais e melhor, tanto a nível pessoal como empresarial. O mesmo não se pode dizer da utilidade que reservamos para estes dados que, na maioria das vezes, ficam armazenados ao longo de anos. O uso potencial destes dados pode ser explorado através da representação gráfica dos mesmos. Estes dados brutos tornam-se, portanto, úteis quando aplicamos métodos das áreas das ciências computacionais, estatística e design gráfico.

É extremamente difícil detetar padrões comparando linhas em tabelas, sendo raras as pessoas que o conseguem fazer sem auxílio de uma ferramenta visual. Como refere Bertin¹, “a comparação de duas tabelas, cada uma com 100 linhas e 100 colunas, é muito mais fácil e até pode ser instantânea se os dados forem representados graficamente.” Até mesmo as crianças conseguem interpretar gráficos de barras e extrair a mensagem da representação visual desses números, o que comprova que a visualização de dados é um processo poderoso, capaz de nos mostrar o mundo através de uma nova perspetiva, revelando padrões e tendências inesperadas nas informações que, de outra forma, permaneceriam ocultas.

A invenção e disponibilização de um acesso fácil à internet foram acontecimentos que, também, revolucionaram esta área, na medida em que veio possibilitar a recolha de dados em tempo real. A visualização de informação passa, assim, a lidar com dados do mundo real, que estão em constante mudança, como é o exemplo da temperatura, chegada de transportes públicos, padrões de tráfego num *website*, entre outros.

Os desenvolvimentos na área tecnológica são cada vez mais rápidos e, ainda hoje, criam novos cenários na área da visualização, que levantam questões relevantes e dignas de estudo e investigação. Face a este cenário, de disponibilização massiva de informação, como poderemos representar conjuntos de dados, recolhidos em grande quantidade e num extenso período de tempo?

1.1 Contexto

Face à produção diária de conteúdos digitais, o tratamento gráfico destes dados torna-se uma necessidade a fim de potenciarem uma extração de novos conhecimentos úteis. Por estes mesmos motivos, a pesquisa na área da visualização de informação está em constante crescimento. Foca-se no estudo e uso de técnicas gráficas eficazes na transmissão de informação quantitativa, qualitativa ou ordenada de forma a facilitar a exploração e análise de dados.

O SAPO é uma das empresas que, diariamente, armazena novos dados em larga escala. A necessidade de criar visualizações interativas para a exploração de dados tem vindo a ser explorada e estudada, tendo já tido alguns resultados, como é o caso dos seguintes projetos: “A Máquina do Tempo”, “Grande Área”, “Lúsica”, entre outros.

SAPO

<http://www.sapo.pt/>

Máquina do Tempo

<http://maquinadotempo.sapo.pt/>

Grande Área

<http://grandeara.desporto.sapo.pt/>

Lúsica

<http://lusica.labs.sapo.pt/>

O registo dos dados no SAPO é feito diariamente e as suas bases de dados contam já com informações datadas de décadas atrás. A visualização de dados em larga escala, referentes a um período cuja durabilidade se aproxima a meio século, através de uma navegação interativa torna-se, assim, numa necessidade que origina a proposta de estágio onde se insere esta dissertação.

1.2 Motivação

“Proporcionar uma solução significativa para a visualização de dados complexos requerer conhecimento em diversas áreas”²⁵. O design gráfico, bem como a multimédia – ou seja, conhecimentos computacionais – são, também, áreas englobadas no campo da visualização de informação, na medida em que, o design gráfico, estuda soluções visuais para a representação dos dados e, a multimédia, trata de as implementar. Desta forma, neste campo, existe uma fusão de áreas científicas e artísticas, característica inerente ao curso de Design e Multimédia, no qual esta dissertação se insere.

O interesse, do SAPO, em ver este projeto ser concretizado assenta na vontade e necessidade de dar tratamento à massiva quantidade de dados recolhidos diariamente, bem como disponibilizar ferramentas visuais que permitam a exploração destes dados de uma forma interativa adequada às necessidades do utilizador.

1.3 Objetivos

O principal objetivo com esta dissertação é a investigação na área da visualização de informação, com vista a explorar soluções que lidem com cenários onde existe a necessidade de tratar e representar dados em larga escala, referentes a um longo período temporal e de utilizar a interação como forma de navegação nos mesmos.

Como objetivo adicional, pretende-se desenvolver um protótipo funcional, que aplica uma solução baseada na pesquisa realizada. À posteriori avalia-se a pertinência, desta solução, na comunicação da história contada pelos dados, através de uma exploração interativa dos mesmos, numa representação com um extenso período temporal.

²⁵ Fry, 2008

1.4 Metodologias

De modo a alcançar os objetivos propostos, numa primeira fase com uma abordagem teórica, é feito um levantamento histórico da evolução das formas gráficas para a representação gráfica de dados e transmissão clara de ideias. Em paralelo, é concretizada uma pesquisa e recolha de projetos na área da visualização de informação cujas características correspondam aos requisitos do projeto. Para cada um destes projetos, é apresentado, através de tabelas, um estudo semiótico orientado para o estudo das marcas gráficas – que representam os dados – e das perceções que estas provocam – perceções que possibilitam uma transmissão clara da ideia quando bem utilizadas. Este estudo é suportado pela investigação realizada e exposta em §2.3.

Seguidamente, numa abordagem mais prática, definimos o tipo de dados que nos interessa ver tratados através da formulação de perguntas sobre a história que estes podem contar. Segue-se a recolha dos mesmos na base de dados do SAPO, através de uma linguagem baseada em *queries*. E, de seguida, a análise destes dados, de modo a escolher uma forma gráfica adequada às suas necessidades. Consequentemente, é gerada uma proposta de design para o protótipo, implementada durante o desenvolvimento do mesmo, através de linguagens para *web*.

Finalmente, o último consiste na avaliação da solução implementada, ou seja, a avaliação do protótipo. Esta, é executada utilizando testes informais com potenciais utilizadores, aos quais é pedida a concretização de um conjunto de tarefas e avaliação das suas dificuldades durante a realização das mesmas, de modo a ser possível apurar problemas que possam ser eliminados ou melhorados.

1.5 Estrutura do documento

Este documento é composto por sete capítulos. Este primeiro capítulo tem como finalidade a introdução do tema da dissertação, revelar os principais motivos que levaram à escolha deste projeto, definir os objetivos e a abordagem, bem como, a apresentação da estrutura do documento.

O segundo capítulo, referente ao Estado da Arte, descreverá a evolução das representações gráficas de dados, introduzirá conceitos de interação aplicados à visualização de informação, enumerará princípios gráficos para design de gráficos e descreverá projetos que se relacionam com o projeto proposto nesta dissertação.

O terceiro capítulo, relativo à Abordagem, pretende clarificar as metodologias adotadas ao longo da realização do projeto. São, também, esclarecidos os objetivos e é apresentado o plano de trabalho.

O quarto capítulo, referente à Proposta – para a visualização de informação –, tem com objetivo informar o leitor, sobre o tema e o público-alvo deste projeto, bem como descrever o processo de definição de dados e respetiva recolha, quando esta foi necessária. A construção da narrativa e escolha da metáfora visual são também abordados neste capítulo, sendo, este, concluído com a descrição de alguns cenários de utilização.

O quinto capítulo corresponde ao Desenvolvimento Prático do projeto, pelo que, relata a evolução do design, a interação aplicada à visualização de dados, as tecnologias necessárias à implementação e, por fim, detalhes técnicos da implementação, incluindo problemas e soluções encontradas.

O sexto capítulo é referente à Avaliação da usabilidade do projeto e aborda a metodologia utilizada para a mesma. Analisa de modo geral os participantes e enumera as questões em pesquisa. As tarefas do teste são explicadas e, finalmente, são analisados os resultados dos testes e sugeridas melhorias.

O sétimo capítulo, referente à Conclusão e perspectivas futuras resume o trabalho efetuado e descreve o que se pretende realizar futuramente.

Segue-se a Bibliografia e a Bibliografia de imagens, estas, enumeram todas as referências consultadas ao longo da investigação e as fontes das imagens utilizadas.

Este documento finaliza com os Anexos – do A ao E – referentes aos primeiros esboços das formas gráficas para a representação dos dados eleitorais, ao *Design Statement*, ao formulário realizado para os testes de usabilidade, ao resumo das respostas ao formulário e à primeira proposta desenvolvida para a ambientação ao d3, respetivamente.

Estado da arte

Este capítulo tem como objetivo apresentar ao leitor o estado da arte e trabalhos relacionados com os temas nos quais a dissertação incidirá. Começaremos por introduzir a evolução das formas gráficas para a representação de dados, desde a representação cartográfica auxiliar às explorações marítimas, até aos gráficos e tecnologias utilizados nos dias de hoje – subcapítulo 2.1. Seguidamente faremos um levantamento de que forma a interação se insere na área da visualização e como esta possibilita a criação de interesse num possível utilizador – subcapítulo 2.2. Será feito um levantamento dos bons princípios para o design da informação – subcapítulo 2.3 – e, no último – subcapítulo 2.4 – faremos uma descrição e uma breve análise semiótica em projetos relacionados.



2.1 Evolução das Formas Gráficas na Representação de Dados

A história da visualização de informação, como uma área de conhecimento, tem menos de duas dezenas de anos até ao ano corrente¹. No entanto, a sua origem teve início com o registar das posições das estrelas e de outros corpos celestes em tabelas, bem como no desenho de mapas, cujo objetivo era auxiliar a navegação e as explorações.

Os topógrafos egípcios usavam uma ideia primordial de coordenadas, que localizava posições terrestres e celestes, em algo semelhante a latitude e longitude cerca de 200 A.C.

Uma das primeiras representações gráficas de informação quantitativa pode ser observada na figura 2.1. Esta, ilustra um gráfico datado do séc.X, que utiliza os dois eixos – horizontal e vertical – para representar a mudança na inclinação das órbitas planetárias de sete corpos celestes em função do tempo². Utiliza, portanto, um plano bidimensional, no qual são apresentadas múltiplas variáveis que alteram o seu valor à medida que o plano horizontal, representativo do tempo, avança.

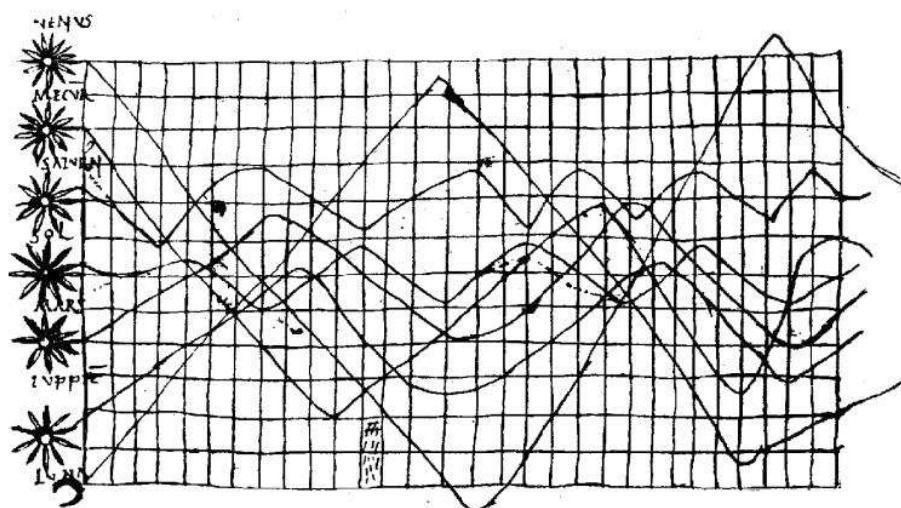


Figura 2.1^a Movimento planetário representado através de inclinações cíclicas. Desenhado por um astrónomo desconhecido, aparentemente no séc.X.

Aparece, nesta época da história da visualização, como um prodígio isolado, sendo a primeira tentativa de mostrar graficamente valores variáveis, tendo o gráfico tempo-série seguinte surgido apenas 800 anos mais tarde³.

No séc. XIV apareceu, num trabalho de Nicole Oresme, a ideia teórica de gráfico de barras e a sua relação lógica com valores tabulados⁴. No entanto, foi William Playfair o primeiro a pôr esta teoria em prática, criando o primeiro gráfico de barras, nascido da necessidade da representação de dados sobre importação e exportação com a presença da dimensão temporal⁵.

Na figura 2.2 observamos o gráfico “*Exports and Imports of Scotland to and from different parts for one year from Christmas 1780 to Christmas 1781*” que, como o título sugere, representa as exportações e importações partilhadas entre a Escócia e diversos países durante um período temporal correspondente a um ano.

Playfair defendia que a representação gráfica dos dados era preferível a tabelas, pois estas, permitiam mostrar a forma gráfica dos dados e facilitar a sua comparação.

¹ Chen, 2006

² Friendly, 2006

³ Tuffte, 1995

⁴ Meirelles, 2013

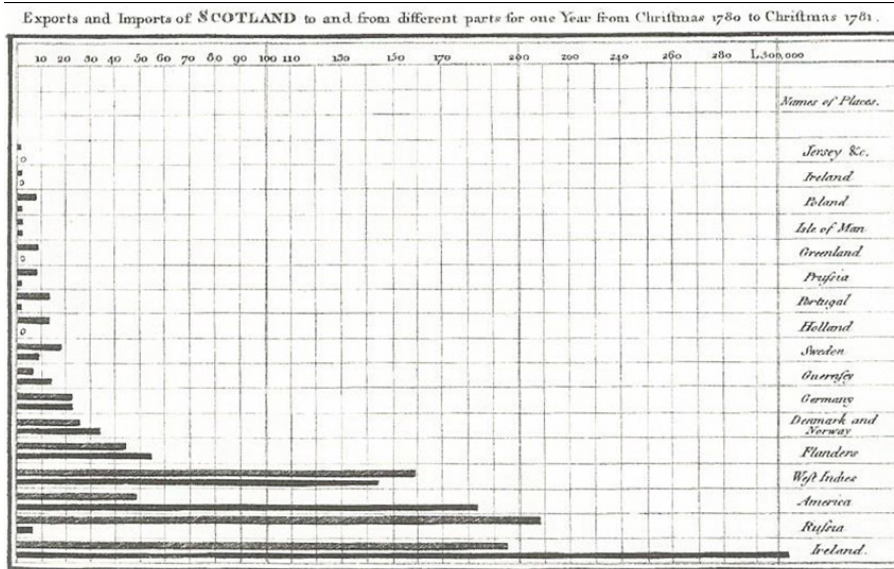


Figura 2.2^b As divisões verticais estão em intervalos de dez mil libras. As linhas pretas representam a Exportação e as linhas texturadas, a Importação.

No caso deste gráfico, não há uma ideia distinta e os dados quantitativos não possuem forma. Portanto, segundo Playfair, este conjunto diminuto de dados não comparativos pertence a tabelas.

O séc.XVII vem instigar grandes avanços à teoria e aplicação prática da visualização de informação. Este período é considerado o nascer do pensamento visual, no qual foram disponibilizados os elementos necessários para o desenvolvimento de métodos gráficos. A ascensão da geometria analítica e do sistema de coordenadas, de teorias de erros mensuráveis e estimativa, o nascimento da teoria da probabilidade e o início da estatística demográfica são exemplos dos acontecimentos que ficaram para a história da visualização de informação.

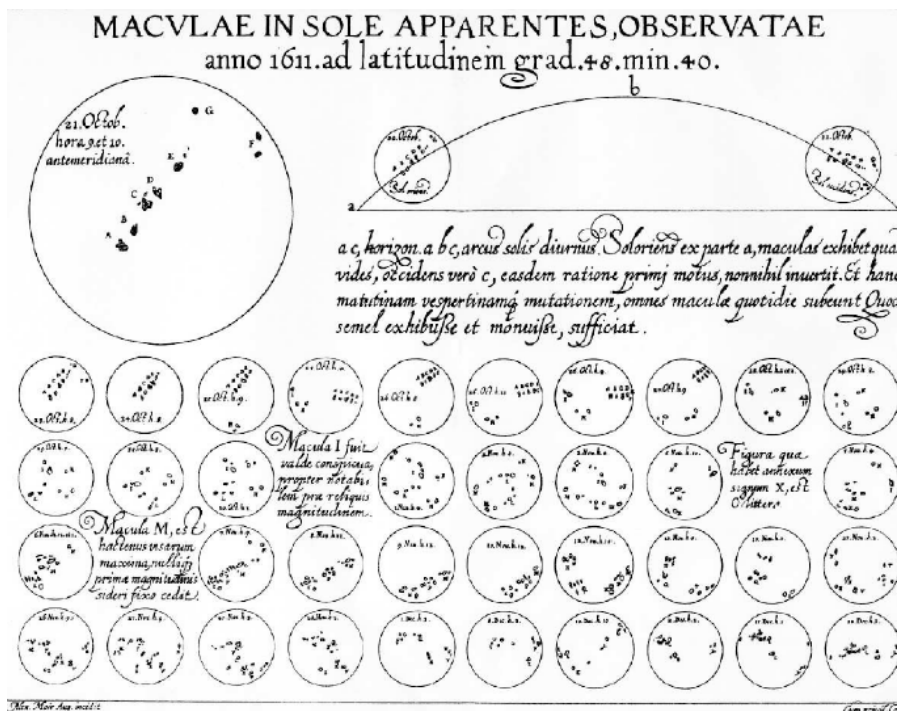


Figura 2.3^a Representação das manchas solares em função do tempo através de pequenos múltiplos.

Também nesta altura, é introduzida, por Christopher Scheiner², a ideia atualmente conhecida por pequenos múltiplos. Estes, que consistem na representação de dados em função do tempo, a partir de várias pequenas representações gráficas semelhantes, facilitam a comparação dos dados em cada instante. C. Sheimer utiliza este tipo de visualização para mostrar mudanças na configuração de manchas solares ao longo do tempo, como se pode ver na figura 2.3.

Os mapas temáticos têm o seu início em meados deste século, exibindo mais do que apenas localizações geográficas. No entanto, os grandes avanços nos métodos gráficos acontecem, maioritariamente, no séc. XIX, entre 1820 e 1860.

Edmond Haley⁴ cria, em 1700, o primeiro gráfico temático, no qual reproduz graficamente os campos magnéticos da Terra, fazendo uso inédito de *isolines*.



Figura 2.4ª "A New and Correct Chart Showing the Variations of the Compass in the Western & Southern Ocean as Observed in the Year 1700", publicado em 1701.

Igualmente neste período, surgiu o primeiro gráfico *time-series*, que representava o tempo, horizontalmente, da esquerda – datas mais antigas – para a direita – datas mais recentes. Esta inovação foi trazida primeiramente por Jaques Barbeau-Dubourg, que cria um gráfico de cerca de um metro e sessenta de comprimento, para a representação

² Friendly, 2006

⁴ Meirelles, 2013

de dados históricos relativos ao período correspondente à origem dos acontecimentos até ao seu tempo presente.

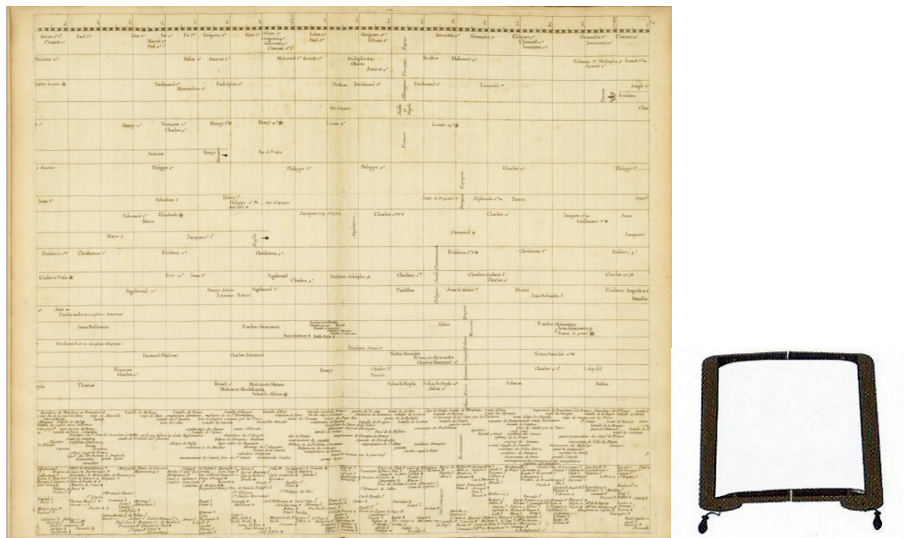


Figura 2.5° "Chronographie Universelle" de Jaques Barbeau-Dubourg. Representa um total de 6480 anos usando, pela primeira vez, uma escala temporal uniforme, em que cada ano corresponde a dois milímetros e meio. Ao lado, podemos observar um aparelho de *scroll* horizontal, criado por Barbeau-Dubourg, com o intuito de facilitar a navegação no documento.

John Priestley⁴ faz também inovações neste campo, usando linhas para representar duração, implementando-as no gráfico "Chart of Biography" – figura 2.6 – , que abrange cerca de 3000 anos e a vida de 2000 pessoas famosas. São utilizadas "linhas de vida" para representar a duração da vida de cada pessoa.

Priestley afirma que, fazemos uso de expressões espaciais, como "longo" ou "curto", para descrever períodos de tempo, que se refletem na representação visual das linhas que, intuitivamente, quantificam um período temporal⁴.

Analisando a figura 2.6, podemos observar uma certa preocupação com a ordenação dos dados, para um melhor entendimento dos mesmos por parte do observador.

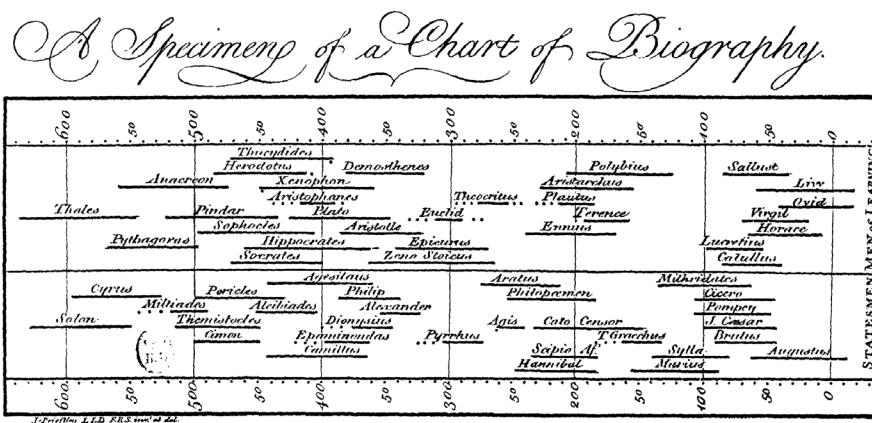


Figura 2.6° "Chart of Biography" de John Priestley publicado em 1765. Para representar dados incertos da vida das pessoas famosas, são colocados pontos no início ou no fim da linha de vida.

⁴ Meirelles, 2013

Os nomes estão organizados por seis temas – Historiadores, Antiquários e Advogados; Oradores e Críticos; Artistas e poetas; Matemáticos e Físicos; Teólogos e Metafísicos; Homens do estado e Soldados –, divididos por linhas e ordenados de cima para baixo e de acordo com o grau de importância, sendo que a ordem se inicia em baixo com os Homens de Estado

Atualmente, os gráficos de barras e gráficos “tarte” – *pie charts* – são amplamente usados para visualizar dados estatísticos, facto que devemos a William Playfair². Estas inovações de Playfair acontecem entre o séc.XVII e o séc.XVIII, com o seu primeiro gráfico de linhas e gráfico de barras de dados económicos, bem como do gráfico de “tarte” e o gráfico circular⁵.

Embora a ideia do registo e enumeração da população ter os seus primórdios nos tempos egípcios, romanos e gregos – que usavam os dados para propósitos administrativos – a primeira coleção sistemática de dados sociais teve início nos finais do séc.XVIII, com o primeiro censo populacional, realizado pela Suécia em 1749, seguida por outros países anos mais tarde⁴.

Apesar das inovações de William Playfair, o uso de gráficos para ilustrar ou analisar dados não pertencentes ao campo matemático e das ciências físicas, foi um fenómeno raro até meados do séc.XIX⁴. Este período, principalmente a primeira metade do séc.XIX, é responsável pelo crescimento explosivo nos gráficos estatísticos e mapas temáticos, preciosidades gráficas herdadas dos séculos anteriores⁴.

A invenção da maioria das formas gráficas modernas, largamente utilizadas nos dias de hoje, teve o seu nascimento nesta época, resultado do aparecimento de gráficos, cujos dados abrangiam vastas e novas variedades de tópicos⁵.

Em 1826, Charles Dupin⁴ cria o primeiro mapa estatístico, exibindo através de escalas de cinzas, desde o preto ao branco, a distribuição do analfabetismo em França, sendo inédita a representação gráfica de fenómenos sociais.

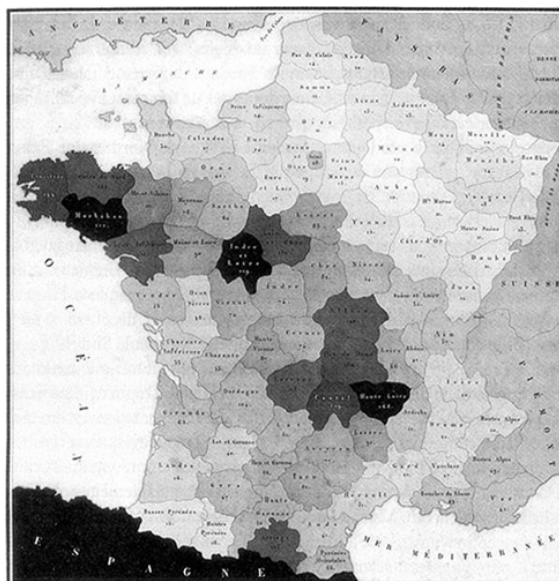


Figura 2.7^c Conhecido por ser o primeiro mapa cloroplético, no qual cada valor único é representado por um único valor de cinza.

² Friendly, 2006

⁴ Meirelles, 2013

⁵ Cruz, 2010

Em 1870, quase todos os países europeus e também os Estados Unidos estavam sistematicamente a coletar, analisar e disseminar estatísticas oficiais do governo sobre a população, o comércio e problemas políticos e sociais, em publicações como atlas estatísticos, exposições internacionais e conferências⁴.

Entre 1835 e 1876, ocorreram oito congressos internacionais de estatística, com grande importância para a visualização de informação, pois discutiram e promoveram o uso de métodos gráficos. Houve, também, uma tentativa de estabelecer padrões internacionais⁴.

O desenvolvimento sem precedentes dos métodos gráficos para a análise de dados, que acontece nos meados do séc.XIX, resulta do reconhecimento, por parte de maioria dos países, da importância da informação numérica no planeamento do bem-estar geral da população. Este período marca também o nascimento de novas disciplinas como estatística, geologia, biologia e economia e é considerada a idade do entusiasmo, mais conhecida por “*Época Áurea*”. Isabel Meirelles nota que “atualmente experimentamos um fenómeno semelhante alimentado pela coleção de dados digitais de todo o tipo e a necessidade de os analisar”⁴.

Num período marcado pelo ponto alto da aplicação de temas humanísticos a cartografias temáticas, surge o famoso mapa de pontos – “*dot map*” – criado por John Snow², em 1855 e que exhibe as mortes causadas pela cólera no centro de Londres, em Setembro de 1854. Através da análise da distribuição das mortes sobre o mapa de Londres, Snow consegue identificar a causa da doença, constatando que a maioria das habitações dos falecidos se situa perto da bomba de água da Broad Street.



² Friendly, 2006

⁴ Meirelles, 2013

Figura 2.8^d “Dot map” criado por Dr. Snow, no qual as mortes de pessoas causadas por cólera são representadas por pontos e as bombas de água por cruzes.

O mapa de Snow tem valor não só pelo propósito de prevenir mais mortes, “mas, também, pela capacidade de mostrar padrões de dados de forma visual, que elucida sobre a eficiência da análise gráfica”⁵.

De acordo com Funckhouser², Charles Joseph Minard é o “Playfair da França”. Minard produz gráficos prodigiosos, entre eles, um muito conhecido e intitulado por “*Carte figurative of the Napoleon’s Russian campaign of 1812*”. Desenhado em 1869, combina mapas e séries temporais, criando uma narrativa em torno das perdas sofridas durante a campanha de Napoleão, na Rússia, em 1812. Nele é removida a maioria da informação geográfica e mantidas, apenas, localizações geográficas relevantes para o gráfico, como é o caso dos rios principais e das cidades. Assim, são representadas seis variáveis: o tamanho do exército, a sua localização no plano bidimensional, a direção do movimento do exército (viagem de ida e regresso) e a temperatura, em várias datas durante o regresso.

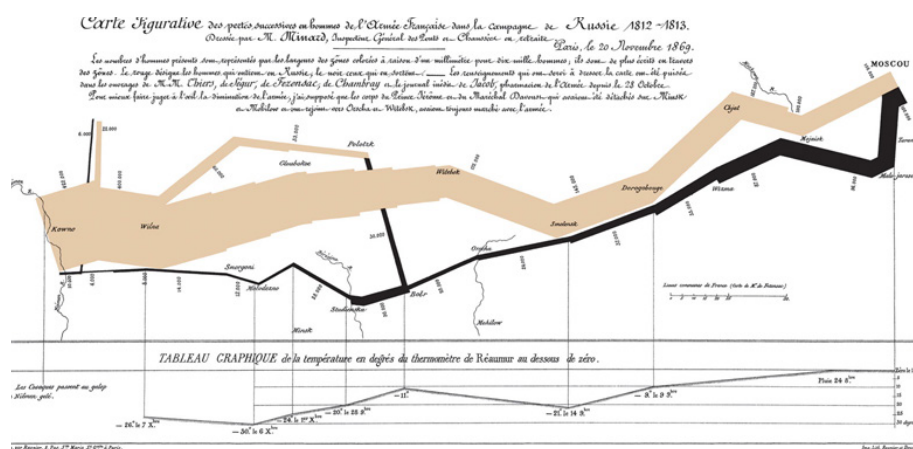


Figura 2.9^o “Napoleon March to and from Russia, 1812-1813”. Combina dados estatísticos com uma *timeline* e dados espaço-temporais acerca do exército francês. A largura da linha é representada de acordo o número de soldados, na qual cada milímetro corresponde a 10 000 homens.

A campanha de Napoleão começa com cerca de 420 000 soldados na linha bege e acaba com 10 000 soldados na linha preta, valores que são dados por Minard ao longo das duas linhas, face à incapacidade visual humana de quantificar absolutamente áreas. O gráfico de linhas, em baixo, representa as temperaturas enfrentadas pelos soldados no caminho de regresso, estando relacionada com a linha preta. Esta relação possibilita a extração de novo conhecimento quanto às causas de morte dos soldados. Por exemplo, 22 000 soldados morreram, atravessando o rio *Berezina*, devido às baixas temperaturas.

Minard³ mostra, com este exemplo, como a complexidade das multivariáveis pode ser integrada de forma subtil na arquitetura do gráfico, de tal forma que os observadores nem dão conta da quantidade de variáveis que estão a analisar. Fazendo um uso acertado das habilidades cognitivas do observador, é possível aumentar o poder explicativo do gráfico, que conta uma história rica e coerente, de uma maneira que apenas números com uma forma visual têm a capacidade de fazer.

Tuffte afirma até que, este, pode bem ser um forte candidato para melhor gráfico estatístico alguma vez desenhado³.

Apesar de Florence Nightingale⁴ não ter inventado os gráficos rosa – *rose*

² Friendly, 2006

³ Tuffte, 1995

⁴ Meirelles, 2013

⁵ Cruz, 2010

charts – o seu diagrama das causas da mortalidade no exército no Leste é um marco histórico da visualização de informação em usos políticos e sociais ou, mais exatamente, no campo da medicina. Nightingale desenha o gráfico com o objetivo de conseguir melhores condições sanitárias para o tratamento de soldados no campo de batalha. Em semelhança ao que acontece com o gráfico de Snow, a representação gráfica de Nightingale, comprova a causa da mortalidade dos soldados, sem deixar qualquer dúvida de que a quantidade de soldados que morreram de doença durante a recuperação de feridas era largamente superior ao número de mortes às mãos do inimigo. Mais um exemplo eficiente do poder dos gráficos para suportar conclusões, o qual convenceu o governo Britânico da necessidade de um sistema de saúde melhor.

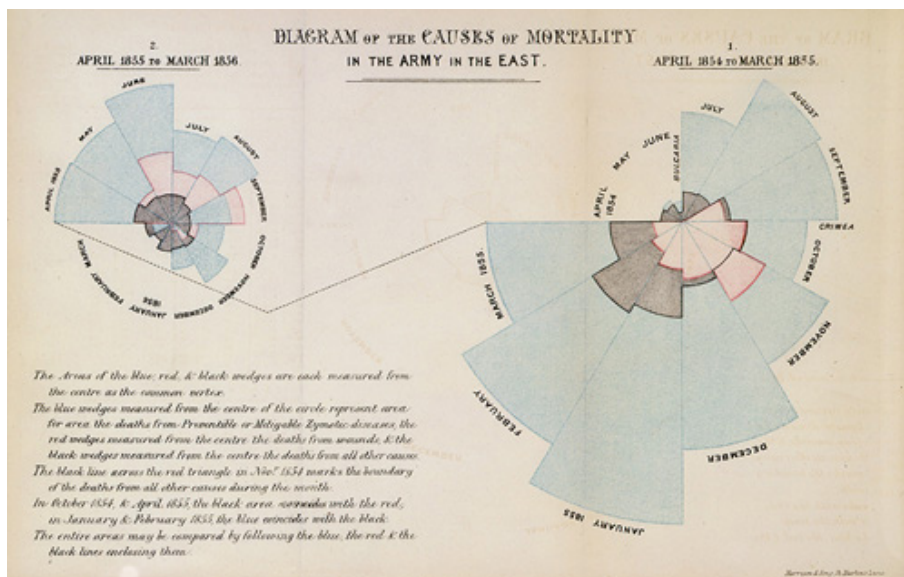


Figura 2.10^f "Diagram of the Causes of Mortality in the Army in the East" criado por uma enfermeira britânica (Florence Nightingale) e publicado em 1858. É um gráfico polar, no qual o raio de cada fatia representa o número de mortes em cada mês. A cor é usada para diferenciar a causa da morte: azul para as doenças evitáveis, vermelho/rosa para feridas da guerra e preto para as fatalidades.

A primeira metade do séc.XX é marcada por alguma pobreza em inovações gráficas, relativamente ao século anterior, ficando o período conhecido por "modern dark ages". O entusiasmo pela visualização, característico do século anterior, foi substituído pelo crescimento da quantificação e de modelos formais, frequentemente estatísticos. Apesar da falta de inovações gráficas, este é um período que populariza os gráficos, através da aplicação dos criados anteriormente. É, assim, visto como um período de divulgação, mais do que de inovação².

Em 1917, Henry Gantt² cria um gráfico para gerir tarefas de projetos, atualmente bastante popular e reconhecido pelo nome de "Diagrama de Gantt". Este diagrama permite a entrada das tarefas relativas a um projeto e mostra a duração da execução das mesmas através de uma barra, cujas extremidades representam o dia de início e o dia de fim de uma única tarefa.

Outro famoso exemplo desta altura é o mapa das linhas do metro subterrâneo de Londres, criado por Harry Beck⁵, em 1931. Primariamente recebido com alguma apreensão por parte do departamento de publicidade face à sua mudança gráfica radical, rapidamente ficou popular entre os passageiros londrinos. Isto mostrou que era exatamente aquilo que o público queria e precisava⁶.

² Friendly, 2006

⁵ Cruz, 2010

⁶ TRANSPORT FOR LONDON, n.d.

Beck percebe-se que, quando estamos no subsolo, a localização geográfica perde a importância, bastando uma sequência correta das estações para uma fácil orientação⁵. Altera a escala real, colocando as estações à mesma distância umas das outras e retira toda a informação geográfica a favor da legibilidade. As curvas sinuosas são substituídas por linhas direitas, cujas inclinações usam apenas ângulos de 90° e 45°⁷. É um gráfico que encontra a sua inspiração nos circuitos elétricos, resultando num diagrama de linhas coloridas que se cruzam e, cujo design, vem facilitar a formação de uma imagem mental da posição relativa das estações do metro, possibilitando um melhor planejamento das viagens e encorajando os cidadãos a utilizar este meio de transporte⁵.

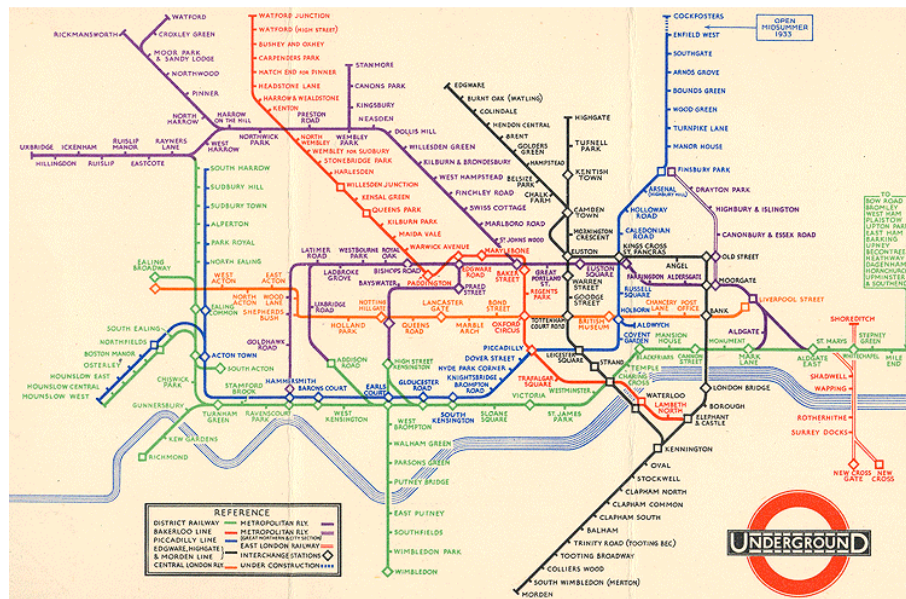


Figura 2.11 ⁹ "London Underground Transport". Criado por Harry Beck, em 1931, é hoje uma das representações *standard* para os caminhos subterrâneos de metros em todo o Mundo.

Entre 1950 e 1975 – já na segunda metade do séc.XX, mais exatamente no primeiro quarto – a visualização desperta para uma época considerada o Renascimento da visualização de dados.

Tukey² reconhece a importância da análise de dados como um ramo legítimo da estatística, distinta da estatística matemática, e suporta esta posição com uma larga variedade de novas apresentações gráficas simples e eficazes, criadas por ele durante este período.

Em 1967, Jacques Bertin⁸ – um cartógrafo francês – publica na França um livro intitulado de "*Semiology of Graphics*". Este livro é um trabalho fundamental no campo do design e cartografia, que sintetiza os princípios da comunicação gráfica, com regras padrão aplicadas à escrita e à topografia.

A criação da primeira linguagem de programação foi também um marco importante desta época, pela capacidade de processar computacionalmente dados estatísticos. Fortran⁹ é a primeira linguagem de alto nível que exigia menos pensamento, menos escrita e que, portanto, facilitava a sua aprendizagem e exigia menos tempo para obter os resultados pretendidos. Mesmo assim, a sua utilização limitava-se apenas aos computadores universitários, estando indisponível ao uso comum.

² Friendly, 2006

⁵ Cruz, 2010

⁷ Fry, 2008

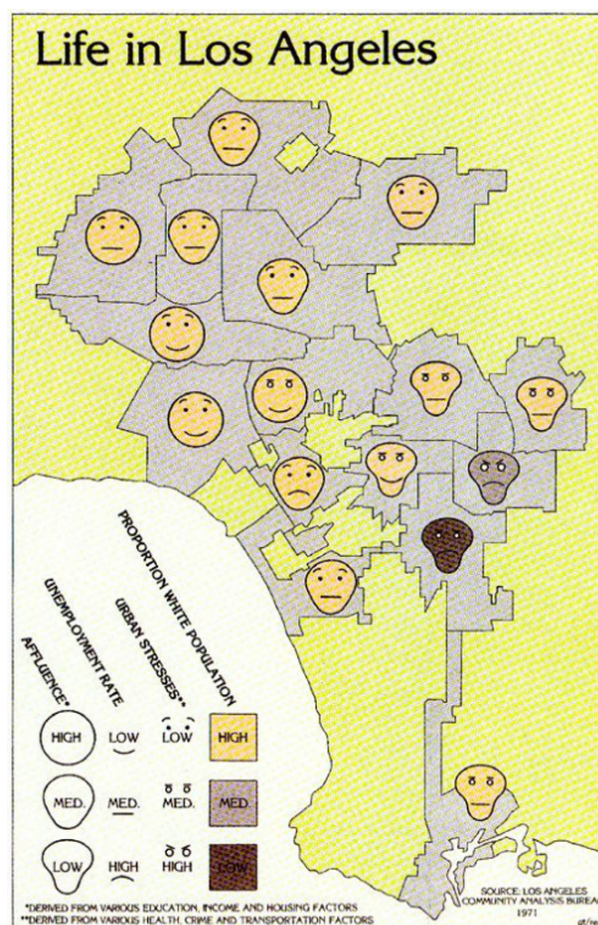
⁸ Bertin, 2011

⁹ Backus, Beeber & Best, 1957

Estas novas técnicas no campo da visualização de informação viabilizaram a construção de novas abordagens gráficas, como é o caso das aplicações interativas e gráficos de alta resolução². “*The Use of Faces to Represent Points in k-Dimensional Space Graphically*” é um artigo publicado em 1973, por Herman Chernoff¹⁰, que põe em prática estas novas técnicas computacionais e, no qual, é apresentado um novo modelo de representação de dados multivariáveis.

O modelo de Chernoff baseia-se na ilustração de caras, cujas características se vão alterando conforme os dados a representar. Atualmente conhecidas como “*Chernoff Faces*”, o observador identifica facilmente padrões, grupos e correlações entre os dados. A sua eficiência tem origem na habilidade inata da mente humana em reconhecer pequenas diferenças nas características faciais e as assimilar de uma só vez. Assim sendo, a cada uma das variáveis é atribuída uma característica facial, para que cada cara seja gerada de acordo com os dados.

Em 1979, o Dr. Eugene Turner¹¹ produz um mapa intitulado “*Life in Los Angeles*” que aplica o modelo de Chernoff para representar as condições de vida nas diferentes regiões de Los Angeles. São mapeadas quatro variáveis: afluência, taxa de desemprego, stress urbano e percentagem de população “branca”, atribuídas a características como forma da cara, curvatura da boca, arqueamento da sobrancelha e cor da cara respetivamente.



² Friendly, 2006

¹⁰ Chernoff, 1973

¹¹ Spinelli, Joseph

G., Zhou, 2004

Figura 2.12^h “*Life in Los Angeles*” criado por Dr. Eugene Turner.

O último quarto do séc.XX veio alastrar a área de pesquisa da visualização de dados. Ferramentas computacionais foram desenvolvidas para um amplo leque de métodos de visualização, que permitiram a exibição de dados de variadas disciplinas e foram, finalmente, disponibilizadas para uso pessoal, deixando de estar limitados às universidades.

Os desenvolvimentos na área da visualização de informação, desta época, aconteceram rapidamente. Entre estes, está o desenvolvimento de sistemas estatísticos computacionais altamente interativos, a emergência de novos paradigmas de manipulação direta para análise de dados visuais e de métodos para a visualização de dados de grandes dimensões e o aumento da atenção dada aos aspetos cognitivos e perceptuais dos dados exibidos. Todos estes progressos nos métodos e nas técnicas dependeram muito dos avanços atingidos na área tecnológica, a qual aumentou a capacidade e a velocidade do processamento computacional, o que possibilitou o desenvolvimento de métodos computacionais capazes de aceder e processar dados massivos, ou fluxos de dados, em tempo real².

Dentro dos progressos feitos nesta área, está uma técnica gráfica desenvolvida por George W. Furnas¹², em 1981, e conhecida por “*Fisheye View*”: uma técnica bastante útil para aplicação em exibições gráficas computacionais que representam uma grande quantidade de dados. Em muitos destes casos, os utilizadores, são forçados a analisar uma grande quantidade de dados em janelas, muitas vezes pequenas. Isto constitui um grande problema, sendo que existe demasiado detalhe, o que dificulta a análise centrada num ponto particular da representação gráfica. A lente *Fisheye* vem resolver este problema, permitindo ao observador concentrar-se em determinados pontos em grande detalhe, enquanto o resto da visualização é mostrada com menor detalhe.

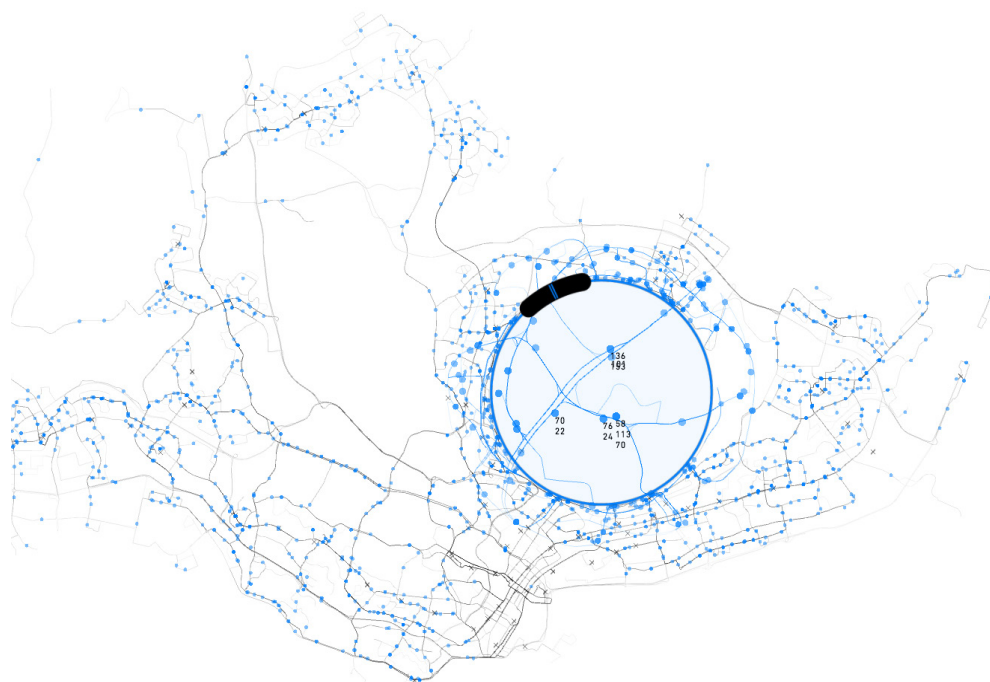


Figura 2.13ⁱ Pormenor da lente *Fisheye* utilizada no projeto *Data Lenses*, realizado por Pedro Cruz, em 2012.

² Friendly, 2006

¹² Furnas, 1986

2.2 Interação na Visualização de Informação

A interação é uma parte essencial da visualização de informação na medida em que, esta existe, mesmo com uma imagem estática. Ao observar um poster, por exemplo, o leitor frequentemente executa naturalmente várias tarefas interativas como: rodar o poster, aproximar ou afastar do olhar ou escrever notas no poster¹³, numa tentativa de perceber o significado da representação gráfica.

Visualizações interativas têm a capacidade de incentivar o utilizador a explorar os dados, de uma forma que as imagens estáticas não o fazem. Quando bem elaboradas, estas visualizações têm a capacidade de imergir o utilizador na navegação e exploração dos dados, de uma maneira que se assemelha à imersão induzida por jogos¹⁴. No entanto, apesar de ter um papel importante na visualização de informação, a interação na visualização de informação é um campo que continua a explorar funções básicas que pouco mudaram desde 1996, quando Ben Shneiderman¹⁶ da Universidade de Maryland propôs um mantra de procura: ‘Visão geral primeiro, *zoom* e filtro, e então detalhes sobre a procura’. Este princípio surge da sua identificação constante em vários projetos e, atualmente, está presente na maioria das visualizações interativas, tendo um grande sucesso ao satisfazer diferentes públicos – desde o público que apenas explora o conjunto de dados através de uma navegação quase aleatória, até ao público que navega na visualização procurando dados capazes de responder a uma pergunta específica¹⁴.

Os seres humanos possuem habilidades perceptivas notáveis que são utilizadas ao interagir com este tipo de visualizações interativas – facilmente conseguem fazer uma ‘digitalização’ mental – ou seja, memorização da imagem – que permite reconhecer ou recordar imagens rapidamente e detetar mudanças no tamanho, forma, movimento, cor ou textura¹⁶.

Em 1996, Ben Shneiderman¹⁶ dá o seu contributo para a investigação da interação na visualização de informação, propondo uma taxonomia de tarefas e definindo sete tarefas de alto nível de abstração – “vista geral”, “*zoom*”, “filtragem”, “detalhar a procura”, “relacionamento”, “histórico” e “extração”- que permitem ao utilizador, fazer uma exploração controlada dos dados. De acordo com as definições do autor, a “vista geral” permite ao utilizador obter uma visão global de toda a coleção de dados. Uma das estratégias que aplica este conceito é o *zoom out* e o autor refere também a ferramenta *fisheye*. A ferramenta *fisheye* possui características que permitem uma visão geral do conjunto de dados, ao mesmo tempo que o utilizador pode ver outros dados detalhadamente, o que acontece através da distorção da representação dos dados.

O “*zoom*” permite ao utilizador controlar o foco e o fator de *zoom* para obter uma vista com especial enfoque nos dados que lhe interessam, sem que a noção de posição e contexto seja afetada. Esta ampliação pode ser feita através de barras de *zoom*, ou alterando o tamanho do campo de uma caixa de visualização, que controla a área da visualização que aparece no ecrã. De acordo com Bederson e Hollan¹⁶ a forma mais gratificante de realizar esta tarefa é através da deslocação do rato até ao local de interesse e emitir um comando de *zoom*, clicando num botão do rato durante o tempo que o utilizador desejar. Nos dias de hoje, esta técnica está bastante desenvolvida na medida em que existem várias ferramentas que nos permitem este tipo de interação

¹³ Yi, Kang, Stasko, & Jacko, 2007

¹⁴ Murray, 2012

¹⁶ B. Shneiderman, 1996

através da localização do rato e da utilização do botão de *scroll* do rato, com a vantagem de não ser necessário manter o botão pressionado. Os dispositivos *touch* são um dos exemplos tecnológicos que têm este tipo de interação presente, ao permitir o *zoom* de áreas, apenas com um gesto com os dedos.

A “filtragem” permite a utilização de consultas dinâmicas para que os utilizadores possam escolher o âmbito de informação a ser exibida no ecrã e concentrarem a sua atenção nos dados que lhes interessam, ao eliminar todos os outros dados que não estejam relacionados com a consulta realizada.

A tarefa “detalhar a procura”, permite a seleção de um elemento ou grupo de interesse e obter novas informações. Como anteriormente, a coleção de dados a ser visualizada é diminuída através da filtragem e a navegação e exploração dos detalhes dos dados individuais, ou de um grupo de dados, torna-se mais fácil. Na altura em que Shneiderman escreveu este artigo – “*The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations*” –, a maneira mais utilizada para oferecer os detalhes consistia na abertura de uma janela *pop-up* que mostraria os valores dos atributos do elemento selecionado.

A tarefa “relacionar” permite a visualização de dados que estão relacionados ou mostrar a relação entre os dados.

O “histórico” suporta ações como refazer ou desfazer para que o utilizador possa ter oportunidade de navegar pelos dados sem ter de voltar ao início, por exemplo. Assim, manter o histórico da interação é importante, pois a exploração de informação é um processo com várias etapas.

Finalmente, a “extração” consiste no conceito de que o utilizador possa extrair a informação das suas consultas e guardar essa informação em ficheiros, cujo formato facilite a sua utilização noutro tipo de tarefas, como por exemplo: envio por email, impressão ou apresentação.

Mais tarde, já em 2007, Yi et al.¹³ define sete categorias de técnicas interativas, que diz serem muito utilizadas na visualização de informação e que estão organizadas em torno da intenção de um utilizador – “selecionar”, “explorar”, “reconfigurar”, “codificar”, “abstrair/detalhar”, “filtrar” e “conectar”. Algumas destas, podem ser relacionadas com as tarefas de Shneiderman – figura 2.14.

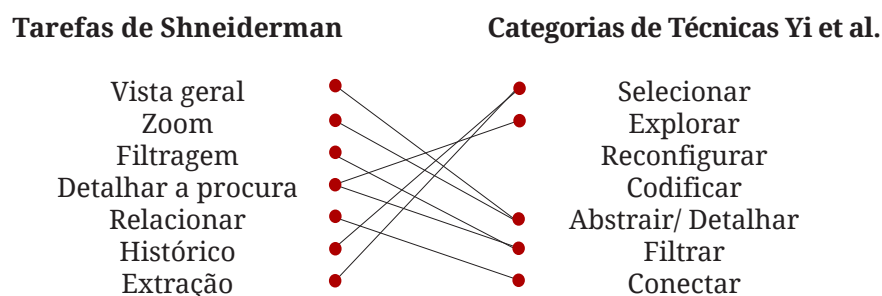


Figura 2.14 Relacionamento entre as tarefas de Shneiderman e as categorias de técnicas definidas por Yi et al.

¹³ Yi, Kang, Stasko, & Jacko, 2007

Para cada uma das sete categorias apresentadas, Yi et al. define uma pequena frase explicativa:

- Seleccionar: marcar algo como interessante
- Explorar: mostrar algo mais
- Reconfigurar: mostrar uma composição diferente
- Codificar: mostrar uma representação diferente
- Abstrair / Detalhar: mostrar mais ou menos detalhes
- Filtrar: mostrar algo condicionalmente
- Conectar: mostrar os dados relacionados

Analisemos uma definição mais elaborada destas categorias para um melhor entendimento das mesmas.

Seleccionar: marcar algo como interessante.

As técnicas interativas de seleção fornecem aos utilizadores a possibilidade de marcar dados de interesse e manter o foco nos mesmos. Quando demasiados dados são apresentados numa representação, ou quando as representações são alteradas, é difícil para os utilizadores seguirem os dados nos quais estão interessados. A técnica de seleção permite destacar visualmente o dado selecionado pelo utilizador e facilmente localizá-lo na representação, quer ela se altere ou existam demasiados dados.

Explorar: mostrar algo mais.

As técnicas interativas de exploração permitem aos utilizadores examinar um subconjunto diferente de dados. Frequentemente, um sistema de visualização de informação apenas exhibe um número limitado de dados de cada vez, principalmente quando trabalha com uma quantidade de dados em larga escala, devido às dimensões limitadas do ecrã e das limitações percetivas e cognitivas humanas do processamento de informação. Normalmente, o utilizador examina primeiro um subconjunto de dados para compreender a representação dos dados, o que possibilita, de seguida, uma exploração mais fácil de novos dados. Yi et al refere a técnica de *panning* como uma das técnicas interativas de exploração mais comuns, na qual há a movimentação de uma câmara através de uma cena ou a movimentação da cena enquanto a câmara permanece quieta. Normalmente, executada quando o utilizador pega na cena e a move com o rato através de arrastamento, também conhecido por *dragg and drop*, ou através do uso de barras de *scroll*, conhecido por *direct-walk*.

Reconfigurar: mostrar uma composição diferente.

Técnicas interativas de reconfiguração oferecem perspetivas diferentes para um conjunto de dados, alterando o arranjo espacial das representações. Perspetivas diferentes do mesmo conjunto de dados facilitam a deteção de características ocultas dos dados e as relações entre eles. Assim, estas técnicas permitem ao utilizador alterar a organização dos dados ou o seu alinhamento com vista a criar novas perspetivas. Ainda dentro desta categoria, podemos identificar técnicas que permitem uma movimentação dos dados mais livre, para que a composição criada se assemelhe mais ao modelo mental criado pelo utilizador e, ainda, técnicas que permitem a redução de obstruções quando existem demasiados dados numa mesma área, podendo haver sobreposições que enganam o utilizador, ao deixarem ver apenas o elemento que ficou por cima. Uma dessas técnicas, é conhecida por *jitter* – figura 2.15 – e cujo funcionamento se baseia na deslocação aleatória através de um pequeno incremento espacial, o que oferece uma noção melhor da densidade de dados nessa região.

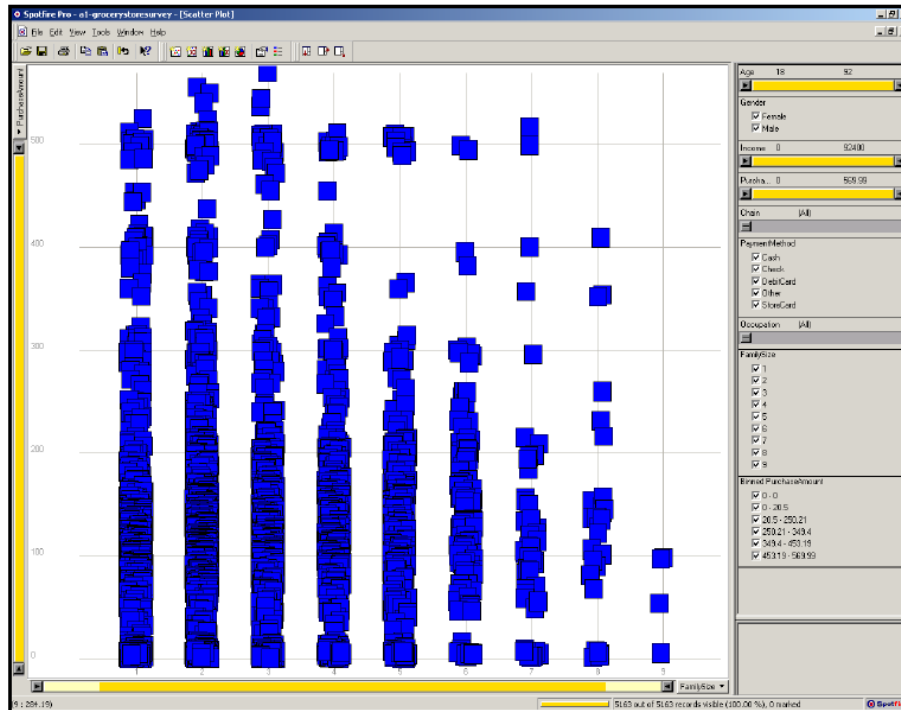


Figura 2.15 | Resultado da aplicação da técnica *Jitter* no projeto *Spotfire*.

Codificar: mostrar uma representação diferente.

As técnicas interativas de codificação servem para que o utilizador possa alterar a representação dos dados, se assim o desejar. Esta alteração reside na representação visual de cada elemento que, normalmente, possui atributos como cor, forma e tamanho. Os elementos visuais desempenham um papel muito importante na visualização de informação – como iremos estudar melhor no próximo subcapítulo – pois estes, afetam a cognição pré-atentiva e estão relacionados com a forma como os utilizadores entendem ou percebem as relações e distribuições dos dados. Os sistemas de visualização de informação que oferecem este tipo de técnicas permitem, habitualmente, a alteração do esquema de cores, do tamanho, orientação, forma ou até mesmo da fonte tipográfica.

Abstrair/ Detalhar: mostrar mais ou menos detalhe.

As técnicas que permitem este tipo de interação servem para que o utilizador possa controlar o nível de abstração de uma representação de dados, sendo que a representação pode mostrar uma vista mais geral ou mais detalhada de uma área ou de um conjunto de dados em particular. O *zoom* é uma destas técnicas que permite ao utilizador alterar a escala da representação, sem que a representação visual seja alterada ou distorcida, apenas replicando a ação de aproximação ou afastamento utilizada também em imagens estáticas.

Filtrar: mostrar algo condicionalmente.

Técnicas de filtragem possibilitam a escolha dos conteúdos a serem representados mediante condições especificadas numa consulta. Os dados que não satisfaçam esta condição, formulada pelo utilizador, ficam escondidos ou são apresentados de forma diferente, sem que haja alteração aos dados, que podem voltar a ser recuperados mais tarde com uma consulta certa. Este tipo de interação, baseado no paradigma da manipulação direta dos dados, ajuda na construção de um sistema mais responsivo, utilizando *queries* textuais orientadas

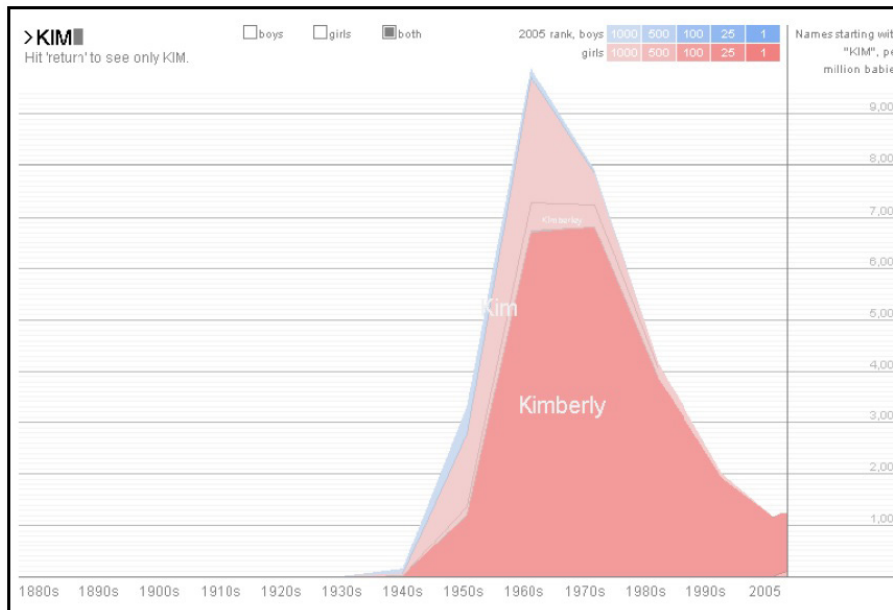


Figura 2.16 Resultado da pesquisa de nomes começados por "KIM" no projeto "Name Voyager" de Martin Wattenberg, 2005.

– figura 2.16 – que podem ser controladas através de *alpha sliders*, *range sliders* e botões de alternância. Atualmente, a aplicação do paradigma da manipulação direta de dados é frequentemente realizada através da digitalização de letras ou palavras.

Conectar: mostrar dados relacionados.

Este tipo de interação está diretamente relacionada com a tarefa "relacionar" definida por Ben Shneiderman e que serve para destacar associações e relações entre os dados representados. Quando diferentes representações do mesmo conjunto de dados, ou seja, diferentes perspectivas são facultadas ao utilizador, pode surgir alguma dificuldade na identificação de um dado específico em cada representação. Para facilitar esta identificação, pode ser utilizada uma técnica de *brushing* – figura 2.17 – a fim de destacar visualmente um dado elemento nas duas perspectivas.

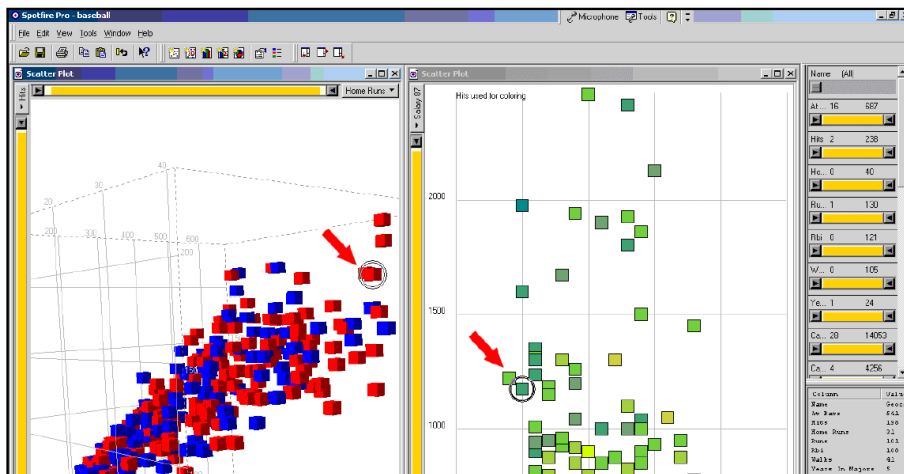


Figura 2.17 Aplicação da técnica de *brushing*. Representações diferentes do mesmo elemento e destaque visual do mesmo, facilitando a sua identificação a um possível utilizador.

Yi et al. conclui a sua pesquisa com a noção de que existe uma certa dificuldade em criar categorias claras, que as categorias enumeradas podem não ser as suficientes para todas as técnicas, pois algumas das técnicas não encaixam numa categoria ou encaixam em mais do que uma categoria. O autor suporta esta posição com o *zoom* semântico que se classifica como uma técnica de “abstração/ detalhe” devido à sua função de *zoom*, pertencendo também à categoria de codificação, na qual a representação muda, dada a alteração de escala.

Tanto Shneiderman como Yi et al. debruçaram-se nas tarefas interativas centradas no suporte das intenções do utilizador mas surge uma nova investigação, em 2007, realizada por Elmqvist et al.¹⁵ na área da interação. Esta investigação promove um novo conceito de fluidez na interação criando princípios de como atingir a fluidez nas visualizações de interação.

Numa tentativa de definir o conceito de fluidez, o autor afirma que esta, “na visualização de informação é um conceito fugaz e intangível caracterizado pela interação suave, perfeita, e poderosa; responsiva, interativa e rápida atualização dos gráficos; proporciona experiências de utilizador cuidadas, conscientes e compreensivas”¹⁵.

Elmqvist et al. define três princípios que caracterizam uma visualização de interação fluída e que são suportados, principalmente, por conceitos estudados na área de Interação Humano-Computador – IHC.

O primeiro princípio foca-se em promover o fluxo e é suportado pelos fatores de Mihaly Csikszentmihalyi. De acordo com Csikszentmihalyi, “a fluidez caracteriza-se por um estado mental de imersão total numa atividade onde os desafios da atividade e as habilidades do participante estão perfeitamente equilibrados levando a uma alta concentração, comprometimento e resultados gratificantes”¹⁵. Esta definição dá origem aos seguintes fatores: “desafio equilibrado”, “concentração”, “passagem a um estado de ação subconsciente”, “transformação do tempo”, “*feedback* imediato”, “sensação de controlo” e “gratificação imediata”. O “desafio equilibrado” tem em conta que a habilidade exigida pela atividade deve corresponder ao nível de habilidade do utilizador. O fator “concentração” sugere que a atividade deverá permitir uma grande concentração num campo limitado. A “passagem a um estado de ação subconsciente” permite ao utilizador conseguir fundir ação e consciência. A “transformação do tempo” tem como objetivo a perda do utilizador na atividade, ao ponto de perder a noção do tempo. O fator “*feedback* imediato” afirma que os utilizadores devem ser imediatamente informados dos seus progressos em relação aos seus objetivos. O fator “senso de controlo” dita que os utilizadores devem sentir que têm controlo sobre a atividade, de modo a afetarem o resultado. E, a “gratificação intrínseca”, afirma que a atividade deve ter uma recompensa para o utilizador.

O segundo princípio promove o suporte da manipulação direta, cujo conceito tem origem nos anos oitenta, sendo, Ben Sheinerman, um dos principais promotores, precisamente a propósito da visualização interativa de informação nos exercícios de manipulação direta suportados no design do *Starfield display*²⁶. O paradigma da manipulação direta¹⁷ promove uma interação através da manipulação de objetos e assenta em quatro princípios: representação contínua dos objetos de interesse; ações físicas em vez de sintaxes complexas;

¹⁵ Elmqvist et al., 2011

²⁶ Jog & Shneiderman, 1995

¹⁷ B. Shneiderman, 1983

operações rápidas, incrementais e reversíveis com impacto visível imediato no objeto; e utilização de um método de aprendizagem de camadas – *layered* – ou espiral que permita o seu uso com o mínimo de conhecimento.

O terceiro princípio centra-se na minimização dos “golfs da ação” – *gulfs of action*. Donald Norman¹⁵ define dois golfs que explicam o desafio de interagir com qualquer sistema: o golfo de avaliação consiste na diferença entre o estado do sistema e percepção do utilizador desse estado; e o golfo de execução consiste na diferença entre as ações permitidas pelo sistema e as intenções do utilizador.

Estes princípios guiam a construção de uma *interface* fluída para a visualização de informação. Estes transmitem ao utilizador sensações de controlo sobre a visualização – quase podendo tocar os dados e manipulando os dados visualizados – e de imersão e envolvimento com os objetos e sua representação visual.

Este tipo de interações “ênfatisam que as nossas habilidades cognitivas são desenhadas especificamente para raciocinar, agir e mover no mundo físico e social natural. (...) As habilidades físicas e sociais são muito mais constituintes do que consideramos como cognição humana, do que o processamento formal de símbolos ao dialogar com um computador através de botões rotulados, menus, hiperligações ou formas”¹⁵. Por outras palavras, a aplicação de interações baseadas em ações do mundo real aos sistemas interativos, tais como é o exemplo das visualizações de informação interativas, são mais vantajosas relativamente à criação de sentido ou entendimento por parte do utilizador e, portanto, têm mais probabilidades de atingir os objetivos da fluidez na interação.

2.3 Design da Informação

O processo de design de gráficos tem vindo a ser alvo de estudo ao longo de vários anos, resultando numa coleção de teorias e vários princípios que guiam a construção de um “bom gráfico”.

Os gráficos ou representações gráficas de informação são, sobretudo, sobre a comunicação de informação. Eles tencionam revelar padrões e relações não conhecidas ou dificilmente identificadas nos dados e a eficácia do seu design na transmissão de informação depende de processos cognitivos e da percepção visual. De facto, as “visualizações de dados podem ser consideradas artefactos cognitivos”⁴. Elas retiram vantagem das capacidades percetivas e cognitivas humanas para gerar novo conhecimento.

Como afirmam Alexandre & Tavares¹⁸ a percepção humana tem a sua base nos sentidos – visão, audição, paladar, olfato e tato. Sentidos esses que são constantemente “estimulados por um fluxo contínuo de acontecimentos envolventes”. A visualização de informação explora especialmente o sentido humano adquirido mais rapidamente pelo cérebro e que possui capacidades de processamento paralelo – a visão.

¹⁵ Elmqvist et al., 2011

⁴ Meirelles, 2013

¹⁸ Alexandre & Tavares, 2007

¹⁹ Ware, 2004

O poder de processamento paralelo do cérebro permite a focagem da atenção num ponto específico da visualização enquanto todo o resto da visualização, num raio bastante largo, é também alvo do sistema visual¹⁸. No entanto, apesar de termos acesso simultâneo a todas as informações representadas visualmente, o nosso sistema visual faz

uma extração faseada das características. Estas fases são explicadas pelo modelo de percepção proposto por Ware¹⁹.

1ª Fase: Processamento paralelo rápido para extração das características básicas.

Numa primeira fase, o processamento pré-cognitivo faz um levantamento rápido e simultâneo das características básicas, também conhecidas por características pré-atentivas. O termo pré-atentivas traduz a noção da velocidade e da facilidade com que certas propriedades são identificadas pelos humanos nas representações visuais¹⁸.

Características pré-atentivas permitem uma execução rápida e eficiente das tarefas, face à rápida identificação do alvo e da verificação das marcas mais importantes. Por exemplo, as propriedades básicas da cor, como a tonalidade, intensidade ou valor e a grossura da linha, ajudam-nos a executar a tarefa rapidamente porque são processadas pré-atentivamente. Estas tarefas são referentes à deteção do alvo, deteção de fronteira, rastreamento de regiões, contagem e estimativa e podem ser organizadas de acordo com quatro categorias – cor, forma, movimento e localização espacial.

Qualquer modificação das características pré-atentivas de um objeto, no espaço da representação visual, poderá alterar o foco de atenção do leitor. No entanto, quando existem muitas distrações na representação visual, a deteção das características pré-atentivas pode ser prejudicada.

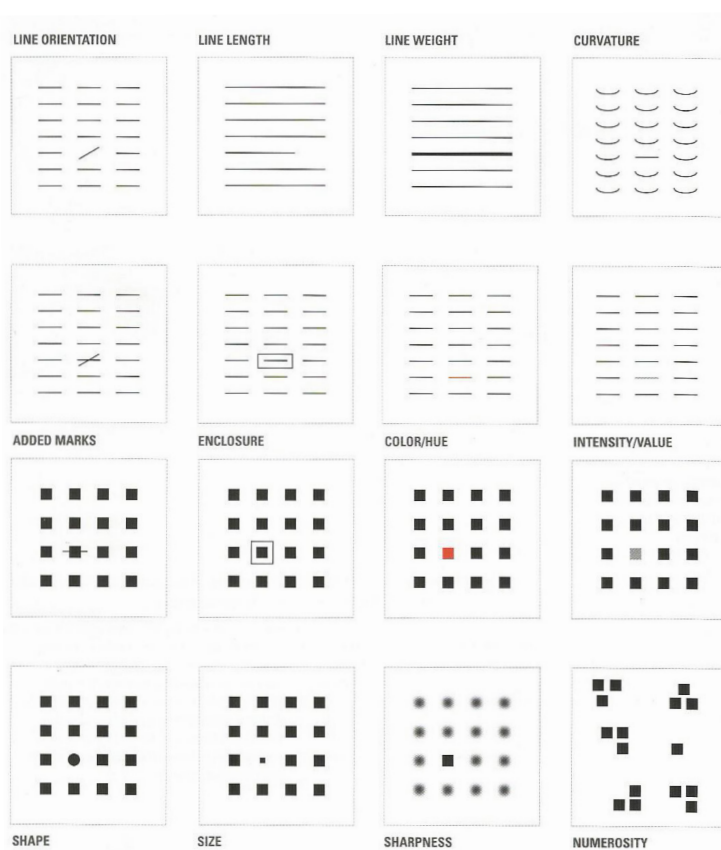


Figura 2.18° Características pré-atentivas, mostrando características para linhas (em cima) e para planos (em baixo).

A utilização consciente das características pré-atentivas na representação das marcas gráficas originam visualizações eficazes que continuam, no entanto, a requerer testes com o público alvo que confirmem a facilidade de execução das tarefas requeridas.

2ª Fase: Processamento lento em série para deteção de padrões e estruturas.

Na segunda fase, o processamento é sequencial e mais lento com o objetivo de detetar padrões nos dados visualizados, através da decomposição do campo visual em regiões. A deteção de padrões é extremamente flexível e influenciada pelas informações recolhidas, através do processamento paralelo e por processos de pesquisa *bottom-up*.

As leis de Gestalt ¹⁹ podem ser usadas como princípios de design que ajudam, de uma forma eficaz, a aumentar a deteção de padrões e deduções perceptivas. A sua teoria baseia-se no princípio de que “não se pode ter conhecimento do todo através das partes, e sim das partes através do todo; que os conjuntos possuem leis próprias e estas regem os seus elementos, e não o contrário; e que só através da perceção da totalidade é que o cérebro pode de facto perceber, descodificar e assimilar uma imagem ou um conceito”. As suas leis estabelecem a forma como os elementos de uma imagem podem vir a ser percebidos em termos organizacionais seguindo princípios de “proximidade”, “semelhança”, “fechamento”, “simplicidade”, “continuidade” e “separação entre figura e fundo”¹⁸.

A lei da “proximidade” determina que os elementos que se encontram próximos tendem a ser agrupados perceptivamente num conjunto, mesmo que não tenham grande similaridade entre si. Nas representações visuais é importante que coloquemos próximos uns dos outros os dados que são conceitualmente relacionados, de modo a facilitar a deteção e pesquisa de dados relacionados.

A lei de “semelhança” determina que os elementos com características semelhantes ou iguais tendem a ser agrupados em conjuntos. A similaridade observa-se principalmente na cor, forma e textura. A semelhança normalmente não se sobrepõe à proximidade.

A lei do “fecho” determina que, os elementos dispostos de maneira a formar um contorno fechado ou formas incompletas, tendem a ganhar maior grau de regularidade ou estabilidade, podendo vir a ganhar unidade. Esta lei é referente à tendência da perceção humana em perceber formas completas.

A lei da “simplicidade” determina que os elementos são percebidos mais facilmente quando apresentam simetria, regularidade e não possuem texturas.

A lei da “continuidade” diz que a perceção humana tende a orientar os elementos que parecem construir um padrão ou um fluxo na mesma direção pois, pela continuidade da direção e pelos ligamentos contínuos entre elementos, são mais fáceis de perceber do que modificações abruptas de direção.

Por último, a lei da “separação entre figura e fundo” afirma que qualquer campo perceptivo pode dividir-se numa figura sobre um fundo.

¹⁸ Alexandre & Tavares, 2007

¹⁹ Ware, 2004

A figura distingue-se do fundo por características como o tamanho, a forma, a cor e a posição. O objeto como figura, só é percebido em primeiro plano com o fundo devidamente separado da mesma.

3ª Fase: Processamento sequencial orientado ao objeto, com informação reduzida a alguns objetos mantidos na memória visual, para formar a base do pensamento visual.

A última fase do processamento das características de uma representação visual de informação, corresponde a um processamento de alto nível que combina as informações recolhidas das fases anteriores com conhecimentos previamente adquiridos. Neste último nível, o que está retido na memória por um determinado período de tempo, permitirá a construção de padrões que irão responder a pesquisas visuais. Por exemplo, se usamos um mapa de estrada para procurar uma determinada rota, a pesquisa visual desencadeará uma procura para ligar contornos vermelhos – que habitualmente representam vias importantes – entre dois símbolos visuais (representando as cidades pretendidas)¹⁹. Assim, chegamos à conclusão que também a experiência passada é fundamental no processo percetivo, de forma a possibilitar associações com conhecimento previamente adquirido. “A perceção visual é, portanto, o resultado da interação intrínseca entre informações externas adquiridas pelo sistema visual e informações internas baseadas no conhecimento previamente adquirido”¹⁸.

Também os símbolos e a forma como eles transmitem significado (ou como o seu significado é percecionado) é alvo deste estudo que contribui para a teoria de como criar bons gráficos e ao qual se dá o nome de estudo semiológico. A semiologia tem sido estudada principalmente por filósofos – aqueles que constroem argumentos baseados no exemplo, em vez de na experiência formal – entre os quais se distingue Jacques Bertin, cujo estudo se centrou na tentativa de classificar as marcas gráficas e na maneira de como elas expressam os dados através de variáveis visuais.

Uma vez que os seres humanos partilham do mesmo sistema visual é provável que todos percecionem o mesmo numa primeira aproximação. Assim, estas regras para o design de gráficos são eficazes para todos os dotados de uma capacidade visual sem deficiências, produzindo o mesmo resultado e transcendendo fronteiras culturais e raciais¹⁹.

A teoria de Bertin associa elementos básicos dos gráficos, como é o caso do plano, ponto, linha e área, com variáveis visuais e tipos de fenómenos, como por exemplo, cor e variação da mesma. Para Bertin⁸, o plano é a base de toda a representação gráfica. O plano é homogéneo e bidimensional. As três figuras elementares do plano geométrico e que nele podem ser implementadas são os pontos, as linhas e as áreas.

O ponto é um fenómeno sem comprimento num plano. O seu centro tem uma posição precisa, numa linha ou num plano, mas a marca permanece sem comprimento. O ponto representa, portanto, uma localização no plano sem comprimento ou área, independentemente do tamanho ou carácter da marca que a torna visível.

A linha é um fenómeno no plano que tem um comprimento mas não tem área. A linha é independente da largura e características da marca que a tornam visível. Apesar de poder variar na posição, essa variação nunca poderá significar área no plano. A linha pode representar a fronteira de um continente ou de uma propriedade, o curso de um navio ou a rota de

¹⁸ Alexandre & Tavares, 2007

¹⁹ Ware, 2004

⁸ Bertin, 2011

autocarro – fenómenos lineares sem uma área teórica.

A área é a única marca com um tamanho mensurável no plano, ou seja tem duas dimensões. Ela pode variar em posição, mas não em tamanho, forma ou orientação sem que o seu significado seja alterado.

Num dado espaço da representação visual a falta de sinais significa ausência de fenómenos, ou seja, quando não existem pontos, linhas ou áreas num dado espaço do plano, não existe variância de variáveis visuais nas marcas, não existe informação.

Qualquer variável visual tem um significado – seja cor, textura ou forma – se elas existem num plano é porque significam alguma coisa, é porque transmitem uma informação específica.

Existem seis categorias de variáveis que podem ser relacionadas com um componente. A categoria “tamanho” pode ser a altura de uma coluna, a área da marca ou o número de sinais iguais. A categoria “valor” utiliza uma escala de vários valores entre preto e branco. A categoria “textura” é dada pela espessura dos componentes – as linhas, por exemplo – numa determinada área. A categoria “cor” refere-se ao *hue* – a tonalidade da cor. A categoria “orientação” refere-se às várias orientações que uma linha ou um padrão de linhas pode adquirir. E a categoria “forma” é dada por uma marca com um tamanho constante, podendo ter um número infinito de diferentes formas⁸.

As variáveis tamanho, valor, textura, cor, orientação e forma, bem como as dimensões do plano, fazem parte dos elementos com características pré-atentivas⁴. Qualquer variável pode ser usada para representar qualquer componente, mas cada variável não é adequada para todos os componentes. A representação correta de uma componente quantitativa pode ser apenas conseguida com a variação do tamanho, ou seja, para conseguir que a componente seja percebida como quantitativa é necessário recorrer à variável “tamanho”. Existem portanto, vários tipos de perceções associados às variáveis. Segundo Bertin, elas são: a perceção associativa, a perceção seletiva, a perceção ordenada e a perceção quantitativa.

A perceção associativa é útil quando se procura equalizar uma variação e para correspondências agrupadas com todas as categorias desta variação combinadas. Bertin quer com isto dizer que existe uma perceção associativa em conjuntos formados por grupos de elementos com formas diferentes, por exemplo. A variação da forma é, assim, associativa. Já a variação de tamanho ou a variação de valor não são associativas.

A perceção seletiva é utilizada para obter resposta à pergunta “Onde está uma dada categoria?”. O olho pode ser capaz de isolar todos os elementos desta categoria e perceber a imagem formada pelas marcas dessa categoria. A variável é seletiva no caso de esta perceção ser imediata e cada categoria formar uma família. O autor não considera que a forma ou a orientação sejam seletivas quando representadas pela área.

A perceção ordenada deve ser usada para comparar duas ou mais componentes ordenadas. Tal comparação pode ser imediata, nesse caso a variável é ordenada. Textura, valor e tamanho impõe uma ordem que é universal e imediatamente perceptível. Textura, valor

⁸ Bertin, 2011

⁴ Meirelles, 2014

³ Tuffte, 1995

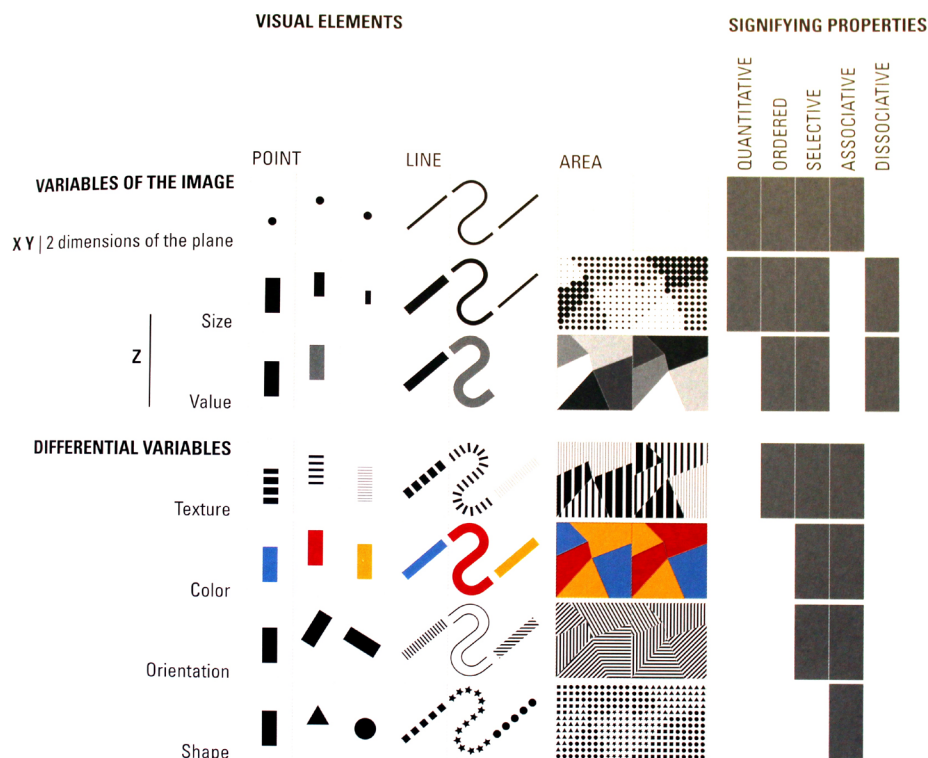


Figura 2.19^c Tabela resumo das percepções das marcas construída por Isabelle Meireles e baseada no estudo semiótico de Bertin.

e tamanho são ordenados para os três tipos de implementação (linha, área e ponto).

Finalmente, a percepção quantitativa está envolvida quando procuramos definir numericamente a relação entre dois sinais. Apenas a variação de tamanho é quantitativa.

“Nós temos muitas técnicas de visualização padronizadas que funcionam bem e estão solidamente incorporadas em práticas de trabalho (...). Em muitas aplicações, um bom design é um design padronizado”¹⁹. De acordo com Edward Tufte³ – um dos autores que mais contribuiu para a padronização de princípios aplicados aos gráficos – a receita para o bom design de gráficos tem dois ingredientes-chave: a simplicidade do design e a complexidade dos dados. Gráficos visualmente atrativos, também ganham o seu poder do conteúdo e das interpretações para além da imediata, dada pela exibição de alguns números. “Os melhores gráficos são sobre o útil e o importante, acerca da vida e da morte, sobre o universo”³.

Em 1995, Tufte publica um livro intitulado “*The Visual Display of Quantitative Information*” que partilha vários princípios que guiaram e guiam os designers de visualizações de informação. Apesar destes princípios terem sido criados em função da visualização de dados em gráficos estáticos, muitos deles continuam a ser válidos também para as visualizações dinâmicas. Assim, o levantamento destes princípios torna-se útil para o processo de design da informação.

Princípios da excelência dos gráficos, segundo Tufte:

- Tem um bom design da apresentação de dados interessantes – um problema de substância, estatísticas e de design.
- Consiste em ideias complexas comunicadas com clareza, precisão e eficácia.
- Dá ao observador o maior número de ideias no menor tempo possível,

³ Tufte, 1995

¹⁹ Ware, 2004

com a menor quantidade de tinta e no menor espaço possível.

- Usa quase sempre multivariáveis.
- Requer registrar a verdade sobre os dados: “A excelência dos gráficos começa com dizer a verdade”.

Tufte afirma que cada parte do gráfico gera, no leitor, expectativa visual sobre outras partes e frequentemente determina o que os olhos vêem. Quando há variação do design num gráfico, existe ambiguidade e alta probabilidade de engano. “O engano resulta da incorreta extrapolação das expectativas visuais geradas em lugares do gráfico”³. O observador confunde as mudanças no design com as mudanças nos dados, sendo que é fundamental mostrar variação nos dados – comunicadora de informação – mas não no design. Os princípios da integridade gráfica ajudam, assim, a evitar situações em que a representação gráfica de dados possa por em causa a verdade, enganando o leitor.

Princípios da integridade dos gráficos:

- As representações de números, como medida física na superfície do gráfico, devem ser diretamente proporcionais às quantidades numéricas a representar.
- Devem ser usadas legendas para combater a distorção e ambiguidade. É importante legendar, sobretudo, eventos importantes nos dados.
- Mostrar variação na informação, não no design.
- A exibição de dados monetários em gráficos *time-series* deve ser feita através de unidades padrão monetárias, que são quase sempre melhores que unidades nominais.
- O número de dimensões da informação representada não deve exceder o número de dimensões dos dados, ou seja, a representação de valores numéricos (com uma dimensão) através de áreas (que são compostas por duas dimensões) contribui para a confusão do leitor e aumenta a probabilidade de transmitir uma informação errada – figura 2.20 .
- Os gráficos não devem citar dados fora do contexto.

“Muito do mundo, hoje em dia, é observado e acedido quantitativamente – e gráficos bem desenhados são muito mais eficientes do que palavras em mostrar tais observações”³.

A maximização da porção de tinta dedicada aos dados – *data-ink* – é também um dos conselhos dados por Tufte, cujo princípio é “maximizar a quantidade de tinta com significado”. A tinta utilizada deve, portanto, apresentar nova informação. É um princípio que, normalmente, se aplica a gráficos já concluídos, para que possam ser refinados. A aplicação deste princípio pode ser feita com a ajuda de dois princípios de eliminação: (1) apagar tinta não dedicada aos dados, com razão e (2), a não ser que a redundância tenha uma razão para existir, apagar dados redundantes com razão.

Os dados redundantes repetem a informação desnecessariamente. Um exemplo disso é a utilização de simetria bilateral, “*box plot*”, “*open bar*” e “*Chernoff faces*” como pode ser visto na figura 2.21. No entanto existem casos em que a redundância se torna útil. A redundância pode dar um contexto e ordem à complexidade, facilitando comparações entre várias partes dos dados e eventualmente, criar um balanço estético. Por exemplo, no gráfico de Marey – figura 2.22 – que representa o horário dos comboios, juntar um meio ciclo extra faz com que todos os comboios nas primeiras 24h do horário tenham uma linha contínua – figura 2.23.

³ Tufte, 1995



Figura 2.20^b Neste exemplo, a representação de um único valor numérico é feita através da largura e altura da nota, facto que resulta no engano do leitor face às porporções erradas.

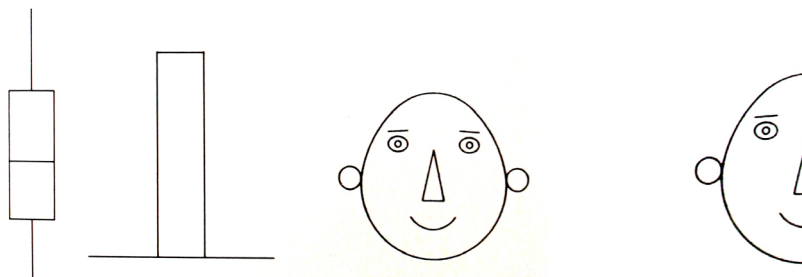


Figura 2.21^b À esquerda temos “box plot”, “open bar” e “Chernoff face”, à direita temos metade de uma “Chernoff face” mostrando que a mensagem é recebida é semelhante com apenas metade da tinta.

Tufte aponta cinco princípios que produzem mudanças substanciais num gráfico, também usados para redesenho de gráficos:

- Devem mostrar dados
- Maximizar a mancha de tinta dedicada aos dados.
- Eliminar mancha de tinta não relativa aos dados.
- Eliminar dados redundantes.
- Rever e editar.

De acordo com o autor, a grelha é também um elemento que deve ser tido em atenção na hora da revisão. A grelha deve ser silenciada ou suprimida, para que tenha uma presença meramente implícita, dando ênfase aos dados. As suas linhas são consideradas *chartjunk* – mancha de tinta sem dados – e, por isso, devem ser eliminadas quando possível.

Considerando o gráfico que foi dado como exemplo anteriormente – figura 2.21 – podemos observar que a grelha é um elemento com uma forte presença, para além de importante, na representação. Neste gráfico a grelha não pode ser eliminada, mas pode ser silenciada através do seu redesenho com linhas mais finas e num tom cinzento. Esta solução, traz os dados mais importantes para o primeiro plano – figura 2.24.

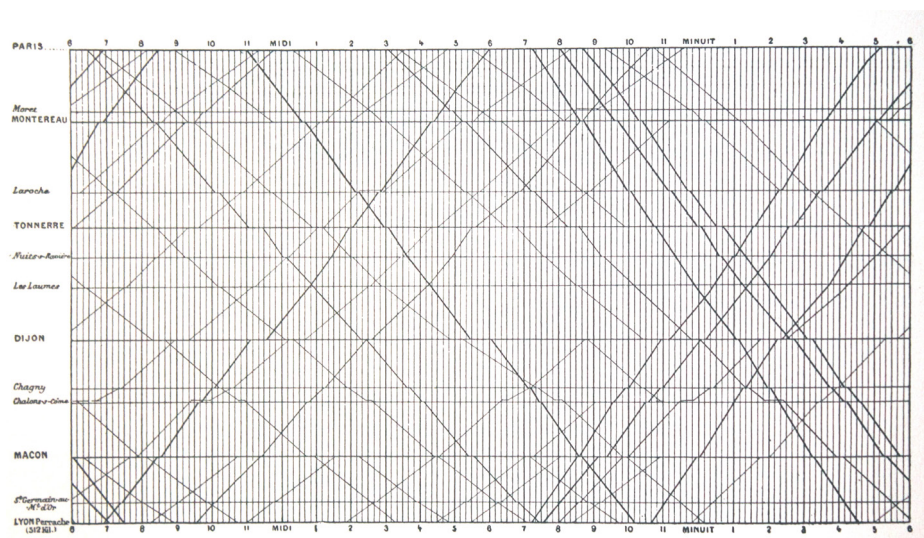


Figura 2.22b "La Methode Graphic" (Paris, 1885) desenhado por E. J. Marey.

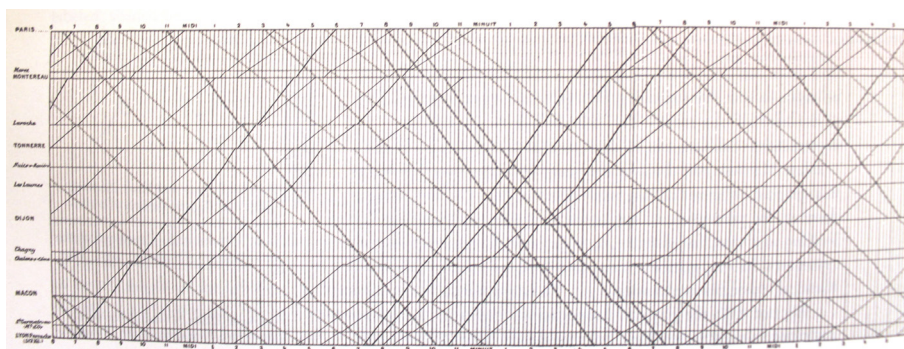


Figura 2.23b Redesenho do gráfico "La Methode Graphic" (Paris, 1885) de forma a que todos os horários das primeiras 24h horas tenham uma linha contínua, evitando quebras na leitura do horário.

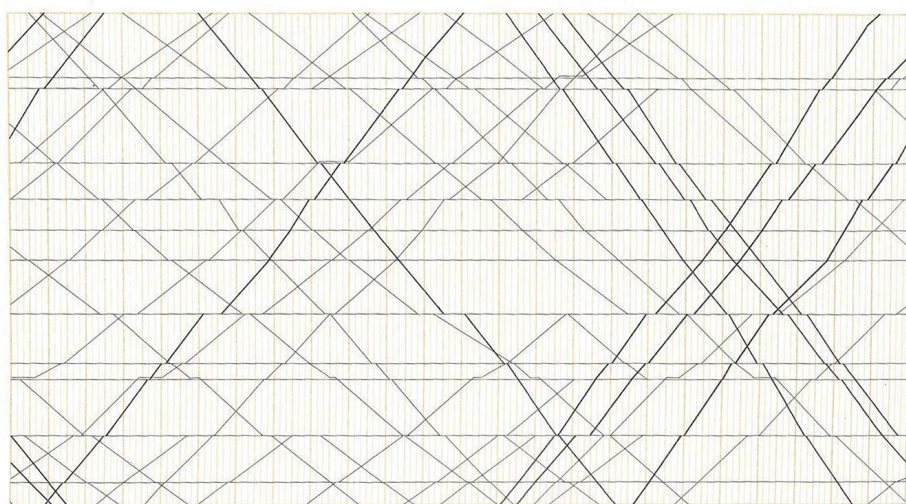


Figura 2.24^b Redesenho do gráfico “*La Méthode Graphic*” de acordo com os princípios de edição de Tufte.

A edição e revisão dos gráficos quase sempre têm um efeito positivo nos mesmos, levando a um melhor resultado. Ao maximizar a informação e ao eliminar *chartjunk* conseguimos gerar alternativas gráficas, enquanto somos impelidos na direção que a revisão deve tomar.

Os gráficos são a melhor maneira de dar sentido a muitos dados complexos que não podem ser geridos de outra maneira. A existência de uma grande quantidade de dados permite a criação de designs ricos em informação, que dão contexto e credibilidade à evidência estatística, ao contrário dos designs pobres em dados que causam suspeitas.

No entanto, este tipo de gráficos com muita informação, deve ser desenhado com um certo cuidado. Quanto maior for o volume de dados, menor terão de ser as representações gráficas dos dados, sejam linhas, pontos ou áreas.

“Muitos gráficos de dados podem ser reduzidos em metade da área publicada sem perda de legibilidade ou informação”³. Este tipo de gráficos – densos em informação e com dimensões diminutas – são nomeados de pequenos múltiplos. Considerados por Tufte como efetivos e poderosos, quando bem desenhados, os pequenos múltiplos são inevitavelmente comparativos, contentores de multivariáveis, baseados numa matriz de dados, desenhados com mancha de tinta praticamente toda dedicada à representação de dados, interpretados de modo eficiente e são frequentemente narrativos.

Resumindo todos os princípios de Tufte, as visualizações de dados atrativas devem ter um formato e design apropriados; usar simultaneamente palavras, números e desenhos; refletir balanço, proporção e uma escala relevante; exibir uma complexidade de detalhes de forma acessível; frequentemente ter uma qualidade narrativa, uma história para contar sobre os dados; ser desenhados de uma maneira profissional, com cuidado na produção dos detalhes técnicos; e evitar conteúdos meramente decorativos, incluindo *chartjunk*.

³ Tufte, 1995

2.4 Trabalhos Relacionados

Os projetos que se seguem foram selecionados de acordo com os requisitos da investigação proposta, sendo eles: a representação de dados em larga escala, a representação de grandes períodos de tempo e a interação. Após a descrição de cada projeto, será feita uma análise semiótica dos mesmos, centrada nas percepções obtidas, através das variáveis utilizadas.

“*Last clock*” é um projeto realizado por Jussi Ängeslevä e Ross Cooper no ano de 2002, que representa o tempo em função de imagens capturadas com uma câmara e exibidas de acordo com o ritmo que marca o tempo. Com uma aparência semelhante à dos relógios analógicos, a *interface* apresenta três círculos concêntricos, cada um representando uma medida temporal – segundos, minutos e horas.

Para captura das imagens que alimentam o relógio em cada nível, a *interface* está conectada com uma fonte de vídeo – por exemplo com a *web* câmara de um tablet pessoal ou uma câmara de vigilância de um local público – que atualiza a imagem em cada segundo, minuto e hora, de forma a coincidir com o ritmo do relógio analógico.

Esta visualização permite a recolha e exibição de um total de 12 imagens para representar as horas, 720 para os minutos e 43 200 para representar os segundos num período de tempo que corresponde a uma volta completa do ponteiro das horas, num relógio analógico. Sendo que o relógio não pára, as imagens estão constantemente a ser substituídas por novas imagens.

O princípio interativo ‘Visão geral primeiro, *zoom* e filtro, e então detalhes sobre a procura’⁷⁰ não se aplica a este projeto. O utilizador não pode explorar os dados através de métodos interativos, no entanto, esta é uma aplicação que atrai o utilizador face à possibilidade de personalização da visualização.

A nível interativo podemos apenas identificar o paradigma da manipulação direta de dados, pois os dados a visualizar poderão ser definidos pelo utilizador que escolhe a direção da câmara e, portanto, as paisagens que serão recolhidas.

Relaciona-se com o projeto proposto nesta dissertação pelo facto de representar um grande número de dados – as imagens – em função do tempo, pela interatividade e possível manipulação das imagens.

Variáveis Visuais	Implementação	Percepção	Variável
Tamanho Dim X, Y	Linha/Área Linha	Seletiva Quantitativa	Categoria Tempo Tempo actual

Tabela 2.1 Análise Semiótica do projeto “*Last Clock*”.





Figura 2.25^k 22:46:50 After football game,
BBC2

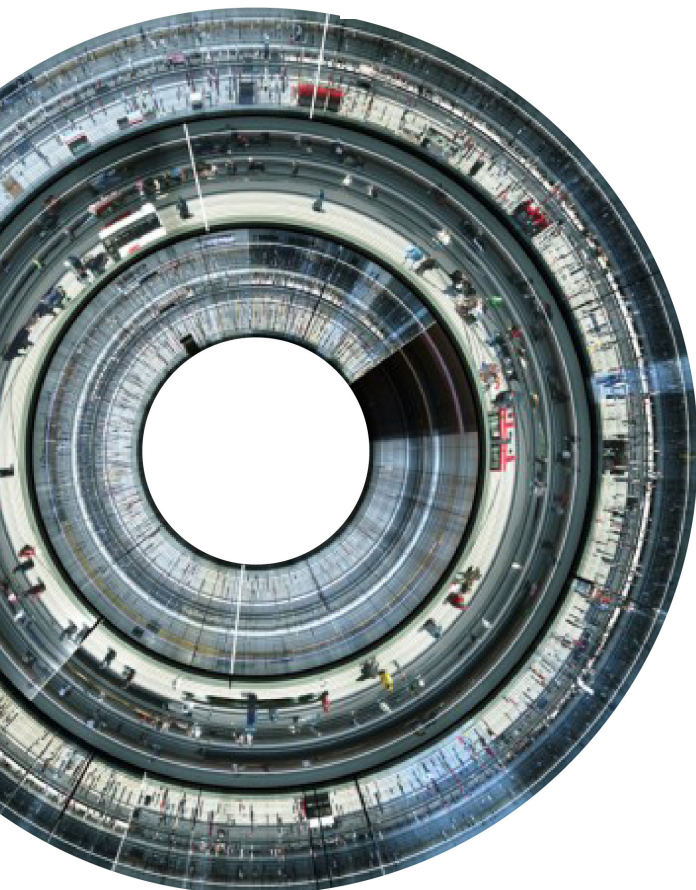


Figura 2.27^k 13:56:02 Kulturhuset
Stockholm, Sweden

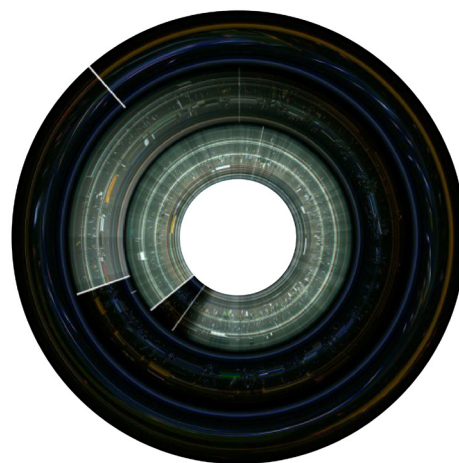


Figura 2.26^k 19:43:52 Splendor Road
Taipei, Taiwan

“NameVoyager” é uma visualização interativa para *web*, realizada por Martin Wattenberg, em 2005 e que exhibe as tendências na escolha de nomes para bebês, desde 1880 a 2013 – um período de tempo superior a um século, mais exactamente 133 anos. O tempo é representado horizontalmente – da esquerda para a direita – e o eixo vertical corresponde à frequência de ocorrência, por milhões de bebês.

Cada banda representa um único nome, à qual é atribuída uma cor consoante o sexo do bebé – rosa para meninas e azul para meninos – sendo que o brilho (e não a tonalidade utilizada no sistema de Bertin) é um atributo adicional que permite fazer distinção de nomes vizinhos, isto é, quando dois nomes femininos se encontram um por cima do outro na representação visual.

A visualização permite ainda uma pesquisa através de consultas, nas quais o utilizador apenas precisa de digitar um nome ou uma letra, para que os dados apresentados sejam alterados de forma a responderem à pesquisa – um exemplo da manipulação direta falada por Shneiderman e citada em §2.2 do documento em questão. Também o seu princípio ‘Visão geral primeiro, *zoom* e filtro, e então detalhes sobre a procura’ é aplicado nesta visualização.

Esta aplicação relaciona-se com o projeto proposto, pois representa milhares de dados num longo período temporal que se aproxima ao período a ser representado neste projeto – aproximadamente um século. Para além disso, é uma aplicação interativa situada na *web*.

Variáveis Visuais	Implementação	Percepção	Variável
Tamanho	Área	Quantitativa	Quant. de Nomes
Cor	Área	Seletiva	Sexo
Dim X	Linha	Ordenada	Tempo
Dim Y	Linha	Quantitativa	Quant. de Nomes

Tabela 2.2 Análise Semiótica do projeto “NameVoyager”.

<http://www.babynamewizard.com/voyager> – disponível em 28 de janeiro de 2015

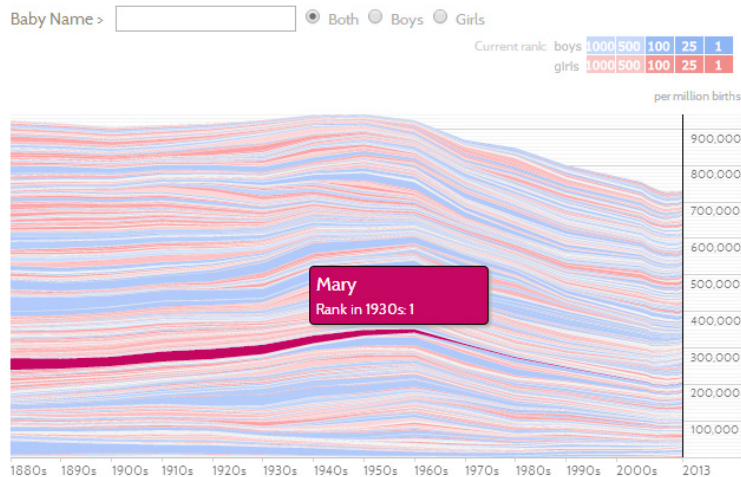


Figura 2.28' Visão geral da visualização interativa “Name Voyager”.

“*The Ebb and Flow of Movies: Box Office Receipts 1986 – 2008*” é uma visualização interativa para *web*, publicada em 2008, no *New York Times* e realizada por Mathew Bloch, Lee Byron, Shan Carter e Amanda Cox. Esta visualização exibe as receitas da bilheteira para sete mil e quinhentos filmes desde janeiro de 1986 a fevereiro de 2008 – aproximadamente vinte e um anos.

A visualização utiliza o método *streamgraph*, criado por Lee Byron e que compõe as *layers* numa forma empilhada e que tem a sua inspiração no método “*ThemeRiver*”, inventado por Havre e colegas em 2000²⁷. É um método que utiliza técnicas de interpolação suave de dados discretos, gerando um design simétrico de *layers* que se centram à volta do eixo horizontal.

Os *streamgraphs* também adotaram técnicas utilizadas no projeto “*Name Voyager*”. O tempo é representado horizontalmente, da esquerda para a direita, o tamanho de cada banda representa a receita da bilheteira por semana e o comprimento representa a longevidade. A área da forma, tal como a cor – que varia na saturação de amarelo e laranja – correspondem ao produto interno bruto do filme.

A manipulação direta de dados também está presente nesta visualização, sendo que o utilizador pode pesquisar dados referentes a um filme em particular. É evidente a aplicação do princípio interativo ‘Visão geral primeiro, *zoom* e filtro, e então detalhes sobre a procura’.

Tal como o projeto anterior, “*The Ebb and Flow of Movies: Box Office Receipts 1986 – 2008*” é uma visualização interativa para *web*, fator em comum com o nosso protótipo. E apesar de ter um período um pouco menor, a visualização também representa os dados em função do tempo.

Variáveis Visuais	Implementação	Percepção	Variável
Tamanho	Área	Quantitativa	PIB do filme
Cor	Área	Seletiva	PIB do filme
Altura	Área	Quantitativa	Receita semanal
Comprimento	Área	Quantitativa	Longevidade

Tabela 2.3 Análise Semiótica do projeto “*The Ebb and Flow of Movies: Box Office Receipts 1986 – 2008*”.

http://www.nytimes.com/interactive/2008/02/23/movies/20080223_REVENUE_GRAPHIC.html – disponível em 28 de janeiro de 2015

²⁷ Havre, Hetzler, & Nowell, 2000

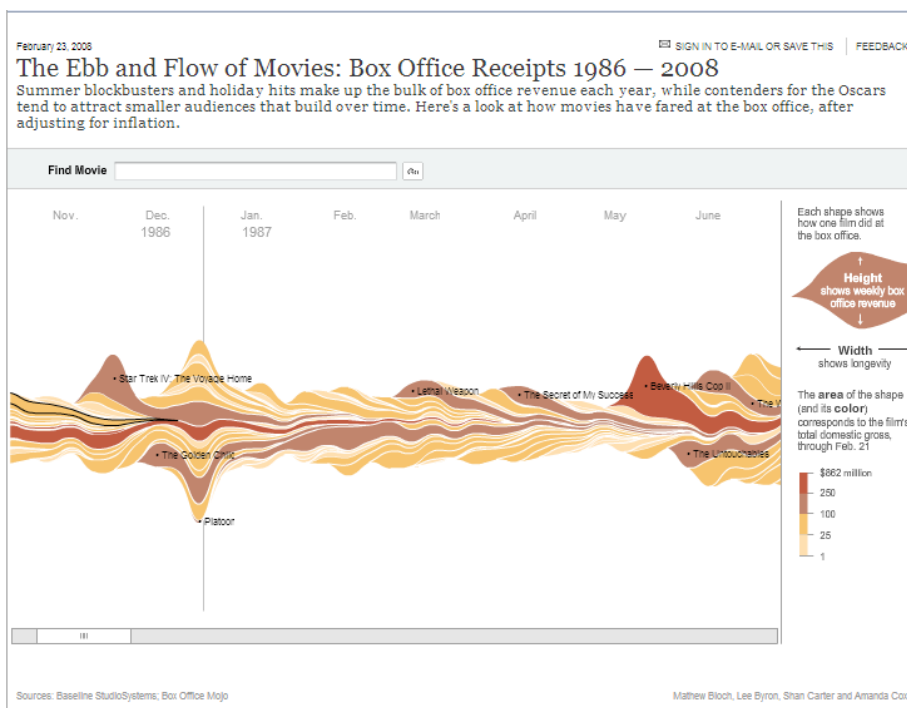
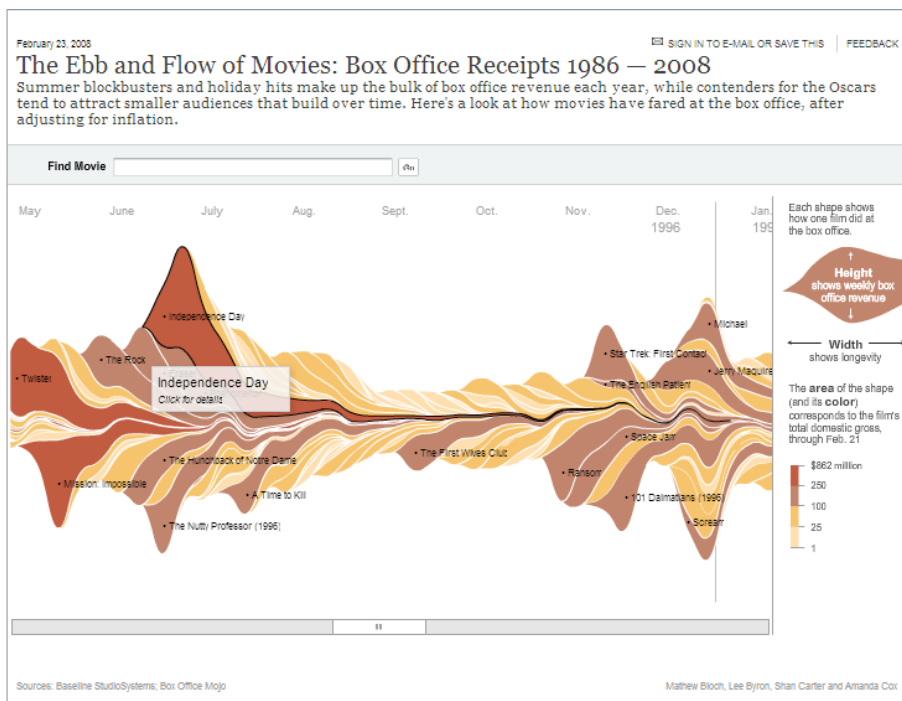


Figura 2.29^m Snapshots da aplicação “The Ebb and Flow of Movies: Box Office Receipts 1986 – 2008”. Em cima representa dados desde Maio a Dezembro de 96 e, em baixo, representa dados referentes ao período de Novembro de 86 a Junho de 97.

“*10 Years of Wikipedia*” foi um projeto realizado pelo alemão Gregor Aisch, em 2011, a propósito do aniversário da enciclopédia. Ao contrário das visualizações anteriores que fazem uso de uma linha temporal horizontal, o autor opta por utilizar um gráfico de linhas polar para representar dados relativos à Wikipedia.

“As estruturas espirais ou circulares são frequentemente usadas quando o objetivo é mostrar uma escala de tempo contínua, como também revelar dados periódicos. Espirais são melhores para representar dados contínuos”⁴.

Neste projeto, a manipulação dos dados que são representados é feita através de botões, que os utilizadores podem pressionar quando pretendem ver novos dados representados.

Este projeto tem em comum com o nosso, o facto de ser uma aplicação interativa para *web* e de representar dados em função do tempo, sendo que apresenta uma nova forma de representar o tempo.

Variáveis Visuais	Implementação	Perceção	Variável
Dim x, y	Linha	Quantitativa	Nr. de entradas
Cor	Linha	Ordenada	Tempo

Tabela 2.4 Análise Semiótica do projeto “*10 Years of wikipedia*”.

<http://visualdata.dw.de/en/wikipedia/> – disponível em 28 de janeiro de 2015

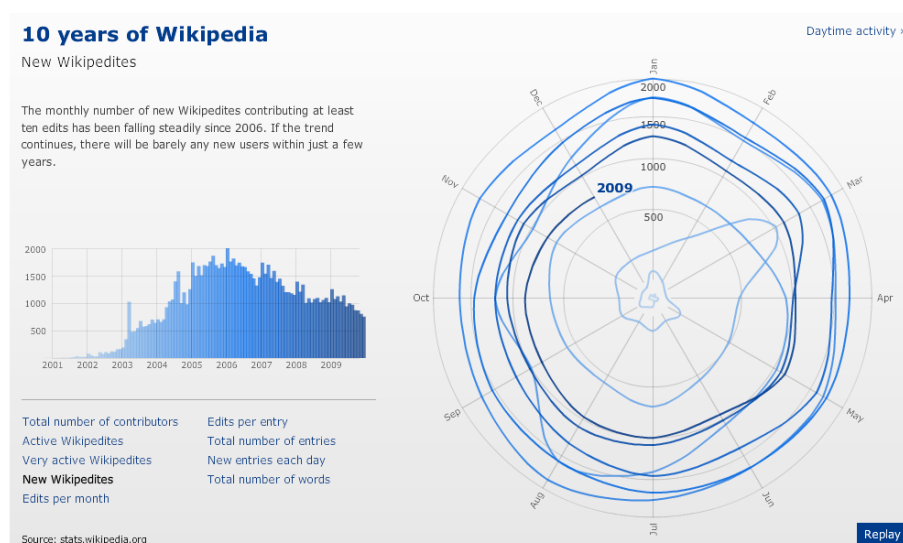


Figura 2.30 Snapshot da aplicação “*10 Years of wikipedia*”.

⁴ Meirelles, 2013

“The Racial Dot Map” é uma visualização desenvolvida por Dustin Cable, em 2013, e que representa as raças dos seres humanos que habitam os Estados Unidos, a partir de pontos coloridos colocados num mapa. Sendo que um ponto representa uma única pessoa – podendo ser “Branca”, “Negra”, “Asiática” ou “Outra Raça/ Nativo Americano/ Multi-Racial” – a quantidade de pontos revela também a densidade populacional de cada área geográfica. Uma vez que os pontos são mais pequenos do que um pixel, a junção de várias pessoas num mesmo pixel poderá originar cores não identificadas nas legendas. Por exemplo, se um pixel representa uma dada quantidade de pessoas “Brancas” (azuis) e de pessoas “Asiáticas” (vermelhas), então a coloração do pixel será roxa, sendo que a sua tonalidade dependerá da quantidade de cada raça.

A fonte dos dados apresentados no mapa provém do resumo dos Censos de 2010, realizados nos EUA, que estão acessíveis ao público através do Sistema Nacional Geográfico de Informação Histórica. A informação é baseada no *“census block”*, a menor área geográfica para a qual os dados são coletados (aproximadamente equivalente a um quarteirão da cidade em uma área urbana).

Apesar da quantidade imensa de dados representados, este projeto não varia a informação consoante um período de tempo. Permite, no entanto, uma exploração dos dados a partir de técnicas de *zoom*, de alteração do código de cor para preto e branco e da opção de visualizar legendas ou não.

Nesta visualização a densidade dos dados representados e a interação são os fatores que se aproximam mais ao nosso projeto. A representação temporal não é um desses fatores, pois apenas são utilizados dados de um único ano.

Variáveis Visuais	Implementação	Percepção	Variável
Dim x, y	Ponto	Seletiva	Pos. Geográfica
Cor	Ponto	Seletiva	Raça

Tabela 2.5 Análise Semiótica do projeto *“The Racial Dot Map”*.

<http://www.coopercenter.org/demographics/Racial-Dot-Map> – disponível em 28 de janeiro de 2015

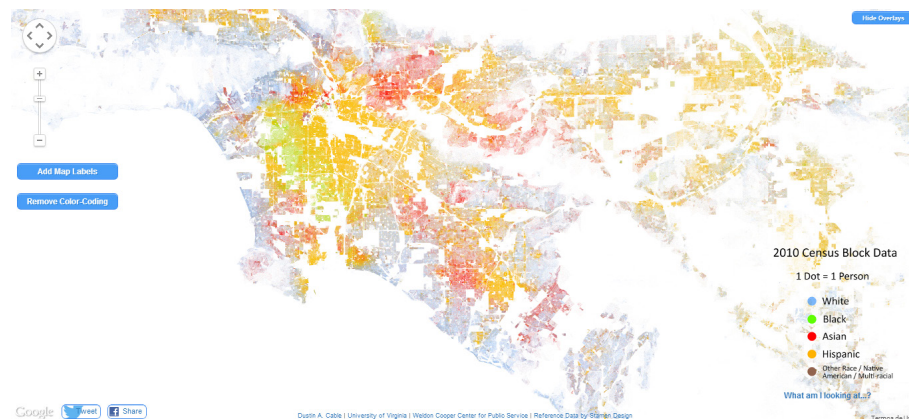


Figura 2.31 Snapshot da visualização interativa *“The Racial Dot Map”* resultando do uso de técnicas de aproximação.

“Lúsica” é um projeto realizado por uma equipa de alunos e docentes da Faculdade de Ciências da Universidade de Lisboa (FCUL), no ano de 2014, no âmbito de uma disciplina lecionada no Mestrado de Engenharia Informática e em parceria com o SAPO Labs. Esta visualização exhibe tendências musicais ao longo do tempo, a partir da recolha de citações de músicas de artistas lusófonos, nas redes sociais, entre Abril de 2011 e Março de 2014.

A forma gráfica dos dados utiliza o método *stackedgraph* que, tal como o método *streamgraph*, empilha a informação em *layers*, com a diferença de que o primeiro método não exhibe uma forma simétrica. Os dados não são representados em torno do eixo horizontal, mas, sim, sobre o eixo horizontal.

A interação permite ao utilizador selecionar um estilo representado no gráfico através das bandas coloridas e observar os álbuns mais citados no mês correspondente à posição do rato no gráfico. O princípio ‘Visão geral primeiro, *zoom* e filtro, e então detalhes sobre a procura’ não se encontra tão explícito neste projeto.

A relação deste projeto com o nosso é o facto de ser um projeto realizado em parceria com o SAPO Labs e, também por se tratar de uma visualização interativa que representa dados em função do tempo.

Variáveis Visuais	Implementação	Perceção	Variável
Cor	Área	Seletiva	Estilo musical
Dim Y	Área	Quantitativa	Total publicações
Dim X	Linha	Quantitativa	Tempo

Tabela 2.6 Análise Semiótica do projeto “Lúsica” .

<http://www.lasige.di.fc.ul.pt/webtools/fpsapo/music/lusica/index.html> – disponível em 28 de janeiro de 2015

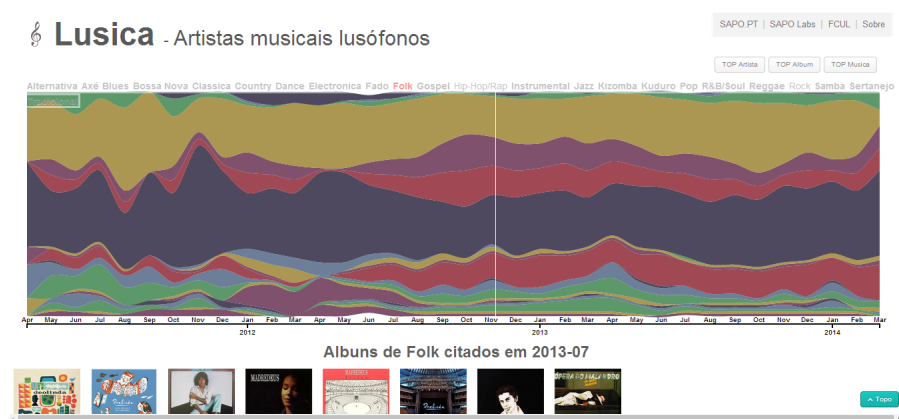


Figura 2.32^a Snapshot da visualização interativa “Lúsica”.

“**Grande Área**” é um projeto interativo de visualização de informação, realizado pelo SAPO Labs, em 2014 e que exhibe dados estatísticos focados nos mundiais de futebol entre 1930 a 2014. O projeto está dividido em duas visualizações de dados intitulados de “Estatísticas” – figura 2.33 – e “Comparômetro” – figura 3.34. A primeira permite a visualização das várias “seleções” ou de “jogadores” de acordo com “eficiência”, “experiência” e “disciplina”. Sendo que a “eficiência” se traduz na organização dos dados por números de golos marcados, a “experiência” corresponde ao número de jogos e a “disciplina” ao número de cartões. Ou seja, quando uma dada seleção ou jogador têm o maior número de golos, então a representação dessa seleção ou jogador terá um tamanho maior quando os dados são organizados por “eficiência”.

Através da seleção de uma equipa, é possível visualizar contra que outras equipas foram marcados mais golos, recebidos mais cartões ou realizados mais jogos. A mesma pesquisa pode ser feita em função dos jogadores. Por exemplo, selecionando um jogador podemos saber contra que seleções esse jogador marcou mais golos – figura 2.35.

Acedendo à página de perfil do jogador ou equipa selecionada existe a possibilidade de analisar, com mais detalhe, estatísticas dos jogos, transferências de jogadores, carreira profissional do jogador e vídeos com os melhores momentos dos jogos da Liga Portuguesa desde 2008.

O “Comparômetro” permite a comparação do desempenho de duas equipas selecionadas pelo utilizador, através de um gráfico rede.

Neste projeto, o princípio de Shneiderman é mais óbvio na visualização “Estatística”, enquanto que, no “Comparômetro” não há possibilidade de explorar dados mais detalhadamente.

Esta visualização interativa está relacionada com o nosso projeto pela ligação com o SAPO Labs e por ser uma tentativa de resolver o problema a que nos propomos: representar uma larga escala de dados em função do tempo, cuja duração é aproximadamente um século. O objetivo da interatividade, como meio de ver novos detalhes e exploração de dados, é também comum a ambos os projetos.

Variáveis Visuais	Implementação	Perceção	Variável
Tamanho	Área	Quantitativa	Golos Marcados

Tabela 2.7 Análise Semiótica da visualização “Estatística”, do projeto “Grande área”.

Variáveis Visuais	Implementação	Perceção	Variável
Cor	Área	Seletiva	Seleção
Dim x,y	Ponto	Quantitativa	golos, cartões, etc.

Tabela 2.8 Análise Semiótica da visualização “comparômetro”, do projeto “Grande área”.



Figura 2.33' Snapshot da visualização interativa "Grande Área", na visualização "Estatísticas".

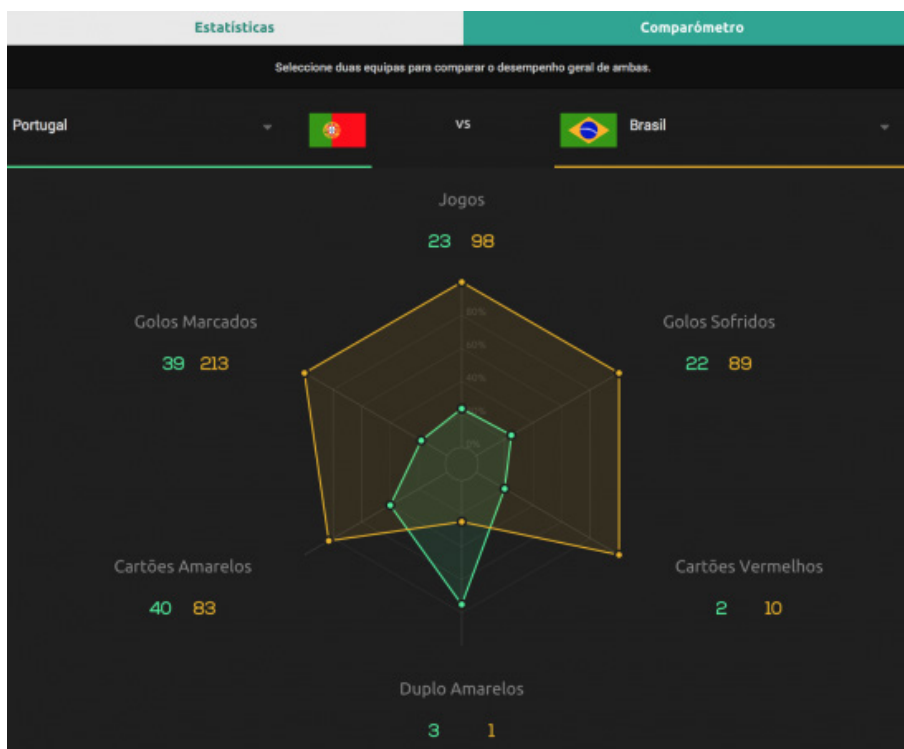


Figura 2.34' Snapshot da visualização interativa "Grande Área", na visualização "Comparómetro".



Figura 2.35^s Snapshot da visualização interativa “Grande Área”, com detalhes sobre um dado selecionado – o jogador.

“**Máquina do tempo**” é um projeto interativo de visualização de informação, fruto de uma parceria entre o SAPO Labs e a Faculdade de Engenharia da Universidade do Porto (FEUP), em 2014. Tem a base dos seus dados nas notícias datadas, desde 1989 até 2014, a partir das quais são representadas relações entre personalidades nelas existentes.

Os utilizadores podem efetuar pesquisas, relativas a acontecimentos, através de uma data ou de uma personalidade. Quando a visualização se centra numa pessoa, o utilizador pode aceder – através do “Perfil” – a uma lista de citações e pode ver artigos publicados em que a pessoa seja citada. Para além disso, pode ser visualizado um gráfico de rede – figura 2.36 – acrescentando outras personalidades relacionadas com os mesmos artigos ou assuntos.

O gráfico utiliza cores para representar a categoria de cada pessoa – economia, educação, local, política, saúde, ciências e tecnologias, cultura, desporto e sociedade – e permite alterar o número de relações que o utilizador deseja observar, para uma análise mais ou menos detalhada.

O projeto lida, no total, com cerca de 7 milhões de notícias que se estendem ao longo de 25 anos. Também, neste projeto, se pode verificar a presença do princípio interativo de Shneiderman.

Esta visualização interativa é mais uma solução elaborada pelo SAPO Labs face à imensidão de dados e à necessidade de os representar graficamente em função do tempo, onde a interatividade permite a navegação nos dados – necessidade que liga este projeto ao proposto nesta dissertação.

Computacionais	Implementação	Perceção	Variável
Tamanho	Área	Quantitativa	Publicações
Cor	Área	Seletiva	Categoria

Tabela 2.9 Análise Semiótica do gráfico de rede do projeto “Máquina do Tempo” .

Variáveis Visuais	Implementação	Percepção	Variável
Dim X	Linha	Associativa	Tempo
Dim Y	Linha/ Área	Quantitativa	Total notícias
Tamanho	Linha	Associativa	Intensidade da relação

Tabela 2.10 Análise Semiótica do gráfico *timeline*, do projeto "Máquina do Tempo".

<http://maquinadotempo.sapo.pt/> – disponível em 28 de janeiro de 2015

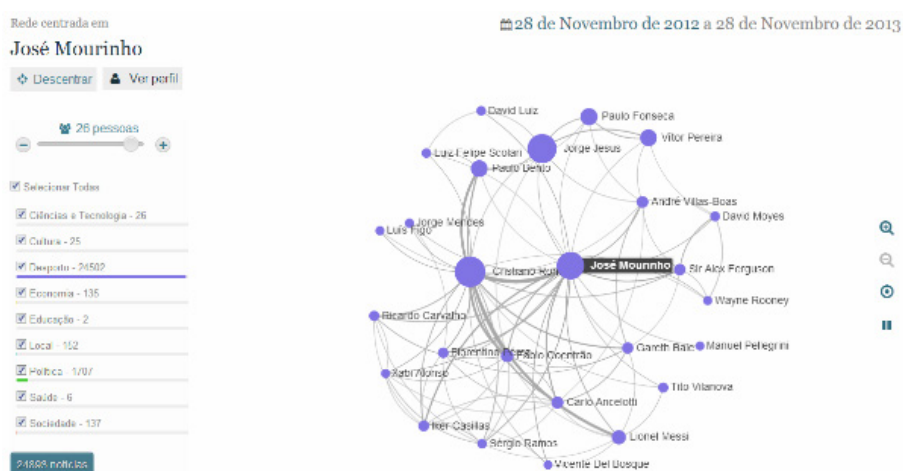


Figura 2.36¹ Snapshot do gráfico de rede, na visualização interativa "Máquina do tempo", cuja personalidade central é José Mourinho.

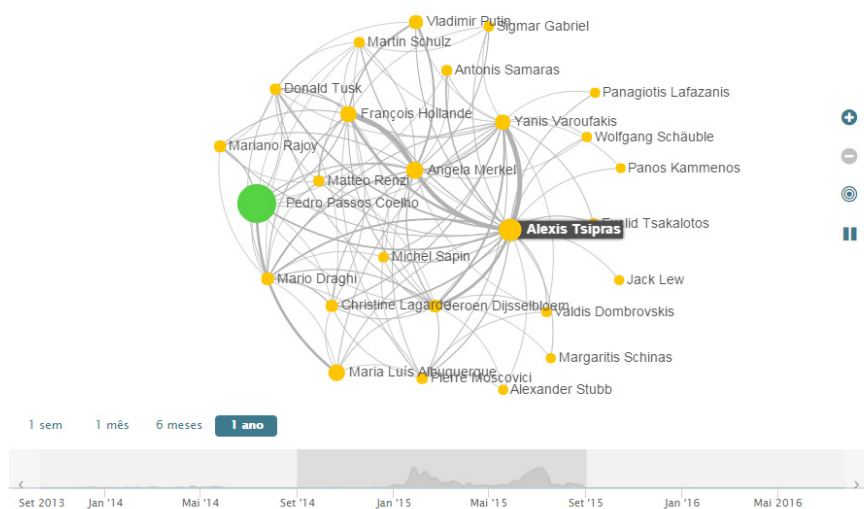


Figura 2.37¹ Snapshot do gráfico rede e *timeline* (mais a baixo), da "Máquina do tempo".

2.4.1 Trabalhos Relacionados com Política

“**Election Results 2008**” é um projeto realizado em 2008, pelo *New York Times U.S.*, durante as eleições presidenciais para mostrar visualmente os resultados dos votos por municípios e estados. Utiliza o mapa para a visualização destes dados e inclui resultados de eleições anteriores, desde 1992, totalizando cinco eleições possíveis de ver representadas. Esta visualização permite ainda, que o utilizador selecione um de quatro modelos de visualização disponíveis – *State winners, County bubbles, County leaders, Voting shifts* –, podendo alternar a visualização dos resultados. Uma das técnicas usadas nesta visualização é conhecida por mapa coroplético – uma das mais populares para a representação de dados estatísticos utilizando símbolos com área mas, na qual, a área não representa quantidade. Para evitar confundir a área com quantidade é empregue uma escala colorida na qual existe uma variação de saturação na cor.

A cor, independentemente da sua saturação, é empregue para distinguir a afiliação política dos votantes, sendo que o azul representa os votos a favor dos democratas e os vermelhos a favor dos republicanos. A área representa quantidade, havendo uma relação de proporcionalidade relativamente à soma dos votos a favor do candidato à liderança em cada município. Para esta representação quantitativa estão disponíveis cinco escalas que permitem alterar a área, logo o tamanho, dos círculos de maneira proporcional de forma a facilitar conclusões face às diferentes perspetivas.

Variáveis Visuais	Implementação	Percepção	Variável
Tamanho	Área	Quantitativa	Votos
Cor	Área	Seletiva	Filiação política
Dim X, Y	Área	Associativo	Região

Tabela 2.11 Análise Semiótica da visualização “Bubbles”, do projeto “Elections Results 2008”.

<http://elections.nytimes.com/2008/results/president/map.html>

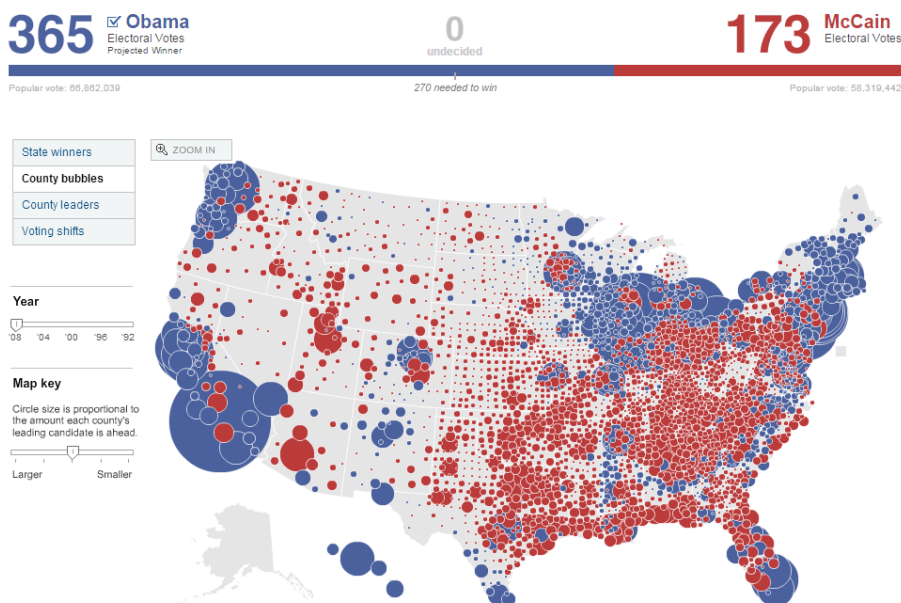


Figura 2.38 Projeto “Election Results 2008” - visualização das contagens dos votos através de áreas circulares.

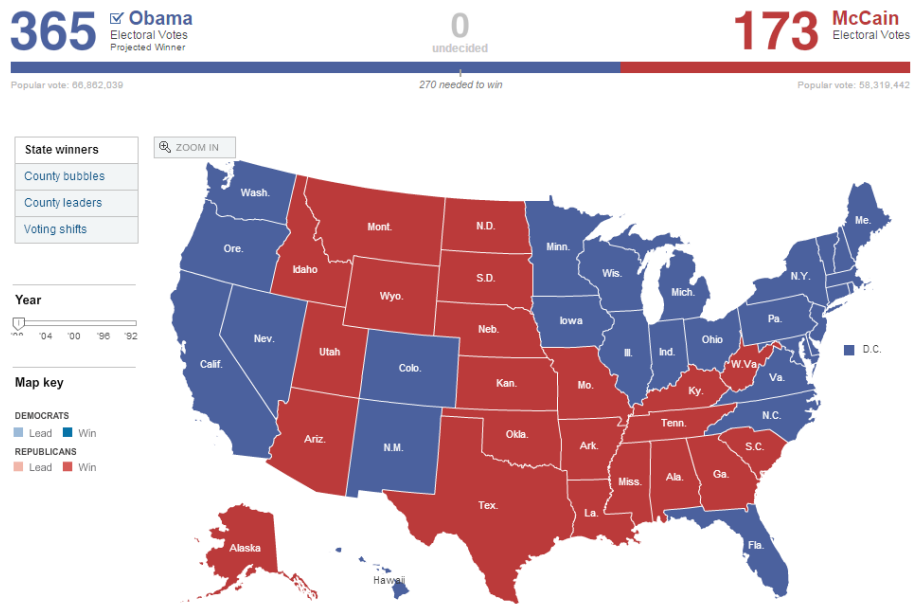


Figura 2.39ª Projeto "Election Results 2008" - visualização do candidato vencedor em cada estado através de um mapa coroplético.

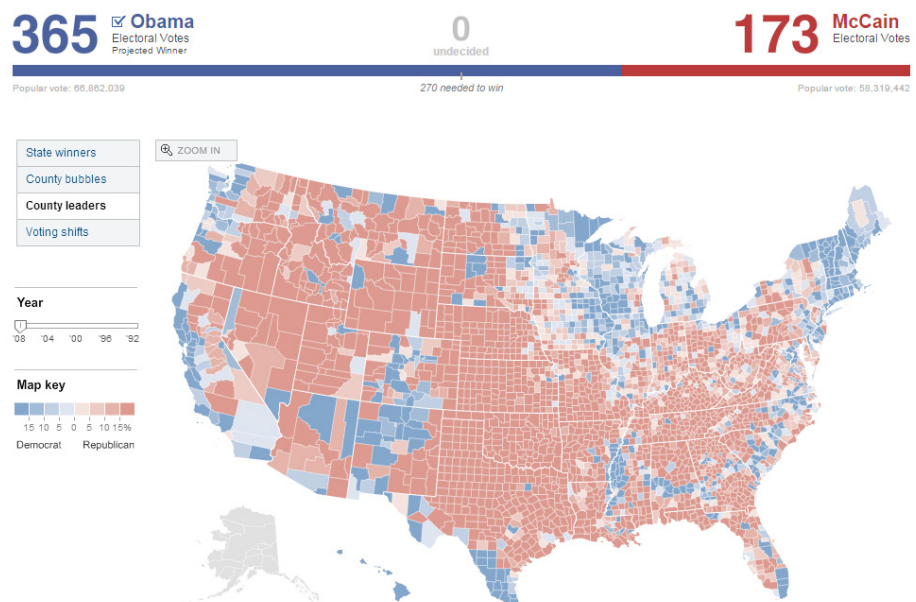


Figura 2.40ª Projeto "Election Results 2008" - visualização do candidato vencedor em cada região através de um mapa coroplético.

“Um país como nós” é um projeto desenvolvido pelo SAPO Labs em conjunto com a Universidade de Aveiro, o LED –Laboratório da Experiência e Design – e com a Universidade da Beira Interior. Este projeto *web* permite ao utilizador visualizar e interagir com dados estatísticos relativos aos concelhos portugueses, através de três modalidades. A primeira modalidade – Retrato – permite ao utilizador criar o seu retrato através de dados estatísticos. À medida que o utilizador vai inserindo informações pessoais, a visualização vai mostrando em que parte das estatísticas se engloba o utilizador.

Na segunda modalidade - Estatísticas - podem ser visualizados dados estatísticos relativos a fatores socioeconómicos como População, Emprego e Economia, Educação e Saúde. Estes dados são aglomerados consoante o distrito escolhido e mostrados por localidade, sendo possível comparar várias localidades do mesmo distrito.

A terceira modalidade - Autárquicas - é a modalidade que mais interessa para este projeto devido ao seu tema. Representa os resultados estatísticos das eleições autárquicas portuguesas, desde 1976 a 2013, utilizando um mapa coroplético, no qual as cores representam o partido com maior votação no concelho selecionado, descartando todos os outros partidos participantes nas eleições e destacando o partido vencedor. Para uma visualização mais detalhada dos resultados das votações nesse concelho, este projeto utiliza um gráfico de barras horizontal que auxilia a visualização principal e fornece as quantidades em percentagem dos votos em cada partido participante na eleição, sendo ainda possível recolher informação relativa à abstenção, percentagem de votos brancos e nulos. Existe ainda a possibilidade de analisar os dados estatísticos a nível global e visualizar a quantidade de câmaras por partido.

Variáveis Visuais	Implementação	Perceção	Variável
Cor Dim X, Y	Área Símbolo	Seletiva Associativa	Partido vencedor Concelho

Tabela 2.12 Análise Semiótica do mapa coroplético, do projeto “Elections Results 2008”.

Variáveis Visuais	Implementação	Perceção	Variável
Cor Dim X, Y	Área Símbolo	Seletiva Associativa	Partido vencedor Concelho

Tabela 2.13 Análise Semiótica do gráfico de barras, do projeto “Elections Results 2008”.

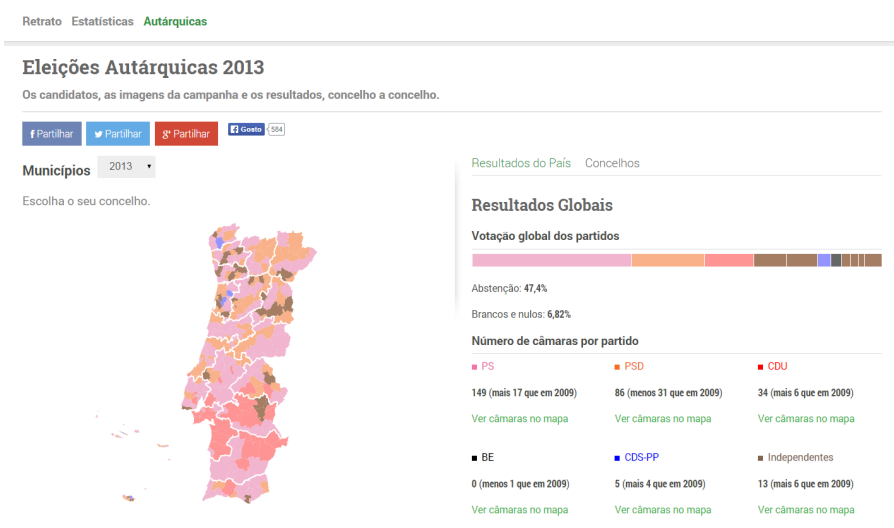


Figura 2.41' Projeto "Um país como nós". Visualização dos partidos vencedores, em cada concelho, nas eleições autárquicas de 2013.

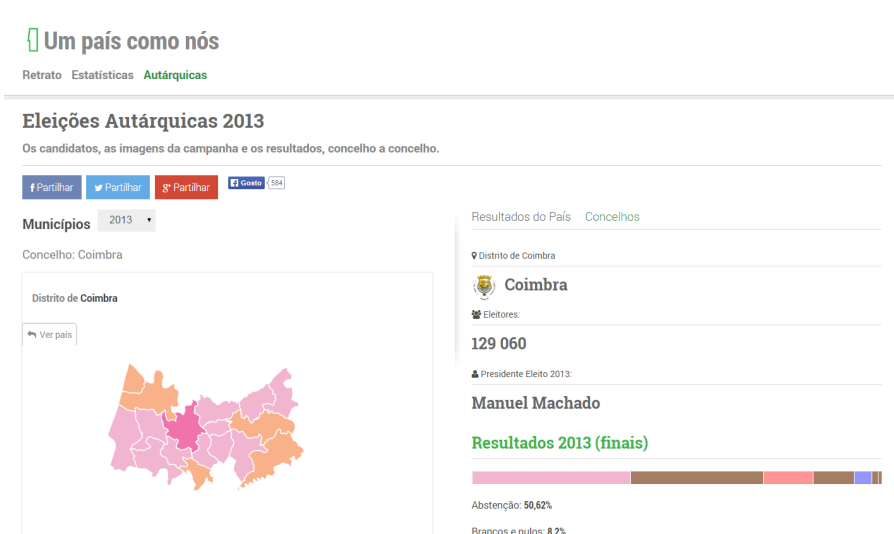


Figura 2.42' Projeto "Um país como nós". Visualização dos partidos vencedores, centralizada no distrito de Coimbra.

"Election 2015: The Guardian poll projection" é um projeto desenvolvido pelo *The Guardian*, publicado em 2015, sobre os resultados eleitorais em Inglaterra do mesmo ano. Este projeto faz uma análise dos resultados recolhidos nas sondagens, mostrando visualmente qual o partido que teria ganho as eleições nesse dia e quais as alianças que conseguiriam o número de votos necessários para governar. É possível visualizar, também, a variação individualizada (diminuição ou aumento) da quantidade de lugares atribuídos a cada partido ao longo do tempo, consoante as sondagens.

A representação da variação de percentagens, através de gráficos de linhas, permite ao utilizador verificar que, apesar de haver variações nas percentagens, existem sempre dois partidos dominantes ao longo

do tempo, desde Dezembro até Maio. A identificação destes partidos é feita através das cores selecionadas para a sua representação.

São ainda apresentados vários gráficos de fluxos para mostrar, no geral e de forma individualizada, quais os lugares no governo que estão a mudar de lado, ou seja de partido, comparando o governo anterior com o governo projetado pelo *The Guardian*.

Ainda sobre o tema anterior – as mudanças dos partidos para cada lugar –, este projeto tem outra modalidade de visualização, que utiliza uma representação similar aos mapas para mostrar as batalhas regionais entre os partidos, mostrando os lugares ganhos ou perdidos.

<http://www.theguardian.com/politics/ng-interactive/2015/feb/27/guardian-poll-projection> – disponível a 30 de agosto

<http://www.theguardian.com/politics/ng-interactive/2015/apr/20/election-2015-constituency-map?src=nav> – disponível a 30 de agosto

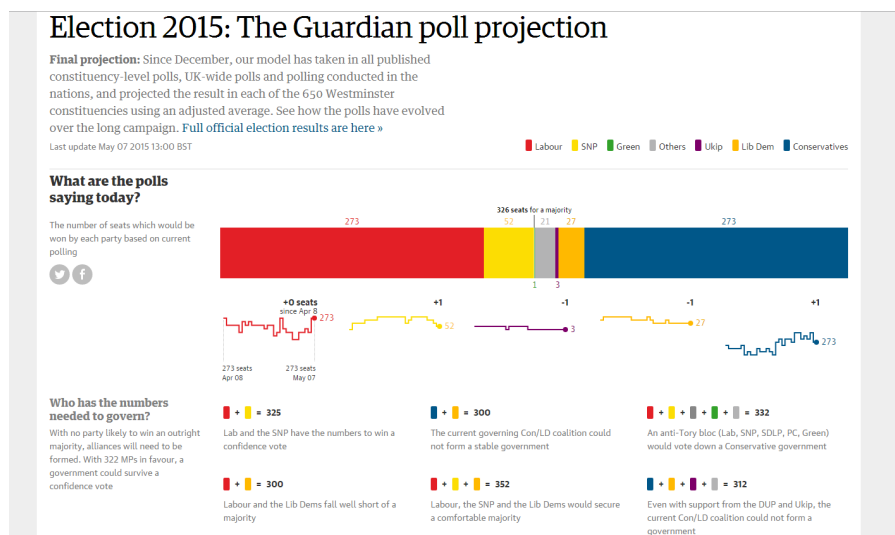


Figura 2.43* Visualização dos resultados estatísticos das sondagens para as eleições de 2015 na Inglaterra, seguido da visualização da variação do número de lugares atribuídos a cada partido e, ainda, estudo das coligações necessárias para obter o número de votos necessários para governar.

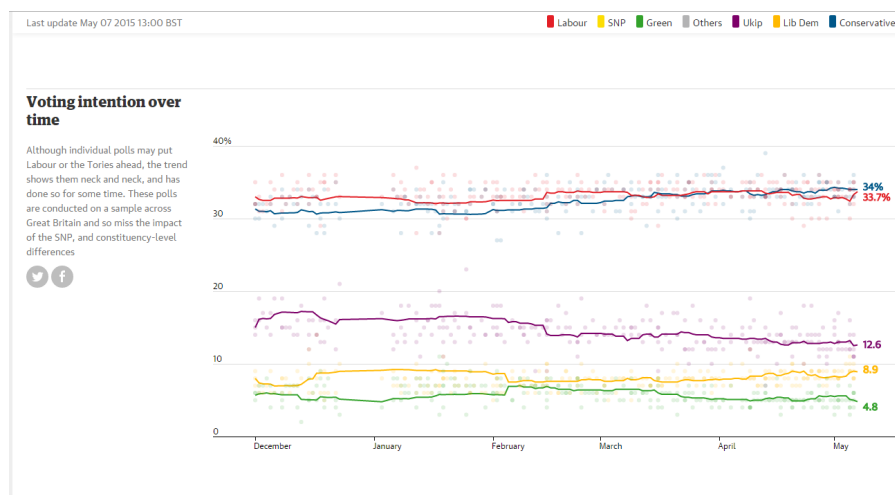


Figura 2.44* Visualização da variação na intenção de voto ao longo do tempo.

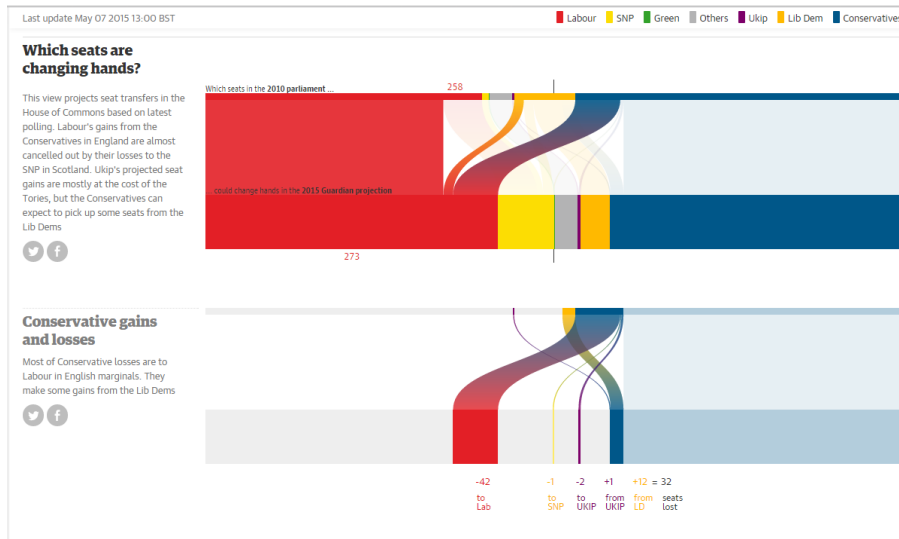


Figura 2.45^z Visualização dos lugares que estão a mudar de partido, através de interação com o gráfico (em cima) e visualização individualizada por partido (em baixo) .

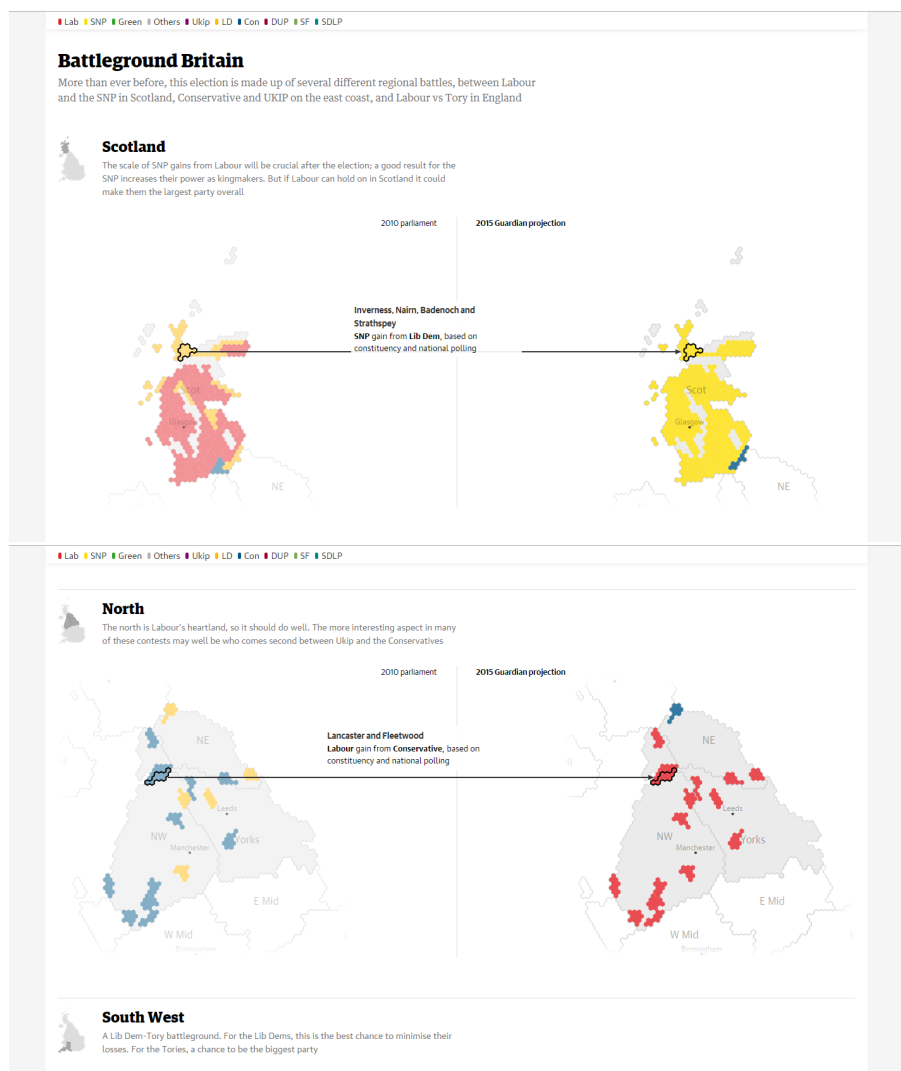


Figura 2.45^z Visualização das batalhas travadas pelos partidos das diferentes regiões da Inglaterra, mostrando perdas e ganhos dos partidos relativamente ao parlamento anterior .

“Um Sistema Político-Empresarial” é uma visualização interativa, desenvolvido pelo português Pedro Cruz, em 2013, que faz uso de uma metáfora biológica para a construção de uma narrativa, em torno das relações entre membros do governo Português com empresas ou grupos. Os dados utilizados são referentes a um período de tempo com uma duração de trinta e oito anos, mais exatamente entre os anos de 1975 e 2013. No entanto, o tempo não é apresentado como nos projetos anteriores, não tem uma representação nem linear nem espiral, não associa sequer os dados a uma data. Todos os políticos – representados como microorganismos vivos e com um tamanho que dita o número de empresas com que dado político se envolveu, bem como as suas atividades – permanência nessas mesmas empresas – são exibidos desde o início, sem que haja uma data associada.

A noção de tempo é dada apenas pelo tempo real que o microorganismo demora a circundar a empresa, que é representada por áreas com tamanho proporcional ao número de políticos que ocupam uma posição nessa empresa. Também a cor é um atributo do político, ou do microorganismo, e representa o partido político ao qual o mesmo está associado.

A interatividade baseia-se no princípio de Ben Shneiderman ‘Visão geral primeiro, *zoom* e filtro, e então detalhes sobre a procura’, sendo que, também faz uso do paradigma da manipulação direta dos dados, através da digitalização de palavras por parte do utilizador.

Este projeto relaciona-se com o nosso projeto por ser uma visualização interativa para *web*, que representa uma grande quantidade de dados – cerca de 103 políticos e 350 empresas –, apesar do período temporal ao qual correspondem não ser muito extenso. Este projeto inova pelo facto de decidir dar menos importância ao tempo, dando ênfase às relações entre políticos e empresas – a história que está a ser contada pelos dados. A noção de tempo é ligeiramente transmitida pelos segundos que os microorganismos demoram a visitar a empresa.

Variáveis Visuais	Implementação	Percepção	Variável
Tamanho	Área	Quantitativa	Tamanho empresa
Tamanho	Pictograma	Quantitativa	Nr. de empresas
Cor	Pictograma	Seletiva	Filiação política
Dim x, y	Pictograma	Associativa	Empresa

Tabela 2.14 Análise Semiótica do projeto “Um Sistema Político-Empresarial”.

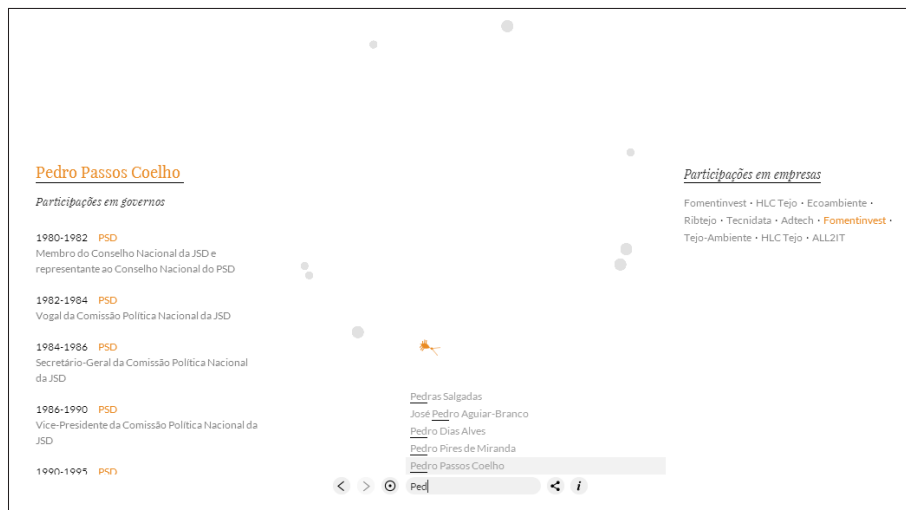
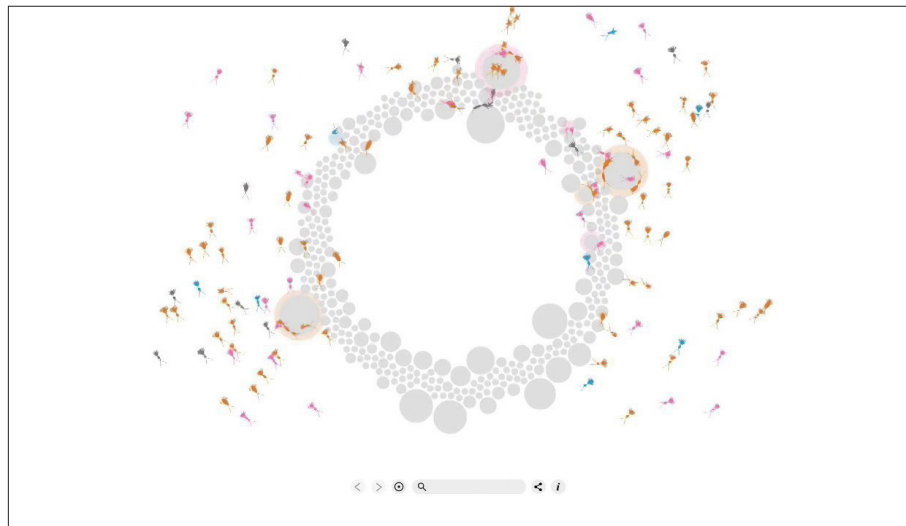


Figura 2.46^p Snapshots da visualização interativa "Um Sistema Político-Empresarial". Em cima, temos uma visão geral inicial. No meio, observamos manipulação direta de dados e, em baixo, temos detalhes resultantes da seleção de um elemento.

2.4.2. Conclusões

Atualmente em Portugal, os resultados políticos são muito divulgados através de gráficos de barras e gráficos Tarte/Queijo, que informam sobre as percentagens das votações. Estas visualizações utilizam como atributos principais a cor e o tamanho para representar os diferentes partidos e a quantidade de votos respetivamente. Como podemos observar no projeto “Um país como nós”, este tipo de gráfico é utilizado para essa mesma função, sendo a informação complementada com um mapa coroplético, permitindo a visualização de padrões por concelhos.

Técnicas similares de representação de dados estatísticos relativos a eleições são utilizadas nos projetos internacionais dos jornais *New York Times* e *The Guardian*, podendo variar ligeiramente nas questões a que pretendem responder. Desta forma, é garantida que a informação representada chega a qualquer pessoa face à sua simples interpretação, sendo que a perceção e compreensão dos dados é quase automática.

Apesar da popularidade das visualizações de informações políticas focadas em dados eleitorais, existem visualizações dentro do mesmo tema – a política – que têm objetivos mais profundos e, de certo modo, delatadores. A visualização de informação “Um Sistema Político-Empresarial” é um dos projetos que faz parte desta categoria, mostrando a relação dos políticos que fazem – ou fizeram – parte do governo português com empresas portuguesas. Para além de explorar uma nova perspetiva sobre a política, contando uma história de forma inovadora, este projeto emprega a criatividade na representação visual, passando uma mensagem secundária relativamente aos dados. Com este lado mais criativo, este projeto cria empatia com o utilizador, compelindo-o a explorar mais e descobrir como interagir, enquanto percebe os dados transmitidos. É este o intuito que se pretende atingir com a nossa visualização, dando espaço para a parte concetual inerente à área criativa – o Design – na escolha da forma gráfica.

Concluí-se que os dados eleitorais, principalmente os que se referem a um único tipo de eleição, não têm um horizonte temporal muito extenso. Por exemplo, desde 1976 ao ano presente são totalizados cerca de quatro décadas, no entanto, apenas em catorze anos desse período aconteceram eleições legislativas em Portugal. São, no entanto, dados que têm tendência a aumentar a sua dimensão temporal de acordo com a passagem do tempo no mundo real.

Este tipo de dados é muito divulgado fazendo uso de gráficos ‘tarte’ ou barra, pelo que, já se tornaram comuns e de acessível leitura para todo o público. Contudo, este projecto não pretende fazer uso destas abordagens, procurando explorar novas formas gráficas que permitam uma leitura semelhante e, mais interessante, do mesmo tipo de dados.

Também foi possível averiguar que a escolha de uma forma para representar o tempo, varia consoante o espaço em que se pretende inserir a visualização e a dimensão dos elementos representados. A melhor forma de representar grandes quantidade de tempo, de forma contínua consiste na utilização de uma *timeline* horizontal/ vertical ou elíptica. Uma distribuição circular do tempo pode representar, também, uma boa solução, mas esta análise só pode ser efetuada quando os dados são definidos, conhecidos e adquirem uma forma visual.

Abordagem

Este capítulo dedica-se à descrição dos objetivos e das metodologias adotadas para a concretização dos objetivos estabelecidos. Iremos também apresentar um planeamento das principais tarefas do projeto.

3

3.1 Objetivos

Um dos objetivos essenciais com esta dissertação é a investigação de visualizações interativas de informação, que lidem com cenários que representam dados em larga escala em função de um longo período temporal e que utilizem a interação como forma de navegação e exploração dos mesmos. O propósito desta investigação é compreender a evolução das formas gráficas para a representação de dados, fazer o levantamento de princípios padronizados aplicados à construção de bons-gráficos e adquirir noções teóricas dos métodos ou princípios interativos que guiam a construção de uma visualização de informação interativa.

Para aplicação das soluções encontradas ao longo da investigação foi desenvolvido um protótipo funcional e interativo que permite a navegação e exploração de dados a um potencial utilizador.

3.2 Metodologias

Na primeira fase, é feito um levantamento histórico da evolução das formas gráficas para a representação gráfica de dados e transmissão clara de ideias. Para isso recorreremos a bibliografia, capaz de nos informar brevemente sobre a origem dos gráficos como formas gráficas representativas de dados e emergência de novas formas gráficas. Seguidamente, centramos a nossa investigação na interação aplicada à visualização de informação, examinando vários artigos de trabalhos realizados nesta área. A recolha de princípios de design de bons gráficos e conceitos semióticos nos gráficos, foi o passo seguinte. Em paralelo, foi feita uma pesquisa e recolha de projetos na área da visualização de informação, cujas características se adequam ao projeto proposto. Utilizamos estes projetos para um breve estudo semiótico das marcas e perceções, apoiado na obra de Jacques Bertin – “*Semiology of Graphics*”.

Na segunda fase, na qual desenvolvemos o protótipo, foi necessário definirmos o tipo de dados que nos interessa ver tratados. O conjunto de dados tem como propriedades essenciais a quantidade elevada de dados e um período temporal extenso. O assunto por eles tratado, a história por eles contada, decidiu-se em função dos interesses de ambas as partes envolvidas nesta dissertação (aluna e empresa).

Seguiu-se a análise dos dados para a escolha de uma forma gráfica adequada às suas necessidades, apoiada nos princípios de design anteriormente coletados. Seguidamente, foi necessária a extração dos dados, em bruto, contidos na base de dados do SAPO, através de uma linguagem baseada em *queries – jQuery*. Estes dados passaram por um processo de conversão para um formato adequado à utilização na visualização e foram filtrados, garantindo, apenas, a armazenagem dos dados relevantes para a visualização. Após estas duas etapas, os dados estavam prontos para serem representados através das formas visuais escolhidas. Estas formas visuais passaram por um processo de transformação face às necessidades dos dados, ou seja, foram refinadas para mostrar informações interessantes.

Os passos seguintes consistiram na criação de interação para o protótipo, a sua implementação na *web* e avaliação da mesma. Para a realização destes passos foram escolhidas as linguagens de

programação mais adequadas, tendo em conta a sua localização na *web*. A avaliação foi apoiada no estudo de Nielsen¹, que comprova que, os melhores resultados numa avaliação de usabilidade, necessitam de cinco utilizadores, no mínimo, para realizem várias pequenas tarefas, possibilitando o despiste dos erros de usabilidade mais relevantes, no protótipo desenvolvido.

3.3 Planeamento

Na primeira fase, que se estendeu até à data da defesa intermédia, o plano de trabalho define uma previsão da calendarização das principais tarefas deste projeto – leitura de referências, escrita da dissertação, definição de metodologias, definição de objetivos, estado da arte, interface gráfica, implementação e testes de usabilidade.

A leitura de referências e a escrita da tese são as duas tarefas cuja duração é maior, sendo parte do projeto desde o início até ao fim, enquanto que os objetivos e metodologias foram definidos na primeira fase do projeto, guiando os temas que necessitavam de ser abordados no estado da arte.

No segundo semestre e segunda fase do projeto, que decorreu entre fevereiro e junho, foi posta em prática a proposta de *interface* gráfica, a implementação do protótipo e os testes de usabilidade. Esta última tarefa, previa-se ser dividida em duas fases para que pudessem ser feitos testes de usabilidade que detetassem problemas na usabilidade e respetivas correções dos mesmos.

Os resultados da primeira fase de avaliação determinaram as melhorias necessárias para uma segunda fase de implementação. A última fase de testes, que se engloba nos trabalhos futuros, servirá para apurar a eficiência das soluções encontradas para resolver os problemas apurados na primeira fase de testes.

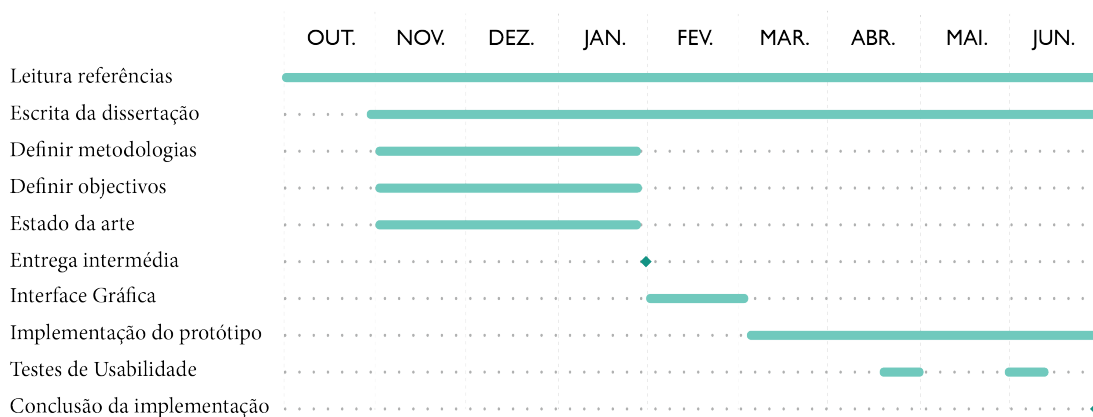


Figura 3.1 Plano de trabalho previstos.

Segue-se a descrição e clarificação das tarefas realizadas na segunda fase deste projeto, das quais, algumas não estavam planeadas.

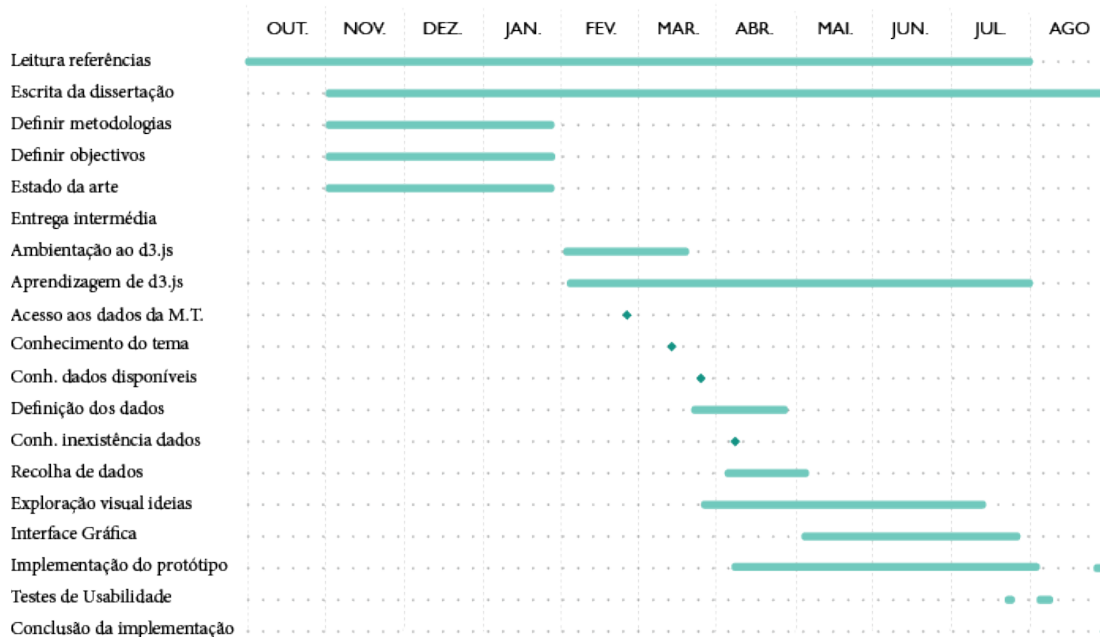


Figura 3.2 Plano de trabalhos realizados.

Ambientação ao *d3*

No início deste projeto estava aponte que a linguagem de programação a utilizar para a implementação da forma gráfica dos dados seria o *Processing.js*. Contudo, no início da segunda fase do projeto – fase do desenvolvimento prático – foi aconselhada a utilização do *d3.js*, visto ser conhecido no SAPO e pensar-se ser mais versátil para *web*. Assim, foi lançado um primeiro desafio com o objetivo de aprendizagem e ambientação ao *d3*, cuja proposta se encontra no anexo E .

Aprendizagem do *d3*

A tarefa aprendizagem do *d3* difere da tarefa ambientação ao *d3*, na medida em que a segunda foi desenvolvida ainda em fase de experimentação, tendo como único objetivo a ambientação. A aprendizagem do *d3* é uma tarefa que começa então com a tarefa de ambientação, mas que acompanha todo o projeto, face à constante pesquisa de soluções para os problemas de código encontrados ao longo do desenvolvimento do projeto.

Conhecimento do tema

O conhecimento do tema para projeto não é uma tarefa, mas sim um evento. Marca a data em que o SAPO transmitiu o tema que queria ver tratado no projeto – as eleições.

Conhecimento dos dados disponíveis

O conhecimento dos dados disponíveis é o evento que marca a altura em que foi recebido o design *Statement* por parte do SAPO. Neste email, era explicado o objetivo do projeto, o público alvo e os dados que poderiam ser disponibilizados - Anexo B.

Definição dos dados

Esta tarefa tem o seu início pouco depois de ser conhecido o tema do projeto. Nesta tarefa foram apontadas ideias dos dados que poderiam ser representados, foi feita pesquisa sobre projetos com o mesmo tema ou temas similares. Foi também, feita pesquisa sobre o sistema político português a fim de perceber ligações políticas e o funcionamento do governo, bem como o resultado das eleições

legislativas. Estas pesquisas ajudaram a formular as perguntas que suportaram a escolha dos dados a trabalhar.

Acesso aos dados da Máquina do Tempo

Este evento marca a data em que foi explicado o processo de recolha dos dados relativos a notícias. Nesta data foram realizados os procedimentos necessários, por parte do SAPO, para que se pudesse ter acesso às respostas dos pedidos realizados através de *jQuery*. Foi ainda partilhado o ficheiro que contém os passos necessários à construção dos pedidos.

Conhecimento da inexistência de dados

Este evento marca a data em que se sentiu a necessidade de começar a trabalhar os dados das eleições legislativas e foi feito o pedido de acesso a esses dados. Contudo, os dados que foram encontrados diziam, apenas, respeito às eleições autárquicas. Ficou, também, a saber-se que os dados das eleições autárquicas tinham sido recolhidos manualmente. Nesta data surgiram, então, tarefas que não estavam planeadas.

Recolha de dados

Apesar de não estar no planeamento elaborado no primeiro semestre, esta tarefa começa um pouco antes do conhecimento da inexistência de dados das eleições legislativas, pois, face à decisão de se representar fatores socioeconómicos portugueses, sabia-se de antemão que seria necessário recolher estes dados dos sites de estatísticas: INE ou Pordata. Contudo, esta tarefa toma proporções de tempo superiores às esperadas, estendendo-se em aproximadamente um mês, face à descoberta da inexistência de dados relativos às eleições legislativas, sendo necessário localizar os dados e fazer a recolha dos mesmos.

Implementação

A implementação do projeto começou no início de abril e estende-se até à primeira semana de agosto, sendo interrompida, entre a segunda e a terceira semana de agosto, para realização de testes de usabilidade. Após esta interrupção, a continuação desta tarefa tem como objetivo melhorar problemas de usabilidade, aplicar alterações sugeridas pelo SAPO e outras sugestões, que resultam da análise dos problemas encontrados durante os testes.

Exploração visual de ideias

Esta tarefa diz respeito ao esboço manual de ideias para a visualização dos dados definidos. Tem início pouco depois de ter sido recebido o *Statement* de design. É uma das tarefas com maior duração, acompanhando a tarefa de implementação e esboçando soluções para problemas emergentes na programação.

Interface Gráfica

Esta tarefa tem o seu início no mês de maio e é finalizada no final do mês de julho. Corresponde ao design de soluções para a ligação das visualizações idealizadas no ecrã e integração de legendas e interação. Tal como a tarefa de exploração visual de ideias, esta é uma tarefa que acompanha a implementação do projeto, facilitando a averiguação do que é possível realizar – ou não – através de código.

Testes de usabilidade

Esta tarefa diz respeito não só à observação e realização dos testes com utilizadores reais mas, também, ao tempo necessário para a sua elaboração. A análise dos resultados já não é incluída nesta tarefa, estando incluída na escrita da tese.

Proposta

Este capítulo tem como objetivo contextualizar o leitor acerca do tema e do público-alvo deste projeto, explicar todo o processo realizado para a definição e recolha dos dados a utilizar, bem como para a construção da narrativa e metáfora que transmitirá a mensagem da visualização de informação. Finaliza com a descrição de alguns cenários de utilização que facilitará, ao leitor, a compreensão do objetivo e funcionalidades deste projeto.



4.1. Tema e Público-alvo

Com vista às eleições legislativas que acontecerão neste ano – 2015 –, o SAPO tem interesse na abordagem deste assunto, propondo-o para o tema deste projeto: as eleições portuguesas. Pretende, então, a criação de um portal *web* que permita aos seus utilizadores tirar partido de visualizações de informação interativas e dinâmicas para explorar dados eleitorais portugueses, à semelhança do que acontece na visualização “Um país como nós” mas, neste caso, com dados das eleições autárquicas.

Há interesse na contextualização de parte da história política do país, relativo às eleições legislativas, no acompanhamento das eleições legislativas do ano atual, na representação de dados relativos a notícias – como acontece na Máquina do tempo – bem como na sugestão de novos dados que possam trazer interesse ou valor à visualização.

O público-alvo é qualquer cidadão português capaz de aceder à internet – ou seja, num utilizador comum – nomeadamente eleitores, jornalistas ou politólogos.

4.2. Definição dos Dados

Numa fase inicial, ainda sem acesso às bases de dados do SAPO, foram enumerados dados, aos quais se poderia ter acesso mais tarde.

- Sondagens / intenções de voto dos principais partidos políticos desde 2011
- Estatísticas eleitorais (resultados das autárquicas e legislativas por concelho) desde 1974
- Notícias com identificação dos candidatos desde 1987
- Distribuições das notícias e citações por candidatos desde 1987
- Tweets / posts do facebook sobre tópicos e pessoas desde 2011
- Todos os dados disponíveis em projetos como a Máquina do Tempo, Um País como Nós, POPSTAR
- Imagens publicadas nas redes sociais sobre candidatos ou partidos (ex. cartazes eleitorais)

Um país como Nós
<http://umpaiscomonos.labs.sapo.pt/Autarquicas>

Máquina do Tempo
<http://maquinadotempo.sapo.pt/>

Popstar
<http://www.popstar.pt/>

Após conhecimento dos dados que estavam disponíveis e com que era possível trabalhar, o passo seguinte foi a definição dos que interessavam trabalhar dentro dos dados disponibilizados. Neste passo é importante decidir qual a mensagem que se pretende passar, qual o objetivo da visualização, quais as questões a que se pretende responder. Para a sua realização, foi importante pesquisar sobre o tema das eleições e perceber de que forma funciona o sistema eleitoral português.

Existem três tipos de eleições: eleições presidenciais, legislativas e autárquicas. As eleições legislativas, também conhecidas como eleições para a assembleia da república, elegem não só o Primeiro-Ministro, mas também, os partidos que representam o novo governo, sendo que a votação define quantos políticos de cada partido vão constituir esse governo. Na totalidade são eleitos 230 deputados. As candidaturas são apresentadas pelos partidos políticos, isoladamente ou em coligação, podendo as listas integrar cidadãos não inscritos nos respetivos partidos – independentes.

Decidiu-se primeiramente, limitar os dados às eleições legislativas,

visto serem, estas, o foco do ano em causa e pelo facto do SAPO já dispor de uma visualização com eleições autárquicas. À medida que o conhecimento desta área aumentava, algumas questões foram surgindo, ajudando a definir os dados que seriam interessantes de representar.

“Quantas vezes, um político já se candidatou? Quantas vezes ele ganhou as eleições?”

Quais os partidos que têm estado no poder?

O que faz de um político um bom político?

Quem passou por cada cargo do governo? (desde 1976)

Por que cargos passou cada político? (vista centralizada no político)

Quais os melhores governantes? O que melhorou ou piorou de ano para ano ou de mandato para mandato.

Que cortes, incentivos foram feitos durante cada mandato?”

No topo do mais importante para a visualização estava a história política – resultados eleitorais, partidos com representação no governo em cada mandato, político no cargo de Primeiro-Ministro, líderes de cada partido no governo. Era importante, também, conhecer cada político, conhecer os seus feitos, a sua popularidade, até mesmo comparar com outros políticos e possibilitar a construção de perfis. A esta necessidade seria possível responder através da representação e exploração das publicações nos *media* relacionadas com cada político. E, finalmente, era importante mostrar como é que o governo afeta o país a nível social e económico, ou seja, mostrar as alterações em fatores socioeconómicos como a natalidade, a mortalidade, a educação, o emprego e o desemprego, a emigração.

Ainda sem acesso aos dados, houve uma primeira tentativa e necessidade de compreender os níveis dos dados a representar para a realização dos primeiros esboços. Os políticos têm associado um partido, um nome e um cargo. As notícias têm um político associado, têm um título, uma data e hora de publicação, uma fonte e um link para a notícia. E os fatores socioeconómicos têm um tema, um valor numérico, uma data. Numa primeira análise destacou-se a diferença no período temporal dos diferentes dados. Enquanto os dados das eleições são referentes a um período temporal de trinta e nove anos (1976 a 2015) e, apesar de apenas catorze serem referentes a eleições, as notícias só começaram a ser recolhidas no ano de 1987 – período de vinte e oito anos. Já o período temporal dos dados dos fatores socioeconómicos variam de tema para tema, podendo por vezes haver temas incompletos, ou seja, com falta de dados em determinados anos.

Outras perguntas foram surgindo aquando a realização de esboços:

“Com que frequência estão a ser comentados partidos ou líderes partidários nas redes sociais? Qual o sentimento – positivo ou negativo – relativo a cada comentário?”

Poderia responder a estas perguntas utilizando os dados do projeto POPSTAR que desenvolveu métodos de recolha, medição e agregação de opiniões políticas e económicas, publicadas no twitter, na blogosfera, nas notícias e através de sondagens. Apresenta, assim, menções a líderes político-partidários nos meios anteriormente mencionados, sentimentos transmitidos através de *tweets* sobre os líderes e intenções de votos nos partidos, recolhidos através de sondagens. Contudo, face aos meios utilizados pelo projeto POPSTAR, o período temporal deste dados é mais curto do que qualquer outro anteriormente analisado,

sendo que apenas tem uma duração de três a quatro anos (2011-2015). Por estas razões e pelo facto de já estarem a ser utilizados dados de mais três temas diferentes – eleições legislativas, notícias e fatores socioeconómicos – estas últimas questões foram descartadas, evitando maior complexidade de dados que poderia confundir o utilizador.

4.2.1. Recolha de Dados

Apesar de todos os dados escolhidos para a visualização terem sido dados como disponíveis, a realidade é que nem sempre existe uma noção completa de todos os dados armazenados pelas bases de dados do SAPO. Assim, houve necessidade de fazer pesquisa, recolher e tratar dados através da elaboração de ficheiros CSV (valores separados por vírgulas) – tarefas que não estavam planeadas e para as quais foi dispensado o mínimo período de tempo possível.

4.2.1.1 Dados Eleitorais

Com os dados eleitorais pretende-se explorar informações como: resultados eleitorais das legislativas desde 1976 a 2015, líderes políticos de cada partido com representação no governo e sondagens de 2015. Para os resultados eleitorais e para os políticos com representação no governo, veio a descobrir-se que apenas existiam armazenados dados referentes às eleições autárquicas. Quanto aos dados relativos às sondagens, estavam disponíveis no projeto POPSTAR através de um ficheiro CSV, pelo que apenas foi necessário o seu descarregamento e posterior compreensão dos campos utilizados.

A pesquisa para a recolha dos dados relativos aos resultados eleitorais incidiu primariamente no site da Comissão Nacional de Eleições²¹, onde a recolha de dados se mostrou uma tarefa dificultada pelo meio de divulgação de dados - PDF. Para além do difícil acesso aos dados, também se verificou que apenas existiam dados relativos a cada distrito e não a cada região, impossibilitando a aplicação das soluções visuais esboçadas antes da recolha dos dados. Enquanto se procedeu à elaboração de ficheiros CSV com base nos dados dos PDF's recolhidos, uma nova pesquisa no PORDATA facilitou a finalização desta recolha através do *download* do ficheiro CSV, relativo aos resultados das eleições para a Assembleia da República – eleitos por partido ou coligação.

Outra informação que seria necessária recolher dizia respeito aos políticos com representação no governo ou candidatos ao cargo de Primeiro-Ministro nos anos das eleições. Esta recolha mostrou-se ainda mais difícil e demorada do que a anterior. Inicialmente recorreu-se à pesquisa dos líderes partidários através das páginas oficiais dos partidos, mas este veio a mostrar-se um método ineficaz uma vez que alguns dos partidos constituintes do governo em anos mais antigos não possuem uma página oficial e a informação sobre estes partidos é escassa.

Recorrendo uma vez mais ao CNE²², foi possível recolher manualmente os nomes dos políticos no cargo de Primeiro-Ministro ao longo da história dos Governos Constitucionais, bem como dos líderes políticos de todos os partidos com representação no governo.

Durante esta recolha destacou-se o facto de nem todos os mandatos terem ocorrido nos anos entre as eleições, ou seja, nem todos os Primeiros-ministros foram colocados nesta posição pelo resultado de uma eleição. Isto aconteceu essencialmente em 1978 e 1979, no entanto, decidiu-se que, como não faziam parte das eleições, estas informações seriam ignoradas. Ainda mais tarde, também através do CNE, chegou-se aos candidatos eleitos mas, mais uma vez, estes dados estavam em PDF, restando a única hipótese de recolher manualmente centenas de nomes de políticos por cada ano eleitoral. Nesta fase foi decidido que se trabalharia apenas com os líderes políticos – candidatos à posição de Primeiro-Ministro – evitando maior gasto de tempo e registando os seus nomes num ficheiro de texto, depois guardado como csv.

4.2.1.2. Máquina do Tempo

A Máquina do Tempo é o projeto de visualização de informação do SAPO que utiliza uma grande quantidade de dados armazenados numa das bases de dados do SAPO. A base de dados que serve este projeto armazena notícias desde 1987 até à data atual. Estas notícias são recolhidas diariamente, notando-se uma subida na quantidade de notícias recolhidas nos anos mais recentes face à divulgação massiva de informação, que acontece cada vez mais rápido e com maior frequência. Face à grande quantidade de dados, era imperativo a utilização destes dados neste projeto, dentro do contexto das eleições. Esta é a base de dados mais importante do projeto e que podemos dizer que armazena dados em larga escala.

A recolha destes dados, para utilização neste projeto, é feita utilizando pedidos (*queries*) base e complementando com a especificação de certos parâmetros. O resultado do pedido utilizado devolve os dados em formato *JSON*, os quais posteriormente são processados e armazenados em *arrays* através de código em *JavaScript*.

Seguidamente é explicado todo o trabalho feito pelos meus colegas no SAPO para o eficaz armazenamento e tratamento das notícias na base de dados, dedicada aos dados da Máquina do Tempo.

Num primeiro nível, quando os dados são armazenados na base de dados, é atribuído um *id* à notícia, são identificadas as entidades presentes na notícia atribuindo um *id* à entidade, é identificada a categoria da notícia – política, desporto, cultura, educação, ciência e tecnologia, local, saúde, economia ou sociedade –, a fonte da publicação, a data e hora da publicação. A base de dados consegue ainda atribuir cargos às entidades, quando identifica a entidade e o cargo na mesma notícia, identificar as fotos e associar-lhe um *id* e identificar as citações associando-as à notícia.

Num segundo nível, numa nova base de dados, os dados são agregados consoante parâmetros específicos. É a esta última base de dados que se faz os pedidos *JSON*, atribuindo aos parâmetros os valores desejados.

4.2.1.3. Fatores Socioeconómicos

A recolha de dados relativos a fatores socioeconómicos decorreu sem grandes contratemplos, visto ser conhecida a localização dos dados desejados: PORDATA ou INE. Após observação e avaliação dos dados armazenados nas duas bases de dados, decidiu-se trabalhar com os ficheiros da PORDATA por estarem mais completos que os ficheiros do INE. Foram decididos os temas a trabalhar: População, Educação, Saúde, Emprego e Contas Públicas. Seguidamente procedeu-se à pesquisa dos dados relacionados com os temas referidos, disponíveis na PORDATA. Foram descarregados ficheiros respeitantes à dívida pública, às despesas do estado na educação, à população residente em Portugal, à população desempregada – por sexo –, à população ativa – total e por sexo –, nascimentos – total e fora do casamento –, óbitos – total e no primeiro ano de vida –, emigrantes – por tipo –, alunos matriculados no ensino superior – por sexo –, matriculas no ensino básico – por modalidade de ensino –, matriculas no pré-escolar – por sexo –, matriculas no ensino secundário – por modalidade de ensino –, docentes do ensino superior – por sexo –, docentes em exercício no ensino pré-escolar, básico e secundário, estabelecimentos de ensino pré-escolares, básico e secundário.

4.2.2 Problemática

Durante a recolha e análise dos dados identificaram-se dois dilemas dominantes. O primeiro consistia na grande variedade de temas disponíveis, que levantava questões de navegação: *“Como ligar os temas diferentes de forma harmoniosa?”*, *“Como representar todos estes dados num único ecrã?”*. E o segundo consistia na diversidade dos períodos temporais em cada tema, questionando a melhor forma de representar os dados de forma interessante: *“Como representar os dados, relativamente ao tempo, evitando conclusões precipitadas pela falta de dados na BD?”*.

O período temporal dos dados eleitorais abrange cerca de quatro décadas, começando em 1976 e terminando no ano corrente, no entanto, apenas catorze desses anos são referentes às eleições e, portanto, apenas esses produzem dados. No caso da Máquina do Tempo, o período de tempo apresenta dados diários de cerca de quatro décadas. E os dados relativos a fatores socioeconómicos referem-se a um período temporal que varia de tema para tema, podendo abranger um período temporal maior que os referidos para os dados anteriores.

O nosso projecto contempla, então, três períodos temporais diferentes face aos dados tratados, pelo que utilizamos o período maior para a definição da dimensão temporal da expressão “dados em larga escala”. Esta passa, assim, a ser definida por um período temporal de, aprodimadamente, meio século. Esta definição poderá compreender períodos temporais mais extensos, à medida que o tempo real avança e em caso de haver manutenção continuada ao projeto.

4.3. Visualização

Inicialmente as propostas visuais para a visualização dos resultados estatísticos, relativos às eleições legislativas, tiveram abordagens muito óbvias, baseadas em mapas coropléticos ou em mapas de símbolos (círculos). Em ambas as abordagens a cor era utilizada para distinguir os partidos políticos e o tamanho dos símbolos ou variação na cor para representar as quantidades dos resultados das votações. Estas abordagens permitiam distinguir partidos, quantidades de votos para cada partido e associar estes dados a concelhos, à semelhança do que acontece nos primeiros projetos descritos no em §2.4.2.

Face ao nome do projecto – “Novos paradigmas” – evitaram-se estas abordagens mais comuns que nada têm de novo, uma vez que, já são tao comuns nos dias atuais. Desta forma, foi essencial definir a narrativa e a metáfora visual que suportaria esta visualização, para ser conseguida uma abordagem mais criativa e diferente do que se tem vindo a fazer relativamente a este tema.

4.3.1. Narrativa

Determinar a forma dos dados é importante para justificar aspetos narrativos da visualização e apenas possível quando se conhece e compreende as características dos dados com que trabalhamos. Como é possível perceber, o governo português tem estado maioritariamente no poder de dois partidos dominantes – PS e PSD –, tal como alguns dos seus líderes têm vindo a ser repetitivamente candidatos ao cargo de Primeiro-Ministro. Este é um facto que queremos enfatizar através da visualização. Deste modo, optou-se por uma visualização estática do tempo que repete os políticos e partidos no governo desde 1976. Utilizou-se uma representação cronológica vertical, em que as datas mais recentes de situam mais acima e as datas mais antigas abaixo.

A história das eleições legislativas portuguesas é contada em primeiro plano, sendo possível explorar a representação por dois caminhos. Um dos caminhos segue a exploração da personalidade do político – os seus feitos, a sua reputação, a sua popularidade – através da visualização das notícias publicadas pelos *media*. O outro caminho, explora a alteração de diversos fatores socioeconómicos em Portugal face ao governo existente em cada mandato. Definida a narrativa, o próximo passo foca-se na forma dos dados. Para isso precisamos de um conceito consistente e forte, que passe uma mensagem eficaz, capaz de captar a atenção dos leitores e de levantar questões.

4.3.2. Metáfora Visual

Uma boa metáfora visual tem o poder de comunicar uma ideia através de uma imagem, permitindo uma fácil descodificação da mensagem e tornando-se o alicerce do projeto criativo, direcionando-o para um resultado consistente. Essa ideia é frequentemente descrita com a frase “uma imagem vale mais que mil palavras”.

Este projeto explorou mais profundamente duas metáforas visuais, acabando por ser escolhida aquela que mais facilmente transmitia uma mensagem explícita aos leitores. A primeira metáfora explorada baseava-se na ideia do governo como uma “Erva Daninha” que se sobrepõe às necessidades do povo. A abordagem desta metáfora

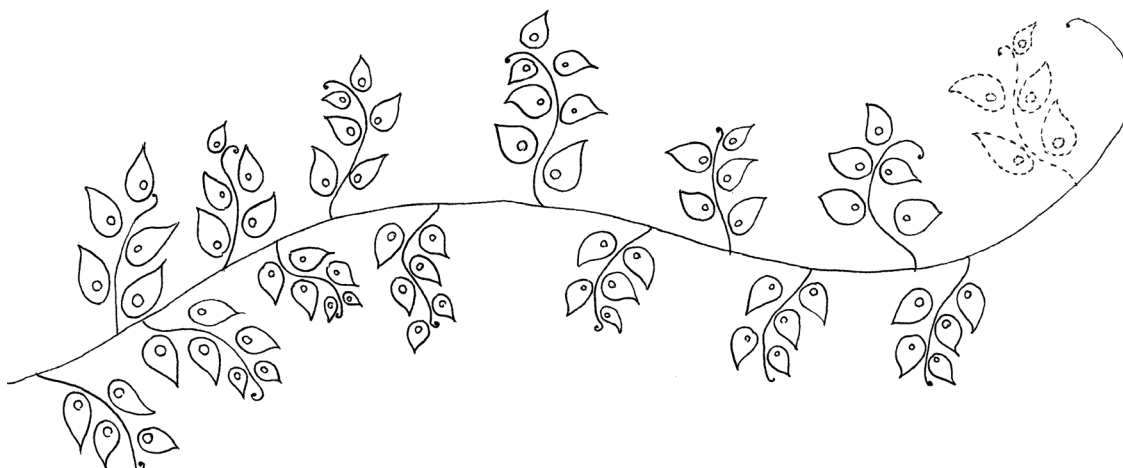


Figura 4.3.2.1 Exploração visual da metáfora biológica "Erva Daninha".

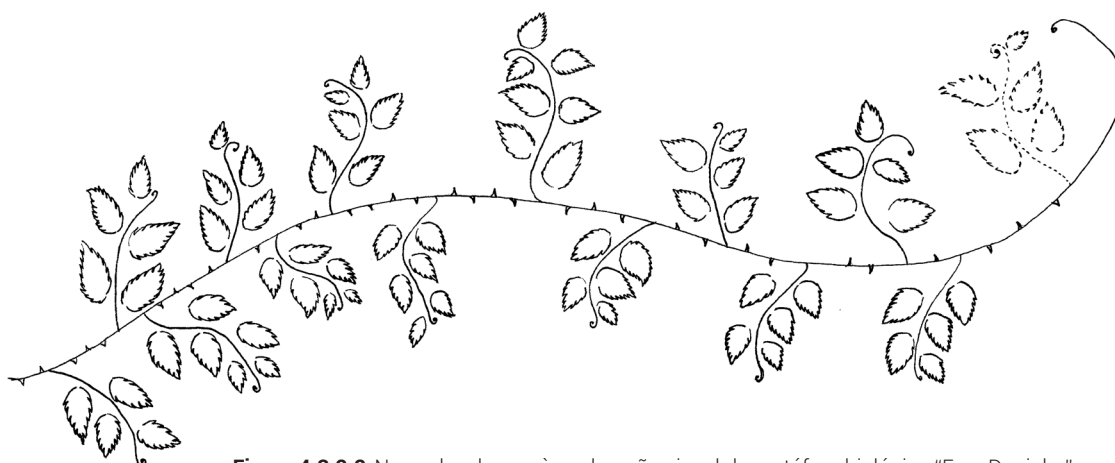


Figura 4.3.2.2 Nova abordagem à exploração visual da metáfora biológica "Erva Daninha" com o objetivo de parecer uma silva.

seguia um caminho mais biológico transformando os políticos – candidatos ao governo – em folhas, as eleições legislativas em caules secundários e a linha temporal no caule principal de uma planta. As folhas seriam representadas com um tamanho proporcional à quantidade de votos, e com a cor representativa do partido, como é possível observar na figura 4.3.2.1. No entanto, esta primeira abordagem não transmitia a ideia negativa conotada às ervas daninhas, pelo que foi alterada a forma das folhas, para umas mais serrilhadas e adicionados alguns espinhos para que se parecesse com uma silva – figura 4.3.2.2. Mesmo assim, a mensagem nem sempre era clara, pelo que se decidiu procurar e explorar uma metáfora mais forte.

A recolha de algumas expressões populares como: "Comem todos do mesmo tacho.", "Eles querem é poleiro.", "São uns ladrões." sugeriam mensagens fortes e de fácil perceção para uma metáfora visual.

Ainda outras citações de filósofos, políticos, entre outros, foram recolhidas:

Na guerra a pessoa só pode ser morta uma vez, mas na política diversas vezes" (Winston Churchill)

"Encontrou-se, em boa política, o segredo de fazer morrer de fome aqueles que, cultivando a terra, fazem viver os outros." (Voltaire)

"Pode-se fazer política com o coração, mas só se pode governar com a cabeça." (António de Oliveira Salazar)

"A política não é uma ciência exacta, mas uma arte." (Otto Bismarck)

"A diferença entre a moral e a política está no facto de que, para a moral, o homem é um fim, enquanto que para a política é um meio. A moral, portanto, nunca pode ser política, e a política que for moral deixa de ser política." (Pio Baroja)

"A política é uma praga tal que eu aconselho todos a não se meterem nela." (Thomas Jefferson)

"É pena que todos aqueles que se dizem democratas, na prática não respeitem o jogo democrático e as posições partidárias diferentes das próprias." (Francisco Sá Carneiro)

"A política é uma guerra sem derramamento de sangue, e a guerra uma política com derramamento de sangue." (Mao Tse-Tung)

"Os partidos são um mal necessário nos países livres." (Alexis Tocqueville)

"Sem política não se organiza uma sociedade. O problema é que a sociedade está nas mãos de políticos profissionais." (José Saramago)

"A diplomacia é uma partida de xadrez em que os povos levam xeque-mate." (Karl Kraus)

"Há quem tema o resultado do voto. Mas temer o resultado das eleições é duvidar do Povo Português." (Francisco Sá Carneiro)

Decidiu-se, no entanto, que as expressões populares têm maior probabilidade e capacidade de transmitir uma mensagem a todo o tipo de público, do que qualquer outra citação. Portanto, para os resultados das eleições legislativas optou-se pela utilização da metáfora baseada na expressão popular "Eles querem é poleiro". Resultando visualmente na representação dos políticos através de pássaros pousados em poleiros que surgem nas datas das eleições legislativas. São representados apenas os líderes políticos cujos partidos obtiveram representação no governo – os líderes e partidos que ficaram no "poleiro" –, sendo que a cor é utilizada para distinguir o partido, a direção do pássaro para distinguir os partidos da esquerda e da direita e o tamanho, bem como a ordem para distinguir o resultado eleitoral – o maior foi o mais votado. Os poleiros e os pássaros surgem, também, distribuídos pela direita e pela esquerda da linha temporal – que se tornou no tronco que suporta os poleiros –, de forma intervalada. A primeira eleição legislativa é representada do lado direito, a segunda no lado esquerdo e assim sucessivamente, por questões de aproveitamento do espaço e equilíbrio da composição visual – como se poderá observar em §5.

4.4. Cenários de utilização

Para que seja possível compreender melhor as funcionalidades desta visualização de informação, foi elaborado um conjunto de hipotéticos cenários de utilização. Estes cenários têm como principal objetivo identificar e compreender os requisitos desta visualização face às necessidades do público-alvo, aplicando os dados disponíveis.

1 - Um jovem eleitor pretende conhecer melhor os candidatos às eleições para formar uma opinião política sobre em quem deve votar nas próximas eleições legislativas. Através da visualização de informação, este nosso eleitor, poderá perceber quais os políticos, e respetivos partidos, que têm ocupado a posição de Primeiro-Ministro, e concretizar o seu objetivo, navegando na representação das notícias de um político selecionado. Aqui, ele obterá uma ideia geral das notícias que têm sido publicadas sobre um dado político, como também poderá aceder à notícia integral.

2 - Um jornalista pretende escrever um artigo sobre um determinado acontecimento referente a um determinado político. Utilizando a visualização como uma ferramenta de ajuda ao seu trabalho, o jornalista poderá explorar a história do político pretendido e perceber, através de uma palavra-chave, se já foram publicadas situações semelhantes noutros anos. Ou poderá apurar o porquê e qual a palavra-chave que mais se adequa a determinado político, consoante a quantidade de vezes em que está presente nas notícias do político em questão.

3 - Um politólogo pretende estudar as alterações em vários fatores socioeconómicos, referentes ao país, consoante os partidos no governo ao longo dos anos. Ao utilizar a visualização de informação, o politólogo poderá ver representadas estas alterações socioeconómicas e poderá complementar o seu estudo, selecionando um político e pesquisando um fator socioeconómico, como por exemplo: natalidade. Com esta pesquisa ele terá acesso às notícias em que esta palavra existe e que poderão ser relativas a incentivos ou decisões tomadas relativamente à natalidade no país.

4 - Um utilizador procura informação sobre as eleições legislativas portuguesas na *web* e depara-se com a visualização de informação “No poleiro desde 1976”. Apercebe-se que a representação central da visualização se foca nos resultados eleitorais das legislativas portuguesas desde 1976 até à data atual e rapidamente consegue perceber os resultados pelos tamanhos dos objetos representados, bem como identificar os partidos que se distinguem pela variação de cores. Numa atitude explorativa, irá reparar que o controlo temporal lhe permitirá aceder à visualização dos fatores socioeconómicos e que, deslocando o controlo no tempo, os fatores irão alterar-se, permitindo-lhe perceber as condições socioeconómicas do país, consoante os governos existentes. Poderá também perceber a duração temporal das consequências de certas medidas aprovadas pelo governo atual ou pelos governos anteriores. Ao selecionar um político irá perceber, também, que a visualização lhe permite navegar pelas memórias dos *media* – notícias publicadas – de cada político e ainda comparar o nível de popularidade de dois políticos em relação ao tempo. Nas notícias representadas pode ainda procurar quais as que contêm palavras que lhe interessem, como por exemplo, saber qual o período temporal em que a palavra “submarino” apareceu mais vezes nas notícias relacionadas com Durão Barroso.

Desenvolvimento Prático

Este capítulo tem como objetivo elucidar o leitor sobre o progresso do projeto ao longo do seu desenvolvimento. Este capítulo abrange a experimentação e definição do design da visualização, relata a interação possibilitada ao utilizador, esclarece as tecnologias utilizadas na implementação do projeto e descreve os problemas encontrados durante a fase de implementação, bem como as soluções aplicadas.

5

5.1. Design

Neste projeto foi necessário pensar e esboçar três visualizações de dados para os três temas diferentes: uma visualização dos dados relativos às eleições legislativas desde 1976, uma visualização de dados relativos a notícias por entidade e uma visualização dos dados relativos aos fatores socioeconómicos. Estas três visualizações individuais passaram por um processo de transformação, sempre procurando otimizar e facilitar a leitura dos dados, até chegarem à sua forma final.

Neste subcapítulo iremos descrever esse processo de transformação, explicando as necessidades dos dados e as soluções encontradas para a resolução de problemas emergentes.

5.1.1 Visualização das Eleições Legislativas

Como referido no capítulo anterior, a forma gráfica desta visualização foi especialmente guiada e despoletada pela criação de um conceito visual, capaz de transmitir uma mensagem forte. Sem este conceito, a visualização estava destinada a uma forma gráfica óbvia e muito utilizada em projetos, que visam mostrar visualmente resultados eleitorais. A primeira abordagem à forma visual destes dados eleitorais – que poderá ser observada no anexo A – é um exemplo demonstrativo de como é importante um conceito para evitar as representações mais óbvias, visto que a tendência é representar dados através de formas já conhecidas e que, portanto, garantem uma interpretação eficiente.

O primeiro conceito centrou-se, como explicado anteriormente, numa metáfora biológica que compara o governo a uma erva daninha – mais especificamente uma silva. Nesta abordagem, a linha cronológica era representada pelo caule principal da planta e deste nasciam ramificações, consoante os anos em que aconteceram eleições – figura 5.1.1.

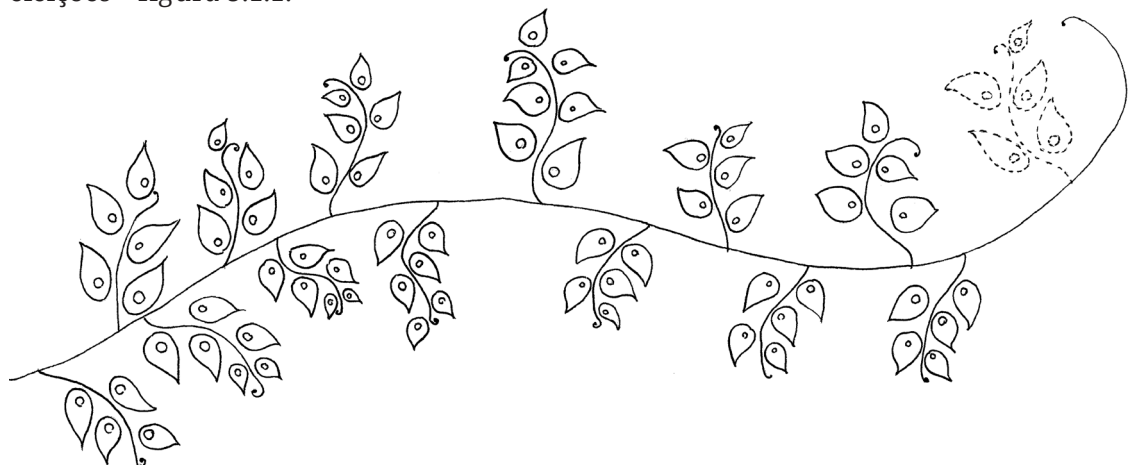


Figura 5.1.1 Exploração visual da metáfora biológica "Erva Daninha", na qual as folhas representam políticos e cada caule com folhas, uma eleição.

Já nesta fase se pensava em representar apenas os líderes partidários, face à falta de dados. Neste caso, as folhas representariam os líderes partidários e o seu tamanho seria proporcional à votação conseguida em cada eleição. A localização das folhas é, também, dada pelas votações seguindo uma ordem decrescente – figura 5.1.2.

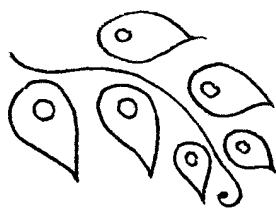


Figura 5.1.2 Pormenor da visualização de resultados eleitorais. As folhas são desenhadas da maior para a menor, intercalando o lado.

Para dar a informação relativa ao partido, as folhas teriam uma gota de orvalho que seria colorida.

Ao representar o ano atual – 2015 – os dados das sondagens seriam utilizados para criar as folhas correspondentes aos líderes partidários, que se estão a candidatar às eleições do ano corrente. Para estabelecer a diferença entre resultados de eleições passadas e resultados previstos, através de sondagens, o traço do desenho seria descontínuo – figura 5.1.3.



Figura 5.1.3 Pormenor da visualização de resultados das sondagens para o ano corrente.

Mais tarde, quando foi percebido que a visualização não estava a passar o conceito pretendido de forma clara, a forma gráfica sofreu alguns ajustes para criar uma sensação mais selvagem – figura 5.1.4.

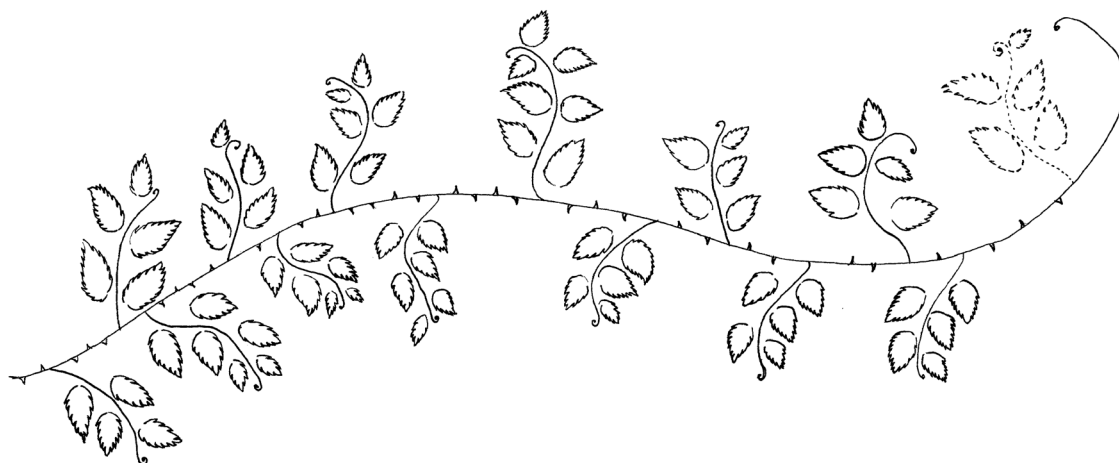


Figura 5.1.4 Nova abordagem à exploração visual da metáfora biológica "Erva Daninha", com o objetivo de parecer uma silva.

O segundo conceito apropriou-se da expressão popular “Eles querem é poleiro” para passar uma mensagem forte e facilmente entendida por qualquer pessoa.

A criação da forma gráfica deste segundo conceito teve algumas influências da forma gráfica projetada anteriormente, contudo muito teve de ser adaptado. Uma das vantagens da forma gráfica do conceito anterior era a forma orgânica e visualmente equilibrada da linha cronológica – dada pelo caule da planta. Com o novo conceito, este equilíbrio perde-se um pouco, uma vez que os novos elementos – pássaros – têm obrigatoriamente uma posição vertical.

Várias tentativas foram feitas para preservar a forma orgânica e o equilíbrio na página, sendo que se concluiu que se tratavam de matérias diferentes e, portanto, teriam de assumir formas diferentes consoante as necessidades do conceito. A nova linha cronológica passa, assim, a assumir uma orientação vertical, como se pode verificar na figura 5.1.5, diminuindo as dimensões, relativamente à linha cronológica anterior, que se estendia pela horizontalidade do ecrã. Os anos representados foram abreviados, possibilitando o aumento do tamanho da letra e, facilitando a visibilidade. Caso contrário, o tamanho das letras teria de diminuir, para possibilitar a escrita da totalidade do ano dentro da mesma área.

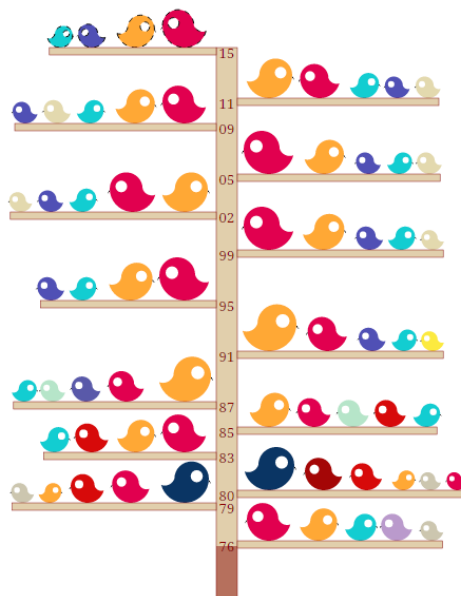


Figura 5.1.5 Nova abordagem baseada na metáfora “Eles querem é poleiro.” Solução de distribuição dos pássaros no ecrã.

À semelhança da forma gráfica anterior, da linha cronológica nascem poleiros horizontais, conforme os anos em que aconteceram eleições legislativas. Neles assentam os políticos, agora representados por pássaros. Os pássaros adquirem a cor do partido que lideraram e o tamanho varia consoante a votação conseguida em cada eleição. No poleiro, os pássaros são ordenados de forma decrescente para que, em caso de existirem tamanhos similares entre os pássaros, seja mais fácil verificar qual foi o que obteve maior quantidade de votos – figura 5.1.6.



Figura 5.1.6 Pormenor dos resultados de uma eleição: ordem decrescente.

Os pássaros assumem ainda uma direção – direita ou esquerda – que está relacionada com o partido. Se o político é líder de um partido da direita, então estará virado para a direita.

Paleta Partidária

Como explicado anteriormente, as cores utilizadas nos pássaros têm como objetivo representar o partido. Assim, foi feita alguma pesquisa das cores associadas aos partidos. Com esta pesquisa notou-se que são vários os partidos que fazem uso da cor vermelha, sendo difícil decidir qual será mais associado à cor em questão e qual terá de adotar outra cor.

Uma vez que estas cores são vistas em conjunto, houve alguma preocupação para não saturar o ecrã de cores diferentes. Assim, por vezes, houve necessidade de alterar ligeiramente a cor para criar harmonia entre todas elas.



Figura 5.1.7 Paleta de cores partidárias.

Como podemos observar na figura 5.1.7, um tom de rosa é atribuído ao PS, o laranja ao PPD/PSD, o azul-turquesa ao CDS-PP. Estas são as cores mais conhecidas e que as pessoas associam mais rapidamente. Todas as outras cores resultam da tentativa de buscar a cor aos seus logotipos, quando estes existem. Para aqueles cuja cor é desconhecida é atribuída uma nova cor, diferente daquelas já utilizadas.

Podemos observar que ao PCP-PEV foi atribuído um azul forte e ao CDU um azul muito parecido. Isto acontece devido ao CDU passar a sua nomenclatura para PCP-PEV, em 1989, sendo o mesmo partido, apenas com nomes diferentes. À semelhança do que acontece com estes dois partidos, o mesmo se passa com o BE e o UDP, que apresentam uma cor neutra muito semelhante.

5.1.2 Visualização das Notícias

Os primeiros esboços para a visualização das notícias determinaram uma linha cronológica circular, com o objetivo de minimizar o espaço necessário. Decidiu-se representar a quantidade de notícias através de círculos, cujos raios variam consoante a quantidade de notícias.

Na figura 5.1.8 podemos observar a primeira solução para a representação das notícias, que atribui um mês do ano a cada circunferência e distribui os dias do mês de forma radial.

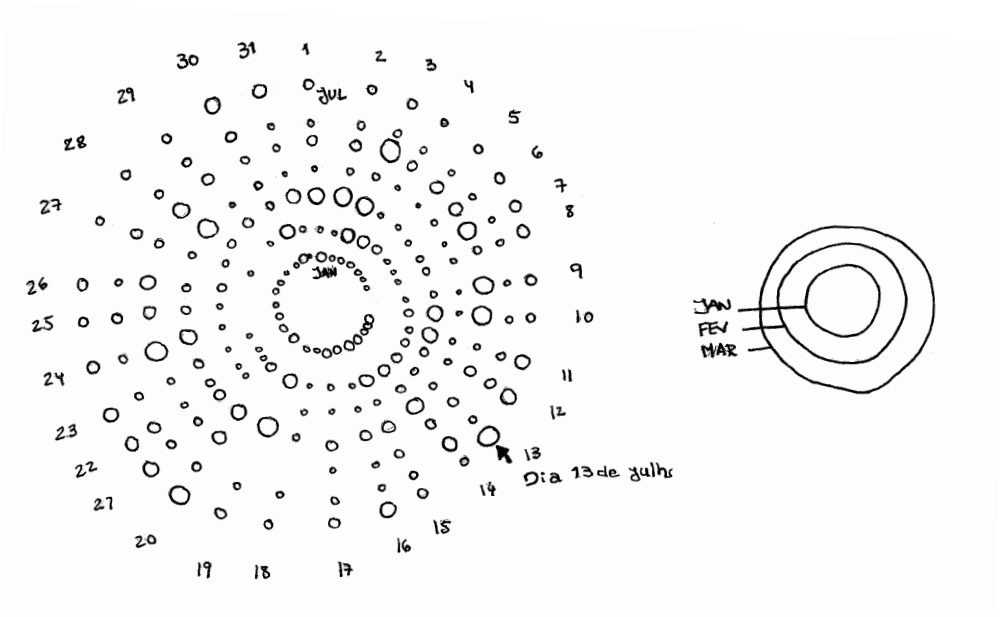


Figura 5.1.8 Primeiro esboço para a representação de notícias referentes a um ano.

Na figura 5.1.9 temos uma solução semelhante à anterior. São representadas cinco semanas, através de circunferências com raios diferentes e os doze meses de um ano de forma radial.

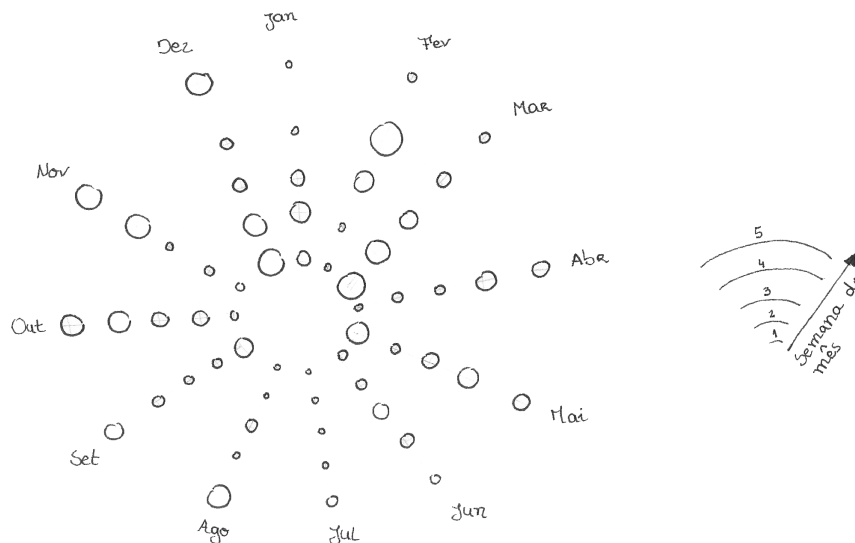


Figura 5.1.9 Segunda solução para a representação de notícias referentes a um ano.

Ambas as abordagens não produziam o efeito esperado, sendo que tinham muitas desvantagens a nível de leitura e interpretação. Para ler e interpretar a quantidade de notícias do ano representado, era necessário quebrar a leitura a cada mês. Isto acontece porque a leitura mensal, neste gráfico, faz-se de dentro para fora.

Para facilitar a interpretação das notícias era necessária uma leitura contínua e sem quebras, pelo que se concluiu que a melhor maneira de o conseguir seria utilizando uma linha cronológica elíptica.

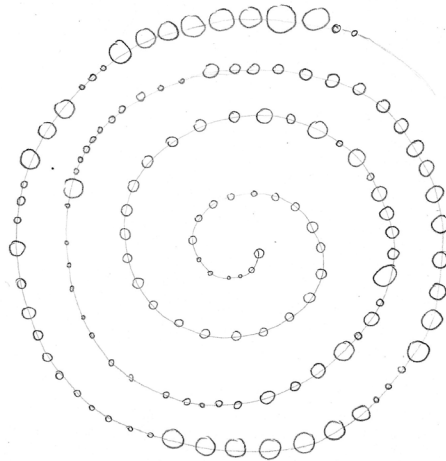


Figura 5.1.10 Terceira solução para a representação de notícias referentes a um ano.

Decidiu-se, para esta cronologia, que se iria representar da totalidade dos anos – desde 1987 até à data atual – evitando a necessidade de um controlo temporal para definir os anos representados. Esta decisão deveu-se também ao facto de já se ter pensado, anteriormente, numa funcionalidade que permitisse ao utilizador alterar a granularidade das notícias, ou seja, controlar o agrupamento de notícias.

Controlo da granularidade

O controlo da granularidade das notícias permite ao utilizador controlar o agrupamento de notícias por períodos trimestrais, semestrais ou anuais. A decisão sobre quais os períodos temporais que são possíveis de visualizar, foi fortemente influenciada pelo resultado visual e pela decisão de representar a totalidade dos anos – cerca de três décadas.

Os pedidos feitos à base de dados do SAPO permitiam, até esta fase do projeto, a recolha da quantidade de notícias publicadas por dias, meses ou anos. Com estes dados seria possível representar vinte e oito círculos – 28 anos –, trezentos e trinta e seis círculos – 336 meses – ou dez mil duzentos e vinte círculos de notícias – 10 220 dias. Como se pode perceber, o aumento do número de círculos com estas granularidades seria exponencial, já para não dizer que a representação de dez mil círculos, para além de se tornar pesada e lenta de processar, obrigaria à utilização de uma escala com tamanhos muito pequenos para permitir a representação de todos os círculos no ecrã. Assim, houve a necessidade de identificar períodos temporais, cujo número de círculos resultantes fosse razoável. O resultado desta pesquisa foram as granularidades atuais – anual, semestral e trimestral – que permite representar cerca de vinte e oito círculos, cinquenta e seis círculos e cento e doze círculos.

Após a definição dos períodos possíveis de visualizar, seria então necessário pensar na forma gráfica do controlo da granularidade. A primeira abordagem, possível de observar na figura 5.1.11, utilizava a forma gráfica de um slider.

Contudo, a sua forma passava despercebida, pelo que se optou por uma nova solução que utiliza três botões diferentes, cada um para a sua granularidade – figura 5.1.12.

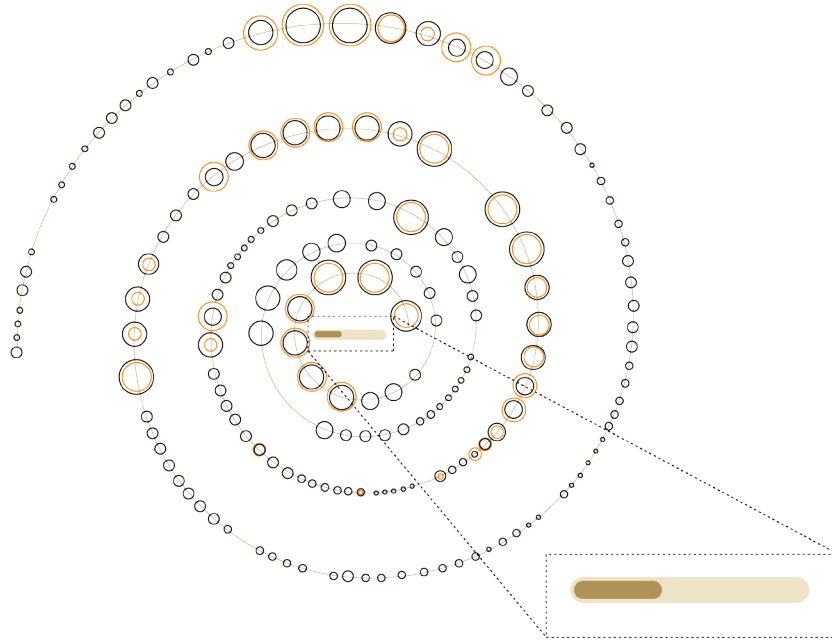


Figura 5.1.11 Primeira solução para o controlo da granularidade.

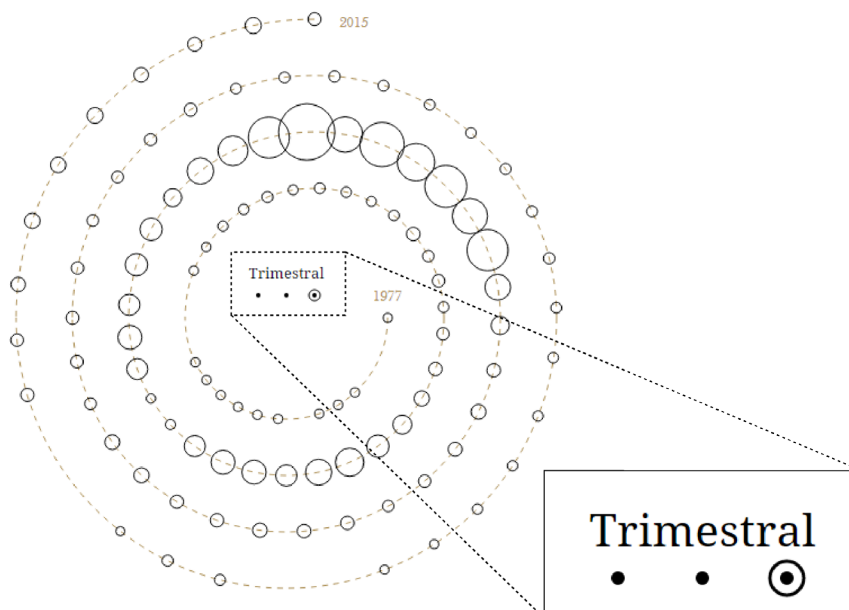


Figura 5.1.12 Solução final para o controlo da granularidade.

Comparação de políticos

Uma das funcionalidades interessantes desta visualização é a possibilidade de comparar a popularidade de dois políticos. Para isso, era necessário encontrar uma solução que possibilitasse a representação das quantidades de notícias, referentes aos dois políticos em simultâneo.

Chegou-se facilmente à solução apresentada na figura 5.1.13, que resolveu eficientemente o problema, permitindo a representação das notícias individuais em simultâneo e fazendo o uso da cor do contorno para distinguir o político que mencionam.

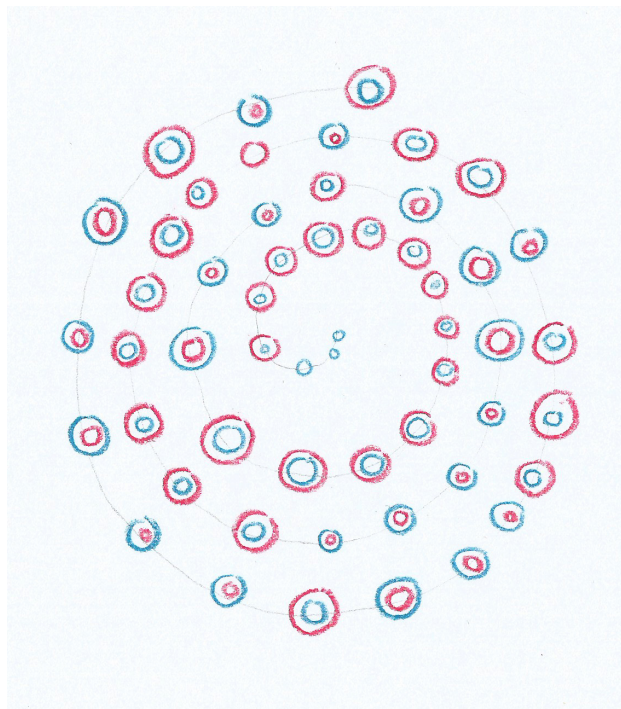


Figura 5.1.13 Solução para comparar notícias de dois políticos.

Legendas

Como aconselhado por Tufte³, devem ser utilizadas legendas para combater a distorção e ambiguidade das visualizações de dados. Neste caso, era importante e necessário para a visualização das notícias – principalmente quando existe a comparação das notícias de dois políticos – a existência de legendas para informar a cor destinada a cada político.

Uma primeira solução e que podemos observar na figura 5.1.14 baseia-se numa abordagem comum. No entanto, foi encontrada outra solução, fazendo uso dos elementos já existentes e permitindo salvar espaço na página. Esta nova solução transforma o político selecionado em legenda, atribuindo-lhe um contorno colorido, tal como nas suas notícias – figura 5.1.15.

³ Tufte, 1995

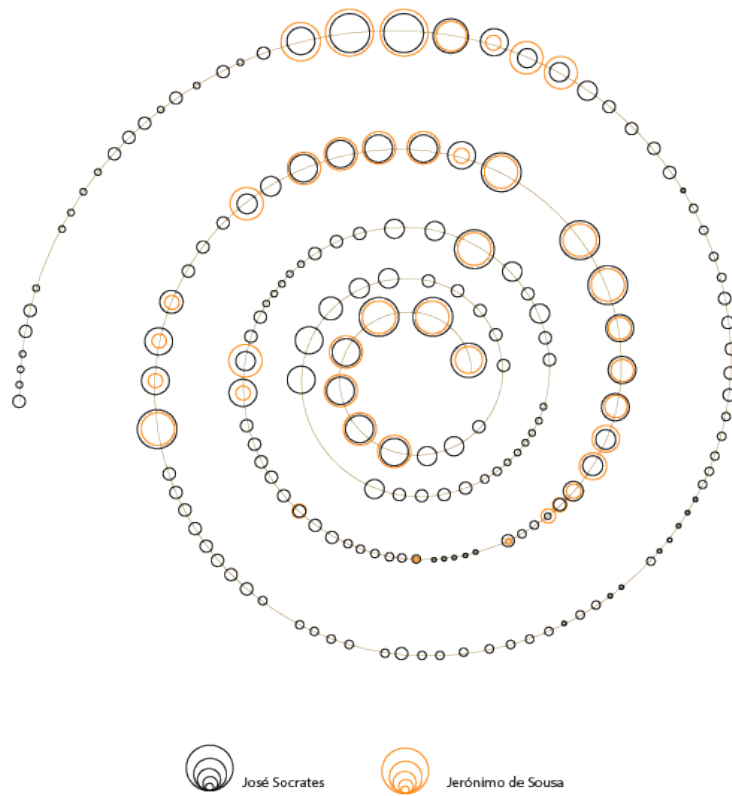


Figura 5.1.14 Esboço da primeira solução para legendar as notícias correspondentes a cada político selecionado.

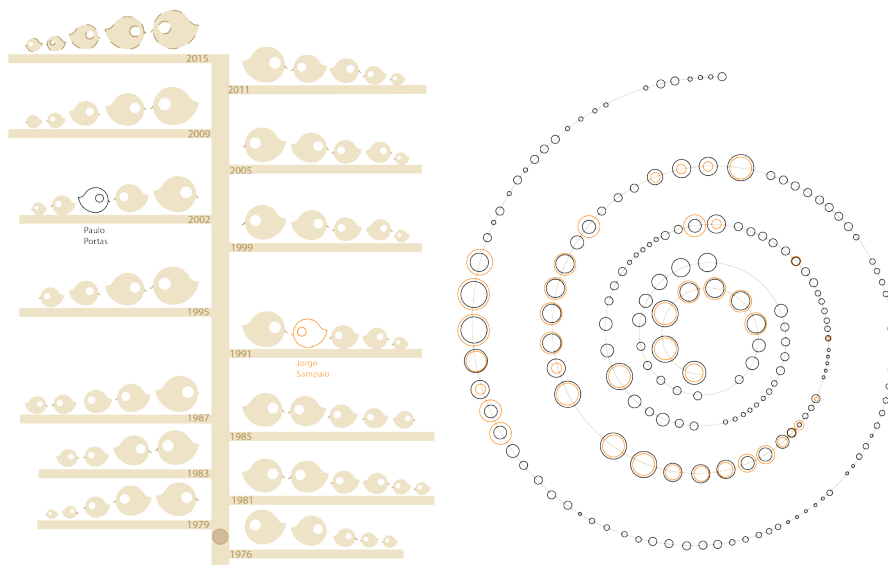


Figura 5.1.15 Solução final para legendar as notícias correspondentes a cada político, fazendo uso de elementos já existentes na página.

Detalhes das Notícias

Através da seleção de um dos círculos apresentados na visualização das notícias, também é possibilitado ao utilizador a inspeção das notícias publicadas pelos *media*, no intervalo temporal definido pelo círculo. Para a visualização destas notícias – título, data de publicação e fonte – relativas a um único político, uma lista de Título ordenados por data, com a respectiva fonte da publicação, foi a solução – figura 5.1.16.

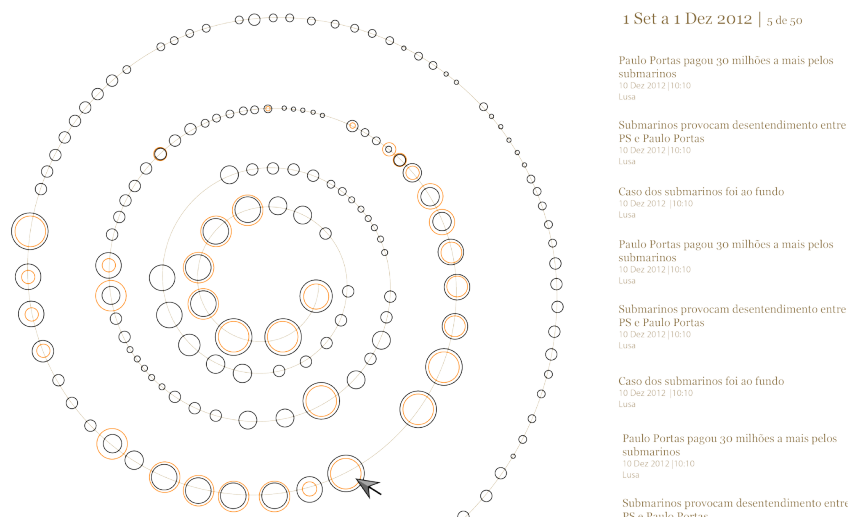


Figura 5.1.16 Experimentação digital de soluções para a visualização de detalhes de artigos noticiosos, referente a um período selecionado.

Para a visualização de notícias de dois políticos, a primeira abordagem – figura 5.1.17 – separava as notícias de cada um, colocando uma legenda com o nome do político no início da lista de notícias referentes a esse político.



Figura 5.1.17 Primeira solução para a visualização de detalhes de artigos noticiosos sobre dois políticos, referente a um período selecionado.

A solução anterior dá preferência a um dos dois políticos selecionados, colocando todas as suas notícias em primeiro lugar. Isto poderia ser um problema quando o utilizador pretendesse explorar as notícias do outro, pois o utilizador não sabe, à partida, que a lista das notícias do outro político é apresentada após a lista do primeiro. Uma segunda abordagem solucionou melhor o problema, fazendo uso do mesmo método utilizado para a legenda das cores dos círculos. Ou seja, utilizou a cor no título, para distinguir as notícias dos dois políticos. Esta solução possibilitou ainda uma ordem cronológica decrescente das notícias – figura 5.1.18.

1 Set a 1 Dez 2012 | 5 de 50



Paulo Portas pagou 30 milhões a mais pelos submarinos
10 Dez 2012, 10:10 | Lusa

Submarinos provocam desentendimento entre PS e Paulo Portas
10 Dez 2012, 10:10 | Lusa

Caso dos submarinos foi ao fundo
10 Dez 2012, 10:10 | Lusa

Paulo Portas pagou 30 milhões a mais pelos submarinos
10 Dez 2012, 10:10 | Lusa

Submarinos provocam desentendimento entre PS e Paulo Portas
10 Dez 2012, 10:10 | Lusa

Caso dos submarinos foi ao fundo
10 Dez 2012, 10:10 | Lusa

Figura 5.1.18 Solução final para a visualização de detalhes de artigos noticiosos de dois políticos, referente a um período selecionado.

Pesquisar notícias específicas

A pesquisa de notícias que mencionam uma certa palavra é uma das funcionalidades mais interessantes para quem está interessado em construir uma opinião sobre um determinado político, tal como verificar em que períodos essa palavra ou esse assunto foi mencionado. Assim, era necessário pensar numa solução para representar estas notícias específicas em simultâneo com as notícias individuais de cada político.

O esboço da primeira solução encontrada, deixava de parte as quantidades destas notícias-chave e assinalava apenas os círculos em que as mesmas existiam através do preenchimento do círculo – figura 5.1.19.

Com a solução anterior percebeu-se que muita informação estava a ser ignorada. Por exemplo, quando dois políticos estão selecionados, interessa saber se a palavra pesquisada é mencionada nas notícias dos dois ou apenas de um e interessa também saber qual a quantidade destas notícias-chave em relação à totalidade de notícias de cada um.

A segunda solução esboçada já respondia melhor a este problema, permitindo representar as quantidades de cada um. Contudo, como podemos observar na figura 5.1.20, esta solução fazia o uso da mesma linguagem gráfica, utilizada para representar a totalidade das notícias, pelo que poderia ser confusa a sua interpretação.

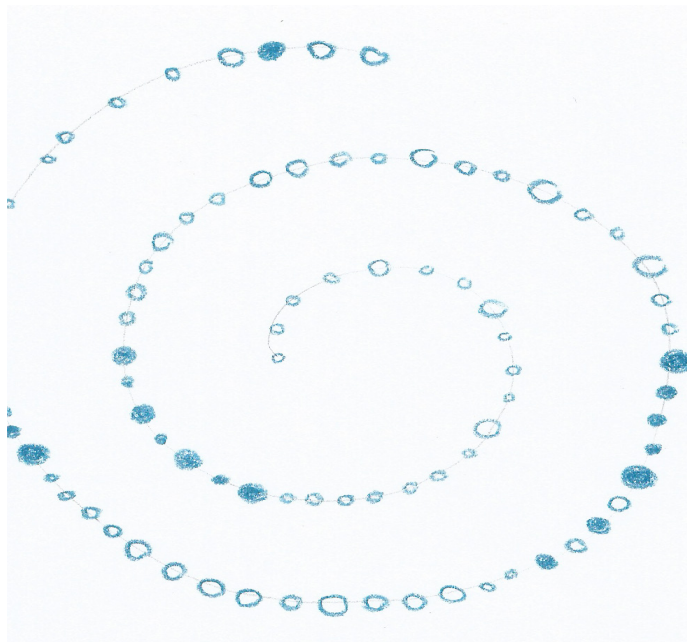


Figura 5.1.19 Esboço da primeira solução para a visualização da existência e frequência de notícias específicas.

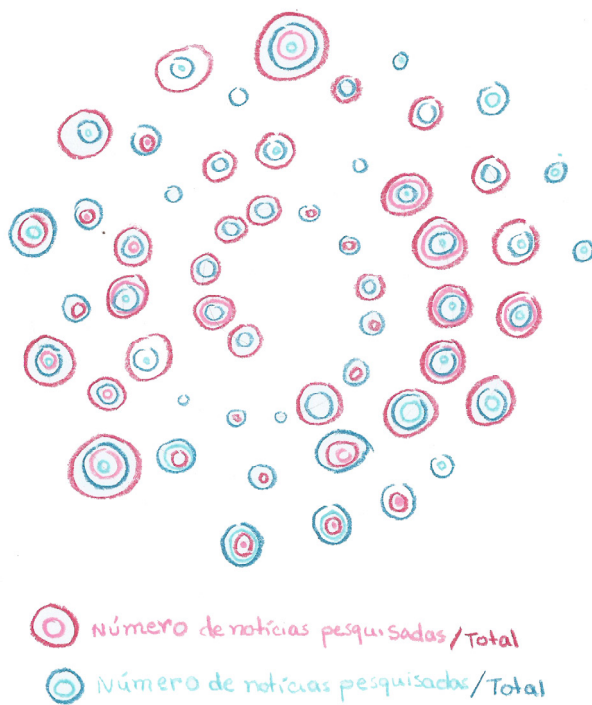


Figura 5.1.20 Esboço da segunda solução para a visualização da existência e frequência de notícias específicas.

Desta forma, pensou-se ainda noutra solução, que continua a ter as suas desvantagens, mas que torna a leitura um pouco mais fácil quando esta é individualizada por períodos. Como podemos observar na figura 5.1.21 a solução encontrada consistiu na aplicação de gráficos tarte – *pie charts* – à totalidade das notícias. Assim, uma leitura individualizada de cada trimestre, semestre ou ano, permite perceber qual a fatia do total das notícias que corresponde às notícias que mencionam a palavra pesquisada pelo utilizador.

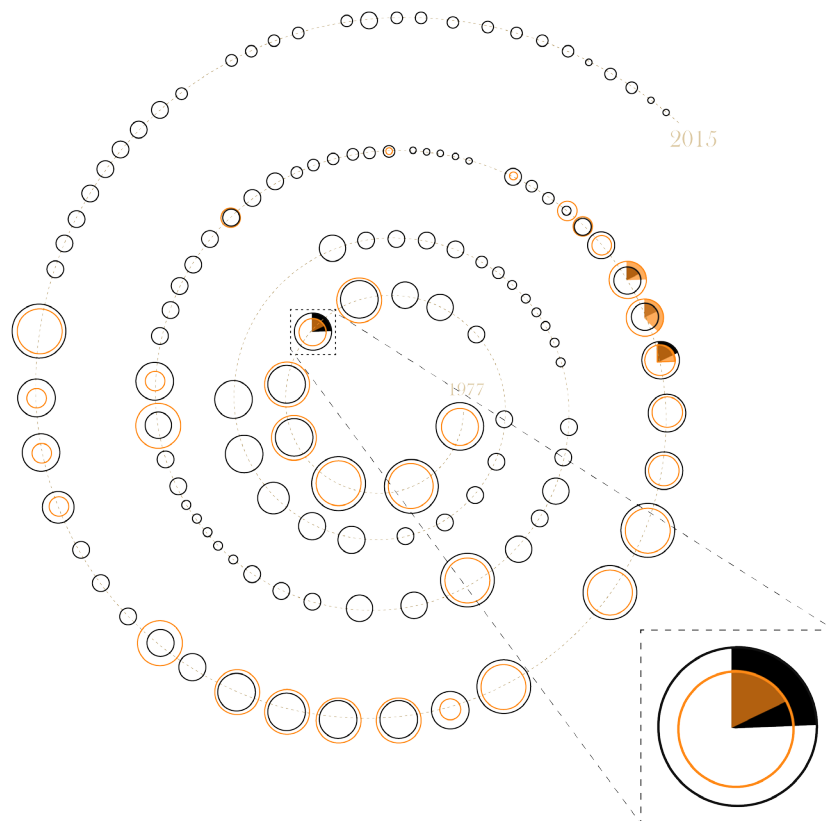


Figura 5.1.21 Esboço da terceira solução para a visualização da existência e frequência de notícias específicas.

A desvantagem desta solução, quando é feita uma leitura geral dos trimestres, semestre ou anos, é que a leitura visual pode enganar, pois um círculo pequeno que contenha uma fatia grande poderá representar uma quantidade menor que uma fatia pequena num círculo grande. Consequentemente, pretende-se através da interação – como será explicado em §5.2 – elucidar o utilizador das quantias reais.

Pesquisa de notícias por *Tags*

Esta funcionalidade foi pensada para aqueles cujos conhecimentos de política são nulos ou poucos. Tal como acontece com a pesquisa de notícias por palavras, esta funcionalidade resulta na apresentação de gráficos tarte que representam a quantidade, em relação ao total, em que as notícias estão ‘tagadas’ com a palavra escolhida. A única diferença para o utilizador, é que em vez de pesquisar uma palavra,

são sugeridas várias, por ordem decrescente de utilização.

Uma vez que a representação das quantidades é igual à da pesquisa das palavras, para esta funcionalidade apenas foi necessário pensar na forma de apresentação das palavras e no botão.

A figura 5.1.2.22 mostra a primeira solução para a exibição das palavras, ou seja, das *tags*. Esta solução enumera as palavras numa lista separada por político, quando estão selecionados dois políticos.

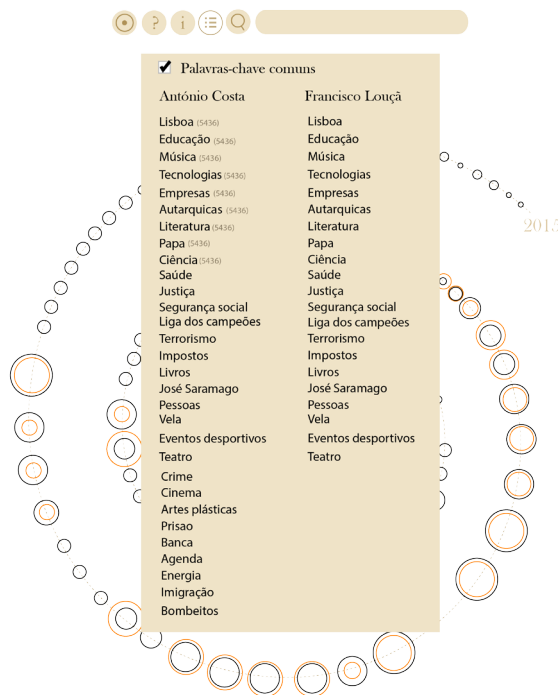


Figura 5.1.22 Experimentação digital da visualização das *Tags* para o político selecionado.

Mais tarde, pensou-se noutra solução, graficamente mais coerente com as decisões tomadas até esta fase do projeto: utilizar a cor atribuída a cada um para distinguir as respetivas *tags*. Foi, ainda, alterada a forma como as palavras se apresentavam, deixando de ser lista e aparecendo seguidas, o que permitiu poupar espaço – figura 5.1.23.



Figura 5.1.23 Experimentação digital da visualização das *Tags* para o político selecionado.

5.1.3 Visualização dos Fatores Socioeconômicos

Como não podia deixar de ser, a forma gráfica desta visualização foi muito influenciada pelos dados escolhidos. Só após a recolha e análise dos dados foi possível compreender o problema que os dados colocavam e, assim, começar a procurar uma solução viável.

O maior desafio destes dados era a amplitude dos valores entre cada tema, tendo sido experimentadas diferentes escalas para representar os valores. O problema verificado é que, independentemente de se utilizar uma escala uniforme ou uma escala de intervalos, a amplitude era sempre demasiado grande para se conseguir representar as variações de forma legível. A alternativa, que permitiu a solução deste problema, consistiu no mapeamento dos valores de cada tema, num intervalo entre o mínimo e o máximo da circunferência, ignorando os valores absolutos – demasiado altos para se ter uma noção de quantidade – e centrando a atenção na variação – subida e descida dos valores. A desvantagem desta abordagem é a impossibilidade de comparação dos valores de dois fatores, pois quando estes têm o mesmo valor na representação isso não significa que partilham do mesmo valor absoluto na realidade.

A primeira solução esboçada para esta visualização usava a forma gráfica dos ‘gráficos-rosa’ para representar os valores de cada tema num único ano. Como pode ser verificado na figura 5.1.23, todos os temas seriam representados em simultâneo, sendo necessária a utilização da cor para distinguir cada tema. A variação consoante os anos seria dada através da manipulação do ano representado, que faria os valores de cada tema atualizarem.



Figura 5.1.24 Primeiro esboço para a visualização do valor anual de cada fator socioeconómico.

Com esta visualização, o objetivo pretendido residia especialmente em poder analisar a variação dos valores de cada tema e poder associar os anos em que há uma queda ou uma subida, com o mandato político.

O que se verificou com a solução anterior foi que para que o utilizador tivesse uma noção da variação de um tema específico, teria de se focar nesse tema, ignorando os outros e memorizar as subidas e

descidas à medida que controlava os anos representados. Esta solução era, portanto, pouco prática para responder ao objetivo pretendido.

A segunda solução veio responder às necessidades deste objetivo, permitindo uma fácil leitura da variação dos temas ao longo dos anos. Esta abordagem permitiu, assim, a representação da variação dos temas em vários anos em simultâneo. Como podemos observar na figura 5.1.25, a utilização da cor continua a estar presente para a distinção de temas, tal como na solução anterior.

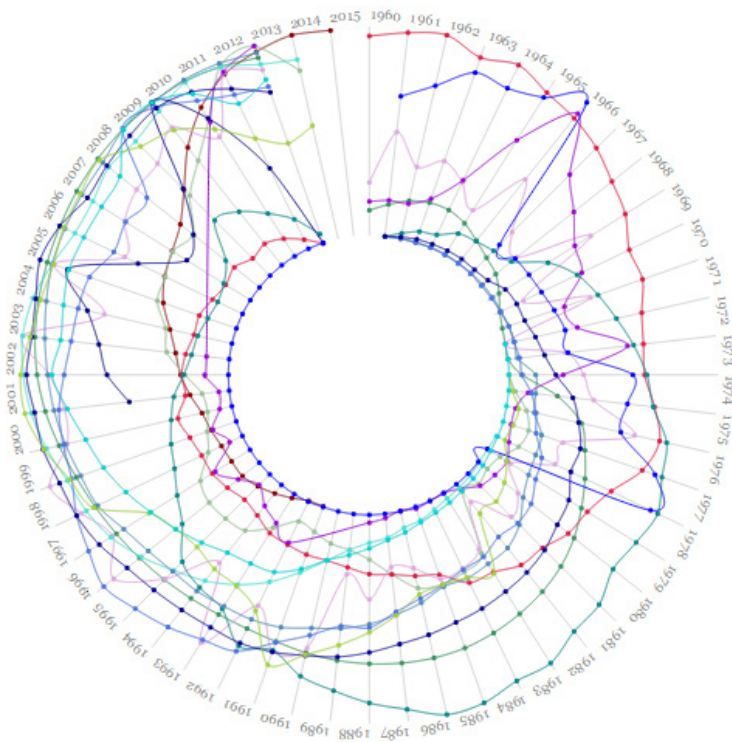


Figura 5.1.25 Solução final para representar a variação anual dos fatores socioeconômicos.

5.2. Interação

A interação com o sistema tem cada vez mais importância para as visualizações de dados face à crescente recolha e registo de informação. É, portanto, essencial para as visualizações de grandes quantidades de dados, pela possibilidade de mostrar conjuntos de dados menores através de camadas. Esta possibilidade é especialmente importante para que o utilizador não seja subterrado em informação logo na primeira impressão, o que dificultaria a leitura dos dados, e é mais provável de causar distrações ou desânimo. Uma interação intuitiva é, então, um dos fatores chave para o sucesso da visualização, pois determina o interesse do utilizador pela visualização, resultando numa atitude de desistência ou de exploração.

A visualização deste projeto representa três assuntos distintos que estão, no entanto, interligados. Como informação principal, a visualização das eleições legislativas estabelece uma relação com os outros dois assuntos – as notícias publicadas sobre os políticos e a variação dos fatores sociais. Desta forma, esta visualização principal influencia e controla as outras duas.

Ao passar o rato pelas eleições – o poleiro que serve de suporte aos pássaros – é apresentada informação relativa aos votos por partido, consoante a eleição em foco. Clicar neste mesmo elemento permitirá manter a informação no ecrã, para comparação com outras eleições.

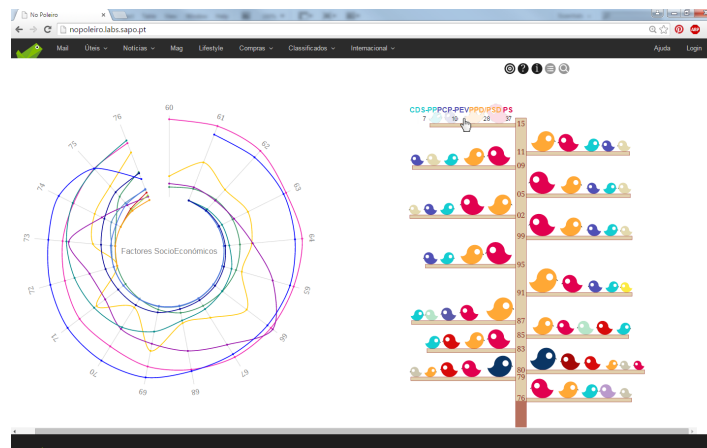


Figura 5.2.1 Interação com o poleiro, mostrando a informação relativa aos resultados das sondagens para a eleição legislativa de 2015.

Ao passar o rato por um político há uma alteração da opacidade do pássaro e é apresentada o seu nome. A existência da opacidade permite aos utilizadores perceberem que podem selecionar o político.

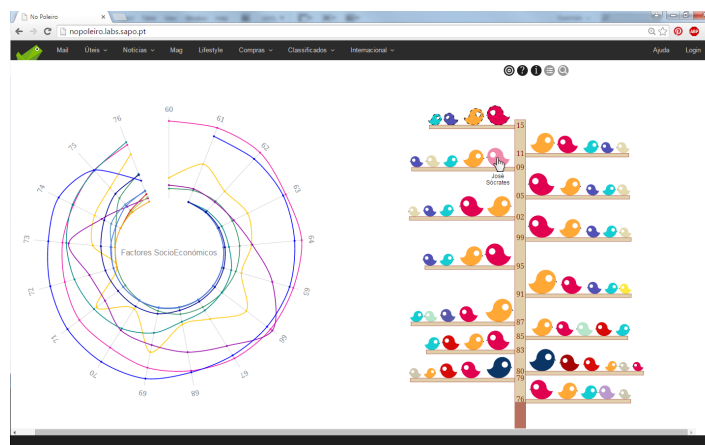


Figura 5.2.2 Interação com o político José Sócrates.

Quando um político é selecionado, a visualização desloca-se na horizontal revelando uma nova visualização de notícias relativas ao político selecionado. O pássaro selecionado adquire, assim, a cor utilizada para representar as notícias que o mencionam, a fim de servir como legenda da visualização das notícias. Esta legenda é

necessária para possibilitar a comparação, quando selecionado mais um político, que adquirirá uma nova cor distinta da utilizada para o primeiro político selecionado.

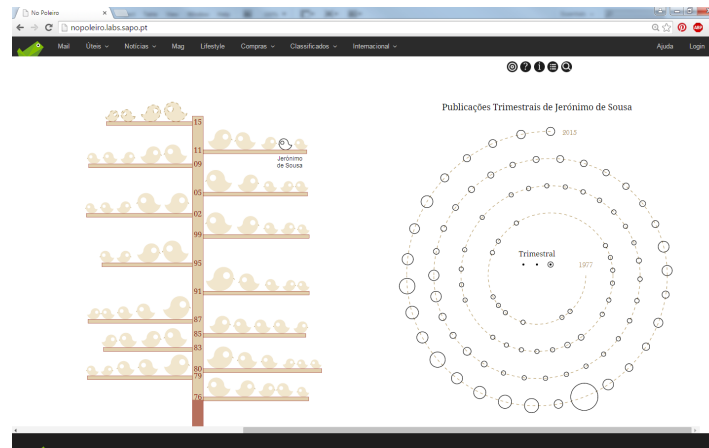


Figura 5.2.3 Resultado da deslocação horizontal e alteração da cor do político selecionado, bem como, de todos os outros.

Na visualização das notícias, quando o rato passa sobre um círculo, o mesmo adquire uma opacidade de 10% e é apresentada a quantidade de notícia correspondente ao círculo. Ao mesmo tempo, é adicionada uma legenda ao centro da visualização com o intervalo temporal correspondente ao círculo em foco.

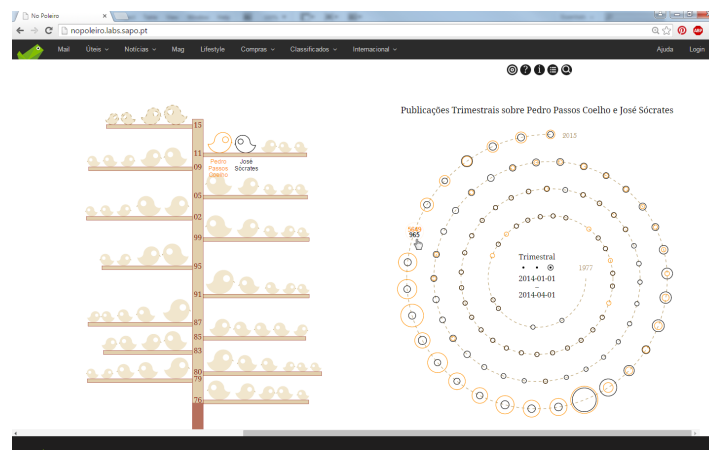


Figura 5.2.4 O mouseover sobre um dos círculos de notícias revela o número de notícias para cada um dos políticos selecionados. Ao centro é dado o período temporal representado pelo círculo em que está o rato.

Quando um dos círculos apresentados é selecionado, a visualização volta a deslocar-se para a esquerda, dando espaço à direita para a exibição das notícias publicadas durante o período de tempo selecionado. No caso de estarem dois políticos selecionados, a cor é utilizada para diferenciar as informações que correspondem a cada um. Como tal, quando dois políticos estão selecionados e o rato se encontra sobre um dos círculos, é dada a informação dos dois políticos ao mesmo tempo. Quando o rato seleciona um círculo, as notícias sobre os dois políticos são apresentadas.

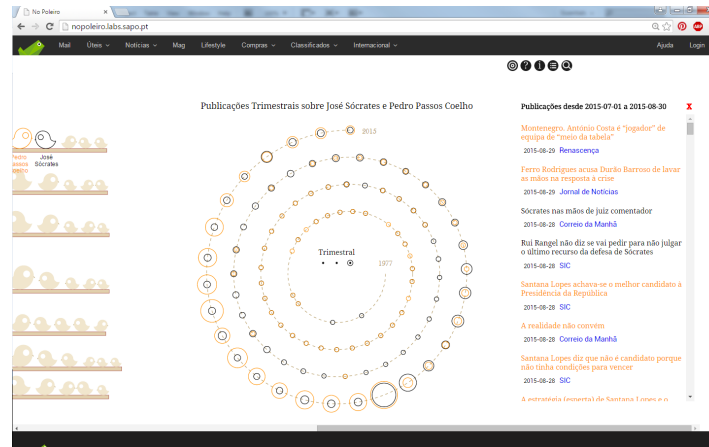


Figura 5.2.5 Ao clicar sobre um dos círculos representados existe uma deslocação para a direita e são mostrados detalhes sobre os artigos noticiosos publicado no intervalo de tempo escolhido.

O utilizador tem também a possibilidade de procurar por notícias que mencionem uma palavra específica. Após a seleção de um político, o botão de pesquisa – anteriormente inativo – altera o seu estado para ativo, permitindo ao utilizador introduzir uma palavra e obter resultados sobre a quantidade de notícias que mencionam a palavra pesquisada.

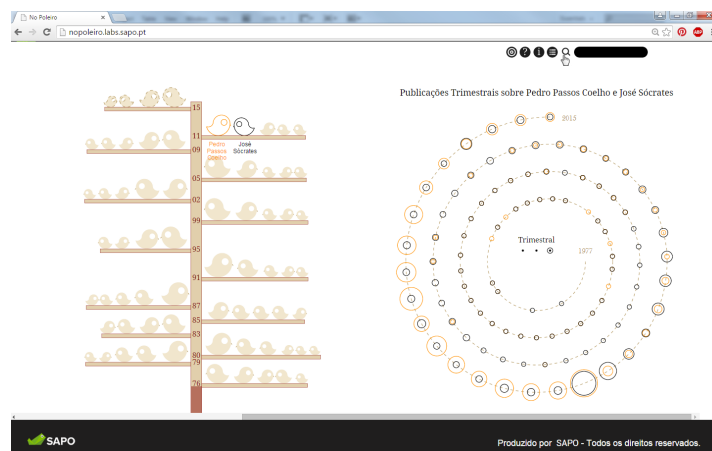


Figura 5.2.6 Ao seleccionar o botão de pesquisa é revelada uma caixa de texto, onde o utilizador pode introduzir a palavra que pretende procurar.

Esta quantidade é dada em forma de gráficos tarte, representando a quantidade de notícias que mencionam a palavra em relação à totalidade de notícias nesse período. Ao passar o rato por estes gráficos tarte, é exibido o período temporal e a quantidade de notícias correspondentes – em percentagem – que mencionam a palavra pesquisada. Quando um destes gráficos é selecionado, são mostradas as notícias que lhe correspondem – à semelhança do que acontece com os círculos, ou seja, a totalidade de notícias.

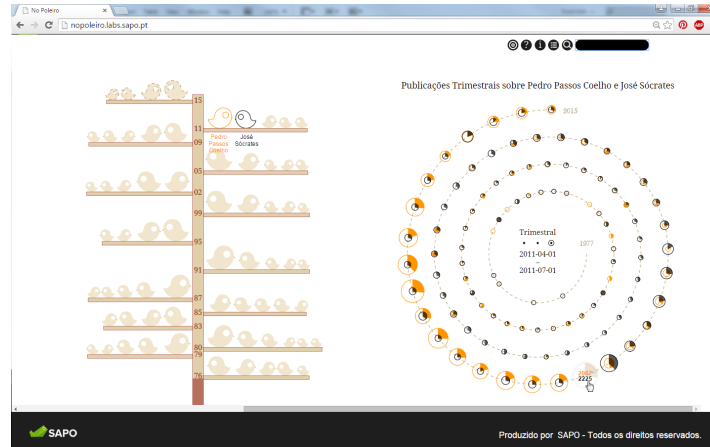


Figura 5.2.7 O mouseover sobre um gráfico tarte revela os valores que correspondem às notícias pesquisadas para os políticos selecionados.

A pensar nos utilizadores que não têm grandes conhecimentos políticos foi criada uma funcionalidade que lhes permite explorar as notícias mais facilmente. Através do botão da lista de tags, os utilizadores podem visualizar as palavras mais utilizadas nos assuntos das notícias, relativas ao político selecionado, escolhendo a palavra que lhes interessa e obtendo um resultado semelhante à pesquisa de palavras. Ou seja, são apresentados gráficos tarte que mostram a quantidade de notícias que foram etiquetadas com a palavra selecionada pelo utilizador.

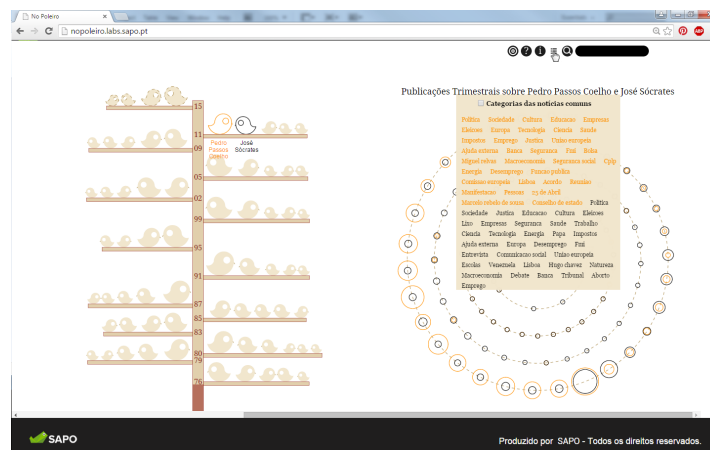


Figura 5.2.8 Apresentação das tags mais utilizadas para cada político selecionado. Para obter esta apresentação das tags individuais, é necessário retirar o 'certo' da checkbox, localizada antes do título "Categorias das notícias comuns".

Voltando à visualização eleitoral, ao passar o rato pelo poleiro vertical – o que define a cronologia – o ponteiro do rato assume uma forma que sugere ao utilizador a alteração da dimensão vertical de um elemento. Este elemento é a barra mais escura, que altera a sua dimensão consoante a posição vertical do rato. Esta dimensão pode ser alterada através do arrastamento da barra ou do clique – implementação necessária para permitir um bom funcionamento nos dispositivos móveis.

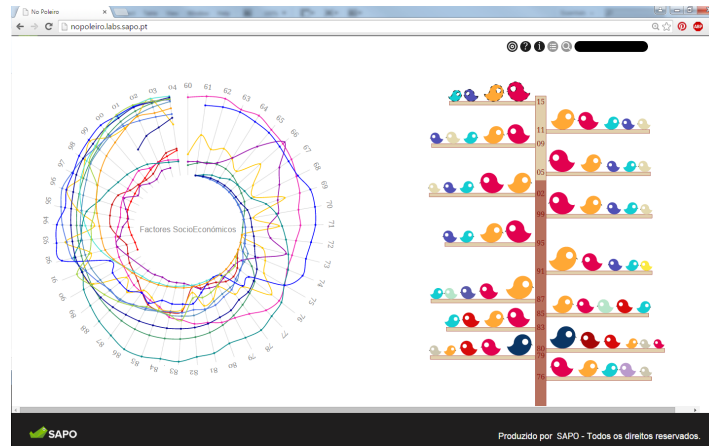


Figura 5.2.9 A manipulação do tamanho da barra mais escura, na *timeline* da representação das eleições legislativas, permite a alteração do período temporal representado na visualização dos fatores socioeconómicos.

O objetivo da alteração da dimensão da barra é a manipulação direta do período temporal, representado na visualização dos fatores socioeconómicos. Esta relação é criada, pois os fatores socioeconómicos portugueses variam ao longo tempo, resultado de medidas de incentivo ou desincentivo por parte dos governos regentes. Nesta visualização, podemos observar uma série de linhas coloridas. Ao passar o rato por cima dessas linhas é exibida uma legenda, no centro da visualização, indicando o tema em foco.

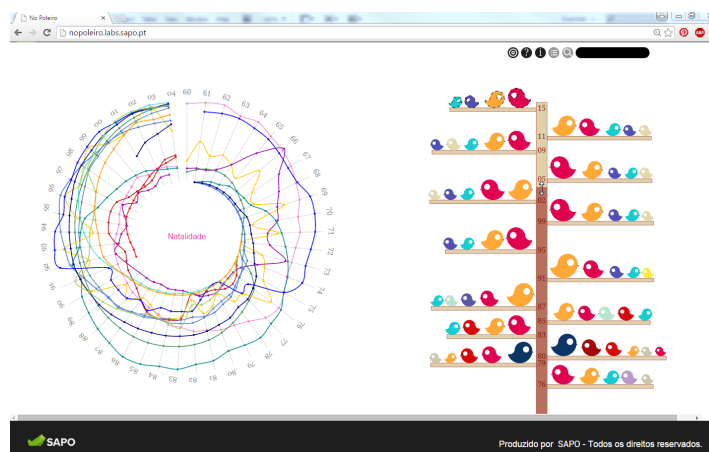


Figura 5.2.10 O cursor do rato, quando este passa sobre a barra temporal, muda de forma para indicar a possibilidade de redimensionar a mesma.

Quando uma das linhas é selecionada, todas as outras adquirem uma opacidade de 10%, permitindo uma análise individualizada do tema selecionado. Esta individualização proporciona uma leitura mais fácil, evitando a dificuldade de acompanhar a linha quando outras estão sobrepostas.

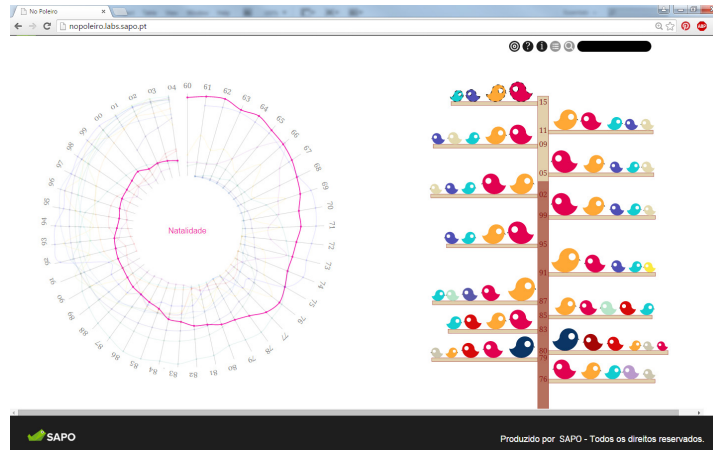


Figura 5.2.10 A individualização de um fator socioeconómico propicia, ao utilizador, uma percepção mais rápida da variação referente ao fator selecionado. Nesta imagem observamos que a Natalidade atinge o mínimo em 2013, o último ano para o qual, existem dados.

5.3. Tecnologias

O objetivo do SAPO, com este projeto, é a criação de um *website*, que permita aos seus utilizadores tirar partido de visualizações interativas de dados, a fim de criar novas opiniões e novos conhecimentos. Assim sendo, este projeto utiliza as populares linguagens da *web*, como o *HTML5*, *CSS* e *Javascript*.

Apesar de ter sido, primeiramente, planeada a utilização de *Processing.js* – uma linguagem já conhecida e utilizada em trabalhos anteriores – para a implementação da forma gráfica dos dados, foi pedido e aconselhado por parte do SAPO a utilização de *D3.js*. Houve, então, a necessidade de aprender a utilizar esta biblioteca. O *D3.js* é uma biblioteca *JavaScript* para a manipulação de documentos baseados em dados. É especialmente utilizada para a criação de visualizações interativas de dados, usando *HTML*, *SVG* e *CSS*. É, também, uma ferramenta *open source* com uma boa documentação na *web*.

Para o levantamento de dados relativos às notícias, foi utilizada uma outra biblioteca do *Javascript* – o *jQuery* – que permitiu a construção de *queries* que retornam os dados pretendidos em formato *JSON*. Foi, ainda, utilizada uma outra biblioteca designada por *Papa Parse* para a extração e armazenamento dos dados dos ficheiros *CSV* em *arrays*.

5.4. Implementação

Neste subcapítulo serão explicados pormenores técnicos, mais importantes, da implementação. Como em todos os projetos, durante esta fase prática surgiram problemas, os quais foram precisos resolver. É neste subcapítulo que estes problemas são descritos e apresentadas as abordagens para encontrar uma boa solução.

comando `display: none`, é inicialmente definido que estas *divs* não devem ser exibidas. Assim, o que se vê inicialmente no *browser* são, apenas, as secções e respetivos elementos, da *div* 'intro' – figura 5.4.3.

A *div* 'container2' é seccionada em quatro *divs* de ajuda, uma *div* 'barra' – onde estarão os botões de reload, ajuda, informação, lista de *tags* e pesquisa – e uma *div* 'visualizacao' – figura 5.4.4.

Por fim, a *div* 'sobre' é seccionada em duas *divs* 'fecha' destinadas a botões, seis Títulos e cinco *divs* 'text' que contém parágrafos. Esta *div* é adicionada à página, imediatamente após a secção 'container2'.

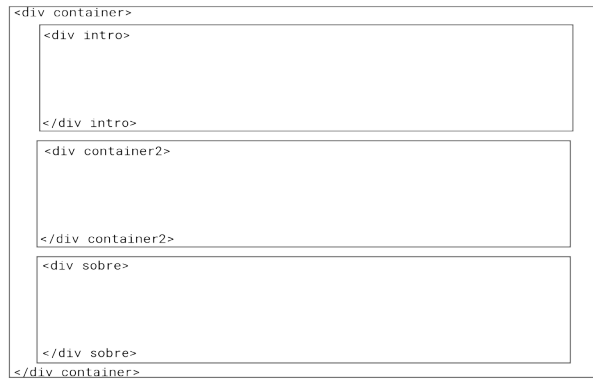


Figura 5.4.2 Estrutura da *div* 'container'.

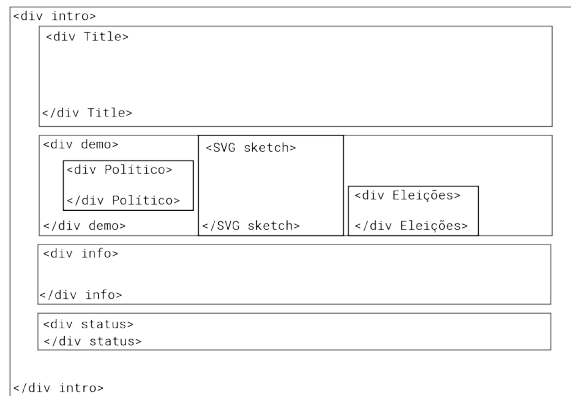


Figura 5.4.3 Estrutura da *div* 'intro'.

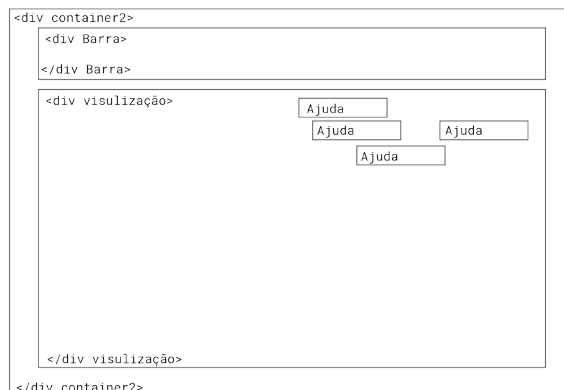


Figura 5.4.4 Estrutura da *div* 'container2'.

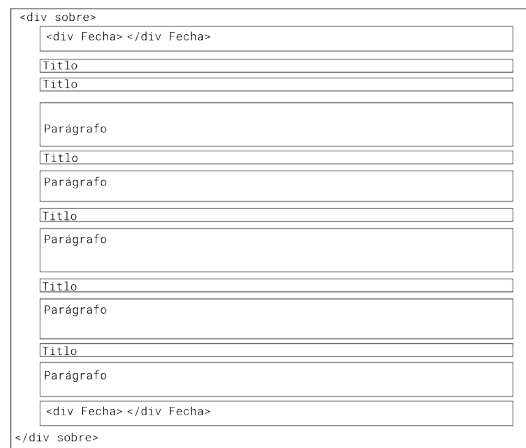


Figura 5.4.5 Estrutura da div 'sobre'.

Extração e processamento de dados

Em simultâneo à apresentação da página de introdução, assim que esta é iniciada, é iniciado também o processo de extração e armazenamento dos dados. Este processo é definido através da linguagem de programação *JavaScript*. Em resumo, podemos afirmar que o *HTML* estrutura a página *web*, o *CSS* dá forma à página, ou seja, define o *layout* e o *JavaScript*, controla o comportamento da página, logo dos seus elementos.

Este projeto utiliza dados com duas localizações diferentes. Enquanto que os dados armazenados em ficheiros CSV encontram-se localizados no servidor do site, os dados das notícias encontram-se numa base de dados do SAPO, logo, a forma de extrair e armazenar os dados será diferente.

A extração dos dados que se encontram em ficheiros CSV foi feita através da utilização de uma biblioteca *JavaScript* designada *Papa Parse*. Esta biblioteca ajuda a converter os dados dos ficheiros CSV em objetos JSON, sendo criado um objeto JSON por cada linha do ficheiro CSV.

O comportamento da função `Papa.parse()` desta biblioteca é o seguinte:

- recebe o caminho do ficheiro
- define os dados da primeira linha do ficheiro CSV dado como nome dos campos do objeto
- define o tipo de variável (caso seja número ou booleano não será extraído como string)
- quando completa:
 - adiciona os resultados ao array `csvFiles` na posição certa
 - se os dados foram todos carregados:
 - exibe o texto "Explorar"
 - exibe o botão explorar
 - esconde o texto "A carregar os dados..."

Esta função é chamada para cada ficheiro CSV cujo caminho está guardado no *array* 'csvFiles'. Este *array* tem, inicialmente duas posições: a primeira guarda o tema do ficheiro CSV, a segunda o caminho do ficheiro CSV. Posteriormente, à medida que o *JavaScript* vai progredindo são acrescentadas mais três posições, resultando num total de cinco. A terceira posição vai armazenar todos os objetos

desse tema criados pelo *Papa Parse*, a quarta e a quinta armazenam o valor mínimo e máximo do tema.

Ao contrário do que acontece com os dados dos ficheiros CSV, que são extraídos e armazenados em *arrays* logo no início, quando o browser exhibe a introdução, os dados das notícias são extraídos, apenas quando são necessários para a construção da visualização das notícias.

A visualização de notícias precisa de dados relativos à quantidade de publicações sobre um político, quantidade de notícias que mencionam uma determinada palavra, quantidade de notícias com uma determinada *tag*, detalhes de artigos noticiosos num dado período temporal, detalhes de artigos noticiosos num dado período temporal que mencionam uma determinada palavra, detalhes de artigos noticiosos com uma determinada *tag* e uma lista de *tags* utilizadas em notícias de um dado político.

Para cada uma das temáticas dos dados enumerados acima existe um pedido diferente, pelo que é criada uma função para cada pedido. Estes pedidos têm, no entanto, uma estrutura similar: uma base comum a todos, seguida da especificação do nome do método (*url*) consoante os dados que se quer pedir e, a especificação de parâmetros pretendidos.

Exemplo: 'http://services.sapo.pt/InformationRetrieval/Epicentro/GetTrendline?timeFrame=TRIMESTER&nodesIDs=43872&beginDate=1987-01-01&endDate=2015-08-21'

pedido_completo = base + método + parâmetros.

Base: http://services.sapo.pt/InformationRetrieval/Epicentro/

Método: GetTrendline?

Parâmetros: 'TimeFrame=TRIMESTER&nodesIDs=43872&beginDate=1987-01-01&endDate=2015-08-21'

Após a execução da função que faz os pedidos dos dados através da função *getJSON* do *jQuery* – outra livreria do *JavaScript* –, a informação recolhida é processada por novas funções que extraem apenas os dados necessários e os guardam em novos *arrays*.

Posição dos políticos

No final do carregamento de todos os dados dos ficheiros CSV é executada uma função que vai criar um novo *array* global 'ClassBirds', cujo objetivo é guardar dados apenas necessários para a criação de novos pássaros. Assim, para cada eleição, serão recolhidos os partidos com lugar no governo e para cada um deles será criado um objeto que armazenará o ano (da eleição), o número de lugares conseguidos, o nome do partido e o nome do líder partidário.

Com este novo *array* temos, praticamente, toda a informação necessária para a criação e desenho dos pássaros, faltando apenas calcular as posições de cada pássaro. Desta forma, logo após o preenchimento do *array* 'ClassBirds' é chamada uma função cujo objetivo é calcular a posição de cada pássaro e chamar a função que o cria. Esta nova função utiliza não só o *array* 'ClassBirds' mas, também, o *array* global 'anos' com quarenta *arrays* que guardam um ano entre 1976 a 2015, a tipologia do ano – eleição ou normal – e, no caso de ser eleição, existe ainda uma outra posição que guarda uma direção – direita ou esquerda.

O cálculo das posições de cada pássaro é processado da seguinte forma:

- percorre o array 'anos':
 - se o ano for ano de eleição:
 - então percorre o array 'ClassBirds':
 - se o ano inspecionado for igual ao ano recolhido do array 'ClassBirds' && a direção for 'direita':
 - chama o pássaro passando-lhe o nome do político, posição x, posição y*, partido, número de lugares conseguidos, o ano e o número do político
 - guarda a posição x no array 'ClassBirds'
 - guarda a posição y no array 'ClassBirds'
 - guarda a direção no array 'ClassBirds'
 - incrementa o eixo do x
 - se o ano inspecionado for igual ao ano recolhido do array 'ClassBirds' && a direção for 'esquerda':
 - chama o pássaro passando-lhe o nome do político, posição x, posição y*, partido, número de lugares conseguidos, o ano e o número do político
 - guarda a posição x no array 'ClassBirds'
 - guarda a posição y no array 'ClassBirds'
 - guarda a direção no array 'ClassBirds'
 - decrementa o eixo do x
 - se o ano não é ano de eleição:
 - eixo x volta ao valor inicial

*Nota: a posição y é calculada através de um valor base – valor inicial – e a subtração de um valor 'encr' que multiplica pelo número de anos já percorridos. O valor da variável 'encr' é calculado antes do cálculo das posições ser iniciado, através da divisão da altura disponível da *div* 'vis' pelo número de anos a representar – trinta e nove anos.

Esqueleto do político

Para ser possível programar o desenho do pássaro foi necessário, primeiro, definir a sua forma. A anatomia de um político, ou seja, de um pássaro será composta por um corpo, um olho e um bico.

Através da figura 5.4.6, pode ser observada a estrutura-base auxiliar ao desenho do pássaro. São, portanto, utilizadas duas circunferências, com tamanhos diferentes, sendo que a dimensão da mais pequena equivale a 40% da dimensão da circunferência maior. É ainda utilizada uma linha auxiliar com um ângulo de 15°, que parte do centro da circunferência maior e a interceta, marcando o início do caminho – *path* – do corpo do pássaro.

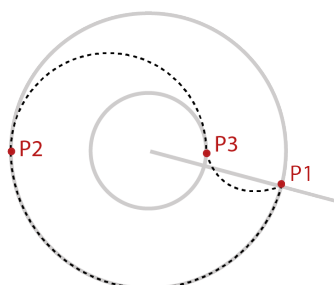


Figura 5.4.6 Estrutura auxiliar ao esqueleto do político.

O caminho para desenhar o corpo do pássaro é, assim, iniciado no Ponto P1, usando arcos elípticos para contornar a circunferência maior até ao ponto P2 – que é definido pelo ângulo de 180° . Contorna depois a circunferência mais pequena até ao ponto P3 – definido pelo ângulo de 360° . O caminho fica completo ao fechar no ponto em que começou – P1. Este caminho é adicionado a um grupo 'birdGroup' que por sua vez é adicionado à *div* 'viz'. A este grupo terão de ser ainda adicionados o olho e o bico para completar o passarinho.

O olho é definido por um círculo que se posiciona no Ponto P4 cuja posição horizontal $-x$ – é dada pela subtração de metade do raio da circunferência maior ao centro do pássaro. A posição vertical $-y$ – é dada pela subtração de um décimo do raio da circunferência ao centro do pássaro.

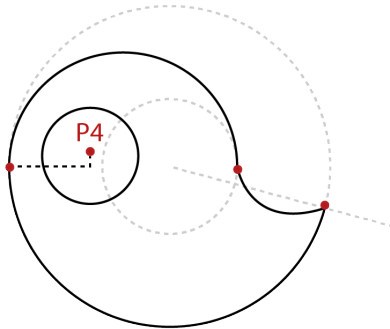


Figura 5.4.5 Cálculo da posição central do olho do pássaro

Finalmente, o bico do pássaro é definido por um caminho que usa arcos elípticos, contornando uma nova circunferência, cujo centro é dado pela subtração da soma do raio da circunferência maior com o raio da nova circunferência, que corresponde a $1/4$ do raio da circunferência maior, ao centro da circunferência maior. Este caminho é iniciado no ponto em que a circunferência completa 18° e acaba aos 90° , não fechando o caminho.

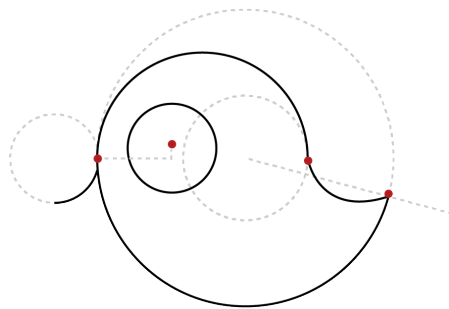


Figura 5.4.6 Cálculo da posição central do bico do pássaro

Dimensão do político

Para definir o raio do círculo é utilizada a quantidade de lugares no governo conseguidos para o partido, do qual o político é filiado. Este valor é passado à função que desenha o pássaro, quando esta é chamada. Contudo, os valores reais, principalmente os valores que correspondem aos partidos menos votados, devem ser proporcionalmente aumentados para que sejam visíveis.

PRIMEIRA ABORDAGEM

A primeira abordagem teve em conta a preocupação de solucionar os valores para os partidos com menos de cinco lugares no partido. Chegou-se à conclusão que um raio de cinco pixéis já era uma dimensão razoável, pelo que se decidiu somar essa quantia – cinco pixéis – ao valor recebido pela função – número de lugares conseguidos pelo partido do político a ser desenhado.

raio do pássaro = valor real + 5;

Contudo, esta solução acentuou outro problema relativo às dimensões. Os pássaros cujos partidos obtiveram maior votação atingiam, agora, dimensões superiores ao espaço disponível entre os poleiros – representação das eleições –, pelo que uma nova abordagem teve de ser pensada.

SEGUNDA ABORDAGEM

Da mesma maneira que os valores menores devem ser aumentados, também os valores correspondentes aos partidos mais votados devem ser proporcionalmente reduzidos para que caibam entre os poleiros. Assim, a nova solução implementava uma escala através de duas condições:

- se o valor real é maior ou igual a 20:
 - então o raio do pássaro = 5 + 35% do valor real;
- se o valor é menos que vinte:
 - então o raio do pássaro = 10 + 35% do valor real;

Após visualização do resultado final, chegou-se à conclusão que esta solução não era a melhor. Acontece que, em casos em que o valor real atingia valores próximos de vinte, as dimensões do político ultrapassavam as dimensões de outros com maior votação, mas, cujos valores, ultrapassavam pouco mais de vinte.

TERCEIRA ABORDAGEM

Apesar da solução anterior ser um fracasso, era já uma solução que se aproximava da solução necessária. Percebeu-se que a soma da base de dez pixéis com 35% do valor real resultava em valores razoáveis à necessidade. Esta solução resolveu o problema dos valores pequenos, sendo que, agora, o pássaro com o raio menor tem sempre um raio de 10, 35 píxeis – resultante da soma da base com 35% de 1 lugar.

raio do pássaro = 10 + 35% do valor real;

Esqueleto do poleiro

Após o cálculo das posições de todos os pássaros, ou seja, da divisão da área de desenho, é chamada a função que desenha o poleiro. O poleiro é definido por uma barra vertical que se estende desde o fundo da *div* ‘vis’ até à base do último pássaro desenhado – no ano 2015, ou seja, localizado o mais acima na página – e catorze barras horizontais que servem de poiso dos passarinhos. Estas barras horizontais partem da barra vertical e acabam assim que termina o desenho do último político nesse ano – figura 5.4.7.



Figura 5.4.7 Dimensões do poleiro horizontal.

Nesta mesma função, em que é desenhado o poleiro, são ainda adicionados os anos das eleições.

Neste seguimento, o desenho em causa é executado em dois passos: primeiro é desenhada a barra vertical e depois as barras verticais, uma a uma, com o ano respetivo. Os elementos são todos adicionados a um grupo, identificado pela class “Poleiro”, que por sua vez é adicionado à *div* ‘vis’.

A lógica da programação, para o desenho da barra vertical, é a seguinte:

- adiciona ao Poleiro um retângulo em que:
 - x corresponde à subtração de 50 pixéis à posição x do primeiro pássaro
 - y corresponde à subtração de 5 pixéis à posição y do último pássaro desenhado
 - o comprimento corresponde a 20 pixéis
 - a altura corresponde à subtração da posição y do último pássaro desenhado ao valor do fim da *div* ‘vis’
 - a cor de preenchimento é um bege
 - a cor do bordo é um castanho
 - e, a grossura do bordo é de 0.5.

A lógica da programação, para o desenho das barras horizontais e dos anos é um pouco diferente, visto que apenas queremos desenhar barras e anos que correspondem aos anos em que existiram eleições legislativas. Assim, é necessário recorrer a condições para filtrar os anos, assemelhando-se ao processo do cálculo das posições dos políticos:

- percorre o array ‘anos’
 - cria um novo array ‘birdpol’
 - se o ano for ano de eleição:
 - então percorre o array ‘ClassBirds’
 - se o ano inspecionado for igual ao ano recolhido do array ‘ClassBirds’:
 - adiciona ao array ‘birdpol’ todos os dados do array ‘ClassBirds’ correspondentes à posição inspecionada

O que este pedaço de código faz é colocar num novo *array* todos os dados – ano, nome do líder político, posição x, posição y, lugares, ordem, partido e cor – dos políticos que foram eleitos para o governo no ano em inspeção, ou seja, a ser desenhado.

Logo de seguida começam a ser desenhadas as barras horizontais, por anos e, de acordo com a direção:

- se a ordem do primeiro pássaro do ano em desenho é direita:
 - adiciona ao Poleiro um retângulo em que:
 - x corresponde à subtração de 30 pixéis à posição x do primeiro pássaro do ano em desenho
 - y corresponde à subtração de 5 pixéis à posição y do primeiro pássaro do ano em desenho
 - o comprimento corresponde à subtração da posição x do último pássaro desse ano à posição do primeiro pássaro
 - a altura corresponde a sete pixéis
 - a cor de preenchimento é um bege
 - a cor do bordo é um castanho
 - e, a grossura do bordo é de 0.5.

- se a ordem do primeiro pássaro do ano em desenho é esquerda:
 - adiciona ao Poleiro um retângulo em que:
 - x corresponde à subtração de 10 pixéis à posição x do último pássaro do ano em desenho
 - y corresponde à subtração de 5 pixéis à posição y do primeiro pássaro do ano em desenho
 - o comprimento corresponde à subtração da posição x do primeiro pássaro desse ano à posição do último pássaro
 - a altura corresponde a sete pixéis
 - a cor de preenchimento é um bege
 - a cor do bordo é um castanho
 - e, a grossura do bordo é de 0.5.

Imediatamente a seguir ao desenho da barra e, antes de se passar para o ano seguinte, é adicionado o ano:

- adiciona ao Poleiro texto em que:
 - x corresponde à subtração de 40 pixéis à posição x do primeiro pássaro
 - y corresponde à subtração de 5 pixéis à posição y do primeiro pássaro do ano em desenho
 - a cor é um castanho
 - o tamanho da fonte é 10 pontos
 - a fonte é a 'Droid Serif'
 - o texto é centrado
 - os eventos com o rato não existem
 - e o texto é o ano que está a ser desenhado

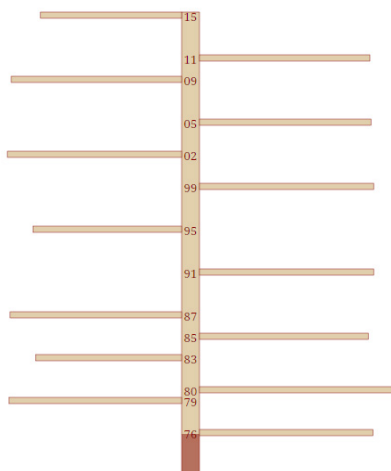


Figura 5.4.8 Resultado final da programação descrita.

Espiral Cronológica

Quando o utilizador seleciona um político – ou seja, um pássaro – é exibida uma nova visualização com a representação das quantidades de notícias que mencionam o político selecionado, ao longo de vinte e oito anos – desde 1987 até à data atual.

Como explicado em §5.1.2, face à necessidade dos dados a representar, conclui-se que a melhor abordagem à representação cronológica seria uma espiral. A programação do desenho desta espiral constitui, no entanto, um desafio cujo processo de solução passou por duas abordagens.

PRIMEIRA ABORDAGEM

A primeira abordagem, para a programação do desenho da espiral, baseou-se no desenho de arcos, cujo raio aumentava a cada círculo de notícias. Na primeira volta de 360°, em cinco pixéis, em dois pixéis e meio e, depois da primeira volta, em 1 pixel.

O cálculo das posições dos círculos utilizava um *array* 'Pub', cujo objetivo era guardar as quantias de notícias anuais, semestrais ou trimestrais, conforme a granularidade escolhida pelo utilizador. Desta forma, era dado um raio inicial de 30 (R), um número inicial de círculos igual a 10 (N), um ângulo de início (A) de 0° e um número de voltas (V) igual a 1. O cálculo das posições era efetuado da seguinte forma:

- percorre o array Pub e por cada posição:
 - calcula a posição $x = R * \coseno\ de\ A + x\ do\ centro\ da\ espiral$
 - calcula a posição $y = R * \seno\ de\ A + y\ do\ centro\ da\ espiral$
 - guarda x no array Pub
 - guarda y no array Pub
 - guarda R no array Pub
 - guarda A no array Pub
 - incrementa A em 2PI a dividir por N
- se A é igual a $V * 2PI$
 - N incrementa em 8 círculos
 - V incrementa em 1 volta
 - se V é igual a 2 (se deu uma volta):
 - incrementa o raio em 2.5 pixéis
 - se V maior ou igual a três:
 - incrementa o raio em 1 pixel
 - noutro caso (se ainda vai na primeira volta):
 - incrementa o raio em 5 pixéis

Após este cálculo, era chamada a função que desenhava a espiral e, a adicionava à *div* 'bolas':

- percorre as posições do array Pub
 - se a posição seguinte `a atual é menor ou igual ao tamanho do array Pub
 - define propriedades do arco:
 - raio de dentro é igual a R
 - raio de fora é igual a R
 - ângulo de início é igual A
 - ângulo de fim é igual ao A do próximo círculo
 - adiciona um caminho "path" à div bolas em que:
 - caminho é o arco
 - o contorno é cinzento

- a espessura do contorno é 0.5
- não tem cor de preenchimento
- e, o centro é igual ao centro da div 'bolas'

O facto de a espiral ser definida através das posições dos círculos, obrigava à alteração da dimensão da mesma, consoante o número de círculos. Pelo que, quando a granularidade das notícias era alterada para anual, a espiral atingiria o seu tamanho mínimo e, quando a granularidade das notícias era trimestral, a dimensão da espiral atingiria o máximo. Para além do problema da alteração das dimensões de uma linha temporal, que representava sempre os mesmos anos – vinte e oito anos –, esta abordagem não era perfeita visualmente. Sendo que o desenho da espiral era quebrado a cada posição de um círculo – figura 5.4.9 – pelo que uma nova solução teve de ser arquitetada.

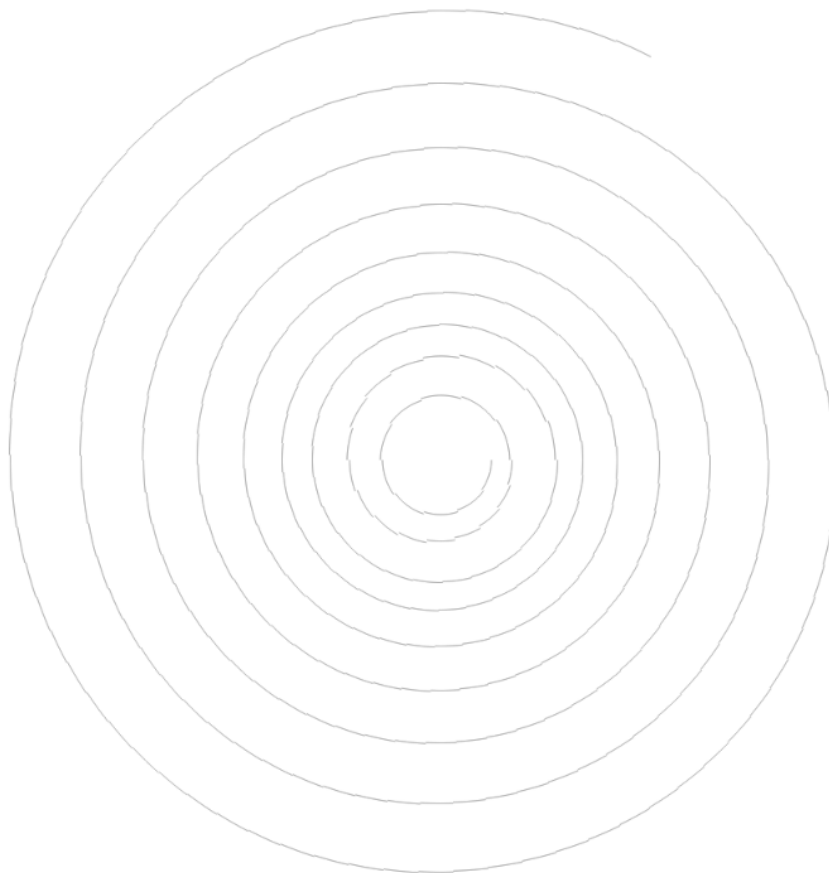


Figura 5.4.9 Resultado final da programação descrita, na qual se pode observar a quebra da linha, principalmente no centro, onde o intervalo é maior.

SEGUNDA ABORDAGEM

Chegou-se à conclusão de que a melhor forma de evitar a quebra da linha, ao desenhar a espiral, seria o desenho de meios círculos, ou seja, arcos de 180°. Assim, para perceber de que maneira estes arcos aumentam de raio sem quebra da linha, foi necessário recorrer ao desenho manual auxiliado de compasso e régua.

Durante a realização do desenho manual percebeu-se que, a espiral desenhada através de arcos de 180°, utiliza dois centros diferentes, um para os arcos inferiores e outro para os arcos superiores – figura 5.4.10.

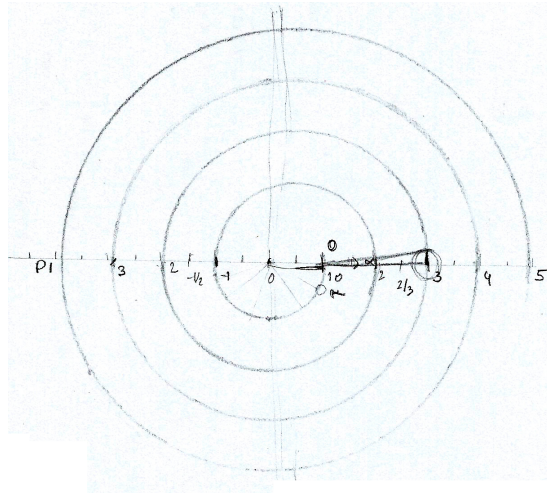


Figura 5.4.10 Digitalização do desenho manual da espiral cronológica.

Este processo de raciocínio ‘manual’ facilitou a passagem do desenho físico a código, pelo que, após alguma experimentação, se decidiu que a espiral teria sete semicírculos e metade do oitavo semicírculo, o que nos permite a construção de uma escala maior para os círculos que representam as notícias.

Para o desenho da espiral, em código, foi necessário primeiro definir algumas variáveis antes do comando de desenho. Utilizaram-se as seguintes variáveis:

WidthBola: guarda o valor do comprimento da *div* ‘bolas’.

HeightBola: guarda o valor da altura da *div* ‘bolas’.

Incremento: é igual a 25. É o valor que determina o incremento do raio relativamente a cada arco.

Theta: é inicializada a zero, mas vai adquirindo novos valores conforme os arcos são desenhados.

Raio: é declarado, inicialmente, com o valor de 90, mas vai aumentando face à soma do incremento cada vez que é desenhado um arco.

H: é inicializada a zero, mas vai somando mais 1 cada vez que um arco é desenhado. Guarda o valor de semicírculos desenhados

PX: declarada, mas não inicializada, servirá para guardar a posição x do ponto em que é iniciado o arco.

PY: declarada, mas não inicializada, servirá para guardar a posição y do ponto em que é iniciado o arco.

PXx: declarada, mas não inicializada, servirá para guardar a posição x do ponto em que é finalizado o arco.

PYy: declarada, mas não inicializada, servirá para guardar a posição y do ponto em que é finalizado o arco.

Após a declaração e inicialização das variáveis descritas, o código comporta-se da seguinte forma:

- enquanto A for menor que 7PI (1260°):
 - se V for par:
 - calcula PX (Raio a multiplicar pelo coseno de theta, mais WidthBola a dividir por 2)
 - calcula PY (Raio a multiplicar pelo seno de theta, mais HeightBola a dividir por 2, menos o incremento)
 - calcula PXx (Raio a multiplicar pelo coseno da soma de theta e

- PI, mais WidthBola a dividir por 2
- Calcula PY (Raio a multiplicar pelo seno da soma de theta e PI, mais HeigthBola a dividir por 2 menos o incremento)
- adiciona à div 'bolas' um caminho:
 - que começa no ponto P_X, P_Y faz uma elipse de raio R,R até ao ponto PX, PY
 - cuja class é "spiral"
 - a cor do contorno é um castanho (#AF9258)
 - a grossura do contorno é 1 pixel
 - o contorno é interrompido a cada 5 pixéis
 - o preenchimento não existe
- incrementa theta em PI
- incrementa H em 1
- incrementa o raio no valor do incremento(25)
- se V for ímpar:
 - calcula PX (Raio a multiplicar pelo coseno de theta, mais WidthBola a dividir por 2, mais o incremento)
 - calcula PY (Raio a multiplicar pelo seno de theta, mais HeigthBola a dividir por 2, menos o incremento)
 - calcula P_X (Raio a multiplicar pelo coseno da soma de theta e PI, mais WidthBola a dividir por 2, mais o incremento)
 - calcula P_Y (Raio a multiplicar pelo seno da soma de theta e PI, mais HeigthBola a dividir por 2 menos o incremento)
 - adiciona à div 'bolas' um caminho
 - que começa no ponto P_X, P_Y faz uma elipse de raio R,R até ao ponto PX, PY
 - cuja class é "spiral"
 - a cor do contorno é um castanho (#AF9258)
 - a grossura do contorno é 1 pixel
 - o contorno é interrompido a cada 5 pixéis
 - o preenchimento não existe
 - incrementa theta em PI
 - incrementa H em 1
 - incrementa o raio no valor do incremento(25)

Este pedaço de código, anteriormente descrito, desenha os sete semicírculos. Faltando a metade do oitavo círculo, que é desenhado de forma semelhante, logo a seguir ao código anterior:

- calcula PX (Raio a multiplicar pelo coseno de theta, mais WidthBola a dividir por 2, mais o incremento)
- calcula PY (Raio a multiplicar pelo seno de theta, mais HeigthBola a dividir por 2, menos o incremento)
- calcula P_X (Raio a multiplicar pelo coseno da soma de theta e PI a dividir por 2, mais WidthBola a dividir por 2, mais o incremento)
- calcula P_Y (Raio a multiplicar pelo seno da soma de theta e PI a dividir por 2, mais HeigthBola a dividir por 2 menos o incremento)
- adiciona à div 'bolas' um caminho:
 - que começa no ponto P_X, P_Y faz uma elipse de raio R,R até ao ponto PX, PY
 - cuja class é "spiral"
 - a cor do contorno é um castanho (#AF9258)
 - a grossura do contorno é 1 pixel
 - o contorno é interrompido a cada 5 pixéis

- o preenchimento não existe
- incrementa theta em PI
- incrementa H em 1
- incrementa o raio no valor do incremento(25)

Escalas das notícias

O crescimento da quantidade trimestral de notícias publicadas nos *media*, ao longo dos vinte e anos representados, é inegavelmente fascinante. Este acontecimento deve-se, principalmente, à facilidade em comunicar cada vez mais rápido graças ao desenvolvimento da tecnologia e das telecomunicações. Assim sendo, é normal que quantidades trimestrais de notícias cheguem facilmente à casa dos milhares. Nesta situação, é óbvio que tem de ser aplicada uma escala para a representação de números tão imponentes.

De acordo com Tuffte³, as boas visualizações devem ter uma escala relevante, pelo que, neste projeto, esta foi uma tarefa que passou por várias soluções, até à solução que melhor transmitia a mensagem certa.

PRIMEIRA ABORDAGEM

A primeira abordagem definia uma escala individual para cada granularidade. Nesta fase do projeto as granularidades disponíveis eram, ainda, mensal, diária e anual. Logo, na função que escala as quantidades das notícias era necessário definir condições para verificar a granularidade e selecionar a devida escala.

A escala definida para a granularidade mensal definia um intervalo de 50 unidades, iniciando com o raio de dois pixéis e aumentando um pixel em cada intervalo, ou seja:

- se a granularidade é mensal:
 - se a quantidade de notícias é maior que zero e menor ou igual a cinquenta
 - então raio igual a dois pixéis
 - se a quantidade de notícias é maior que cinquenta e menor ou igual a cem
 - então raio igual a três pixéis
 - se a quantidade de notícias é maior que cem e menor ou igual a cento cinquenta
 - então raio igual a três pixéis
- (...)

E assim sucessivamente até o raio ser igual a quinze – raio máximo para os círculos.

Para a granularidade diária e para a granularidade anual, o mesmo acontecia diferindo apenas no intervalo. Para a quantidade diária de notícias era utilizado um intervalo de 5 unidades e para a quantidade anual de notícias era empregue um intervalo de 150 unidades. Esta solução – a escala – mostrou-se pouco informativa, pelo que foi necessário pensar numa nova escala.

SEGUNDA ABORDAGEM

Ao contrário da primeira abordagem, a segunda definia uma única escala de intervalos para todas as granularidades – anual, semestral, trimestral e mensal. O processo do código para escalar as quantidades era similar ao anterior, com as mesmas condições que verificam intervalos de 50 posições.

Esta solução provou ser ainda pior do que a anterior, visto que se perdia muita informação, ao utilizar uma escala igual para todas as granularidades. O que acontecia era que a granularidade anual representava os círculos de notícias maiores e a granularidade diária os menores, pelo que, o tamanho dos círculos, em cada granularidade, pouco diferia.

Abandonou-se, então, esta abordagem, em prol de uma que reforçasse a variação das quantidades de notícias em cada granularidade.

TERCEIRA ABORDAGEM

A terceira abordagem volta, assim, às escalas individuais mas, desta vez, com intervalos tratados. Chegou-se à conclusão que a melhor maneira de definir os intervalos para as escalas seria descobrir o mínimo e o máximo da quantidade anual, semestral, trimestral e mensal de notícias. Com estes valores, foram calculadas as amplitudes para cada granularidade – diferença entre o máximo e o mínimo. Calculou-se também a amplitude da nossa escala, ou seja, do valor mínimo e máximo do raio que os círculos que representam as notícias poderiam atingir – uma escala de 2 a 15, logo com 14 unidades de amplitude.

Por fim dividiram-se as amplitudes de cada granularidade pela amplitude da nossa escala. Pelo que obtivemos os seguintes resultados:

Intervalo anual ≈ 1700
 Intervalo semestral ≈ 965
 Intervalo trimestral ≈ 675
 Intervalo mensal ≈ 285

As condições para escalar as quantidades de notícias passavam agora, a utilizar os seguintes intervalos:

- se a granularidade é mensal
 - se a quantidade de notícias é maior que 0 e menor ou igual a 285
 - então raio igual a dois pixéis
 - se a quantidade de notícias é maior que 285 e menor ou igual a $285 \cdot 2$
 - então raio igual a três pixéis
 - se a quantidade de notícias é maior que $285 \cdot 2$ e menor ou igual a $285 \cdot 3$
 - então raio igual a três pixéis
- (etc)

O mesmo acontecia com as outras granularidades alterando apenas o valor do intervalo, consoante indicado acima.

O resultado da aplicação desta abordagem deixava a maioria dos círculos com dimensões pequenas, visto que as quantidades máximas

poucas vezes são atingidas. Após a terceira tentativa falhada, a conclusão era óbvia: a utilização de uma escala com intervalos não era a solução adequada aos dados.

QUARTA ABORDAGEM

Seguidamente à verificação de que o problema das abordagens anteriores não estava no intervalo, mas sim no escalamento por intervalos, a quarta abordagem teve obrigatoriamente de seguir por um caminho diferente.

A utilização de um método de mapeamento da quantidade de notícias para valores pretendidos – de 2 a 15 pixéis – foi a melhor solução encontrada para possibilitar uma boa perceção da variação da quantidade de notícias.

Para este mapeamento, o *d3* tem uma função que produz escalas lineares, recebendo um domínio (*domain*) e uma variação (*range*). O domínio recebe os valores máximo e mínimo da quantidade das notícias – sendo que o mínimo é 1, uma vez que não queremos representar zero notícias. Por sua vez, a variação recebe os valores máximo e mínimo que os raios dos círculos podem atingir. Neste caso, o aumento do raio entre cada arco da espiral cronológica, permitiu aumentar o raio máximo e mínimo dos círculos de notícias. Assim, quando antes o intervalo de raios era entre 2 e 15, agora passou a ser entre 4 e 25.

A aplicação desta solução, utilizando o valor máximo de notícias sobre o político Pedro Passos Coelho – político que atinge o máximo da quantidade de notícias relativamente a todos os outros políticos representados –, concluiu que a utilização de um número máximo geral impedia uma boa transmissão das quantias para os outros políticos. Rapidamente este problema foi solucionado, individualizando os domínios, ou seja, para cada político o máximo de notícias era alterado, permitindo a utilização de todo o leque de raios e transmitindo claramente a variação das notícias do político selecionado.

Ao receber os dados do pedido das quantias de notícias ao servidor do SAPO, depois do tratamento destes e armazenamento nos devidos *arrays* – sendo que as quantias relativas ao primeiro político são armazenada no *array* ‘Pub’ e as quantias relativas ao segundo político selecionado são armazenadas no *array* ‘PubE2’ –, a função cujo objetivo era o mapeamento dos dados era chamada e, a esta, era passado o *array* que precisava dos valores mapeados.

Esta função, de agora em diante designada por ‘Scale’, executa quatro passos. O primeiro tem como objetivo recolher as quantias das notícias e armazena-las num novo *array* ‘num’. Este passo é necessário, pois a quantidade máxima é encontrada a partir de um *array* de números, o que não acontece com o *array* ‘Pub’, pois este guarda, também, as datas a que correspondem as quantias de notícias. Assim, este processo acontece da seguinte forma:

- cria novo *array* ‘num’;
- percorre as posições do *array* recebido:
 - se o número de notícias guardado no *array* recebido é diferente de zero:
 - então adiciona ao *array* “num”, o número de notícias

O segundo passo é tão simples como aplicar uma função matemática do *Javascript* – *Math.max.apply()* – que recebe o *array* ‘num’ e devolve o número máximo nele contido. Uma vez obtido o número máximo podemos passar ao terceiro passo que consiste em definir o comportamento da escala através da função *scale.linear()* do *d3*:

```
var linearScale= d3.scale.linear()
                    .domain([1,máximo obtido])
                    .range([4,25]);
```

Por fim, resta apenas aplicar esta escala linear aos valores contidos no *array* recebido pela função ‘Scale’:

- percorre as posições do array recebido:
 - se o número de notícias guardado no array recebido é diferente de zero:
 - então o raio é igual a *linearScale* do número de notícias;
 - se não:
 - então raio é igual a zero;
- adiciona o raio ao array recebido

No que diz respeito à escala das notícias, esta solução veio resolver a maioria dos problemas experienciados até esta fase do projeto, no entanto, verificou-se que ainda não estava completa, pois levantava um novo problema. Face à individualização das escalas para todos os políticos, acontecia que, quando dois políticos eram selecionados, os círculos de notícias dos dois atingiam o raio máximo, o que causava situações em que um círculo com menor quantidade de notícias fosse maior que um círculo mais pequeno com maior quantidade de notícias. Desta forma, concluiu-se que a escala aplicada ao político com maior quantidade de notícias – de entre os dois políticos selecionados – deveria ser também aplicado ao outro, possibilitando a comparação.

A nível de código, a diferença consistiu em executar o primeiro, segundo e quarto passo anteriormente descritos para cada um dos políticos selecionados – quando um segundo político está selecionado. Era, portanto, necessário obter os máximos das quantidades de notícias relativos a cada um, verificar qual desses máximos era o maior e aplicar a escala do político com maior número de notícias às quantias de notícias do outro político, redesenhando os círculos do primeiro quando este calhava ser o político com menores quantidades de notícias.

Áreas Seleccionáveis

Na visualização das notícias, o aumento dos raios – permitido ao intervalo entre os arcos da espiral – foi, não só uma vantagem, mas também uma necessidade. Apesar deste projeto estar a ser realizado, principalmente, para computador, um dos objetivos futuros é a possibilidade de utilização desta visualização, também, em aparelhos móveis. Posto isto, deve ser tido em conta que elementos muito pequenos podem ser difíceis de clicar, quando é feito um *zoom in*. Decidiu-se então, que a solução mais confortável para o utilizador, ao precisar de passar o rato sobre – ou clicar – nos círculos de notícias, seria a criação de áreas circulares maiores e invisíveis por cima destes – figura 5.4.11.

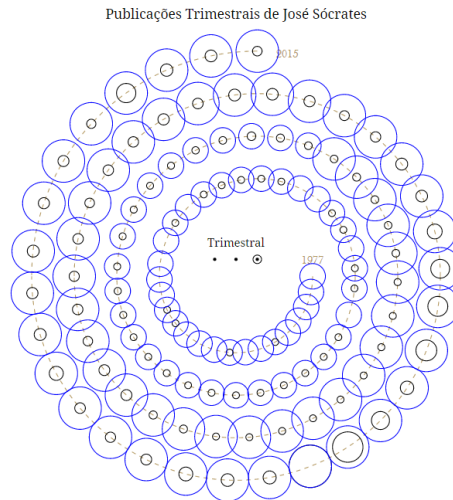


Figura 5.4.11 Áreas selecionáveis a colorido sobrepostas aos círculos das notícias. Estas áreas não são visíveis durante a utilização da visualização.

Assim, por cada círculo desenhado, é adicionado à *div* ‘bolas’ outro círculo, cuja cor de preenchimento e contorno é transparente. Interessa explicar que a interação com o rato – *click*, *mouseover* ou *mouseout* – é facilmente conseguida graças ao *d3*. É tão fácil definir a interação com o rato, tal como definir a cor do elemento ou a sua posição. A dificuldade em definir a interação com o rato, depende apenas do que acontece – ou seja, da função que é executada – quando essa interação existe.

Gráficos Tarte

Ao fazer *Enter* de uma palavra introduzida no campo da pesquisa, vai ser feito um pedido, ao servidor do SAPO, para obter a quantidade de notícias que mencionam a palavra inserida. Quando dois políticos se encontram selecionados, este pedido duplica, a fim de se obter o número de notícias que mencionam a palavra inserida no total das notícias de cada político. Estas quantidades são armazenadas num ou dois *arrays* globais diferentes, consoante o número de políticos selecionados, pelo que, quando chega o momento de representar estas quantidades, facilmente conseguimos aceder aos dados.

Antes de criarmos o desenho – *path* – é feita uma filtragem aos dados para que não sejam criados desenhos sem raio. Apesar destes desenhos invisíveis não fazerem diferença ao utilizador, que não os vê, eles são criados e adicionados à página, contribuindo para o aumento do tempo necessário ao desenho de todas as quantias destas notícias específicas. A triagem é feita da seguinte maneira:

- percorre o array com as quantidades das notícias pesquisadas:
 - se o número de notícias pesquisadas e o número total de notícias do político for diferente de zero
 - então calcula a ângulo correspondente à fatia das notícias específicas

Logo após a triagem são desenhados os elementos, tendo em atenção que, as notícias do segundo político selecionado, devem ter opacidade para que se consiga observar as notícias do primeiro:

- especifica os atributos do arco:
 - raio interior : zero
 - raio exterior : raio do círculo das notícias totais
 - ângulo inicial: zero
 - ângulo final : ângulo correspondente à fatia das notícias específicas

- adiciona um caminho à div 'bolas':
 - em que o caminho é definido pelo arco especificado
 - cuja posição é igual ao centro do círculo das notícias totais
 - a cor de preenchimento varia conforme a cor da seleção do político
 - a opacidade varia conforme a cor da seleção do político

À semelhança do que acontece com os círculos das notícias, também para estes gráficos tarte, foram desenhadas áreas selecionáveis com um raio de maiores dimensões – figura 5.4.12. O processo de desenho é semelhante ao descrito anteriormente, diferindo, apenas, no raio final do arco, que passa a ter dimensões maiores para facilitar a interação com o rato.

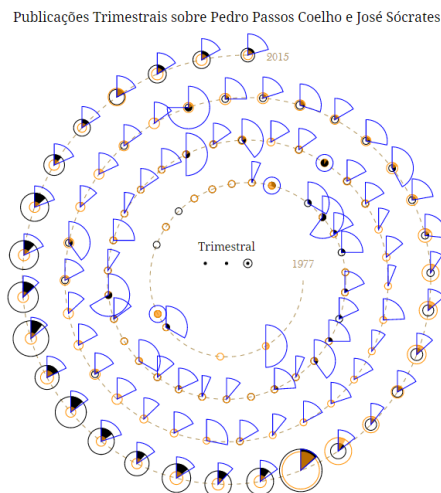


Figura 5.4.12 Áreas selecionáveis a colorido sobrepostas aos gráficos tarte. Estas áreas não são visíveis durante a utilização da visualização.

Filtragem de notícias

Quando o utilizador pretende ver as notícias – título, data e fonte – que fazem parte de um círculo de notícias, instintivamente clica no círculo. Esta ação vai dar origem a um pedido de notícias ao servidor e, estas, serão processadas e armazenadas em *arrays* diferentes, consoante o número de políticos selecionados.

Relativamente à resposta a este pedido, por vezes, são recebidas notícias duplicadas que resultam de uma edição feita a uma dada notícia, sendo que o título e a fonte se mantêm, mas não a data de publicação ou, mais corretamente, a data de edição. Assim, foi necessário criar uma função que percorre os títulos de todas as notícias armazenadas e verifica se existem títulos duplicados. Quando encontra um duplicado, então, remove esta notícia do *array* em que estava armazenada.

Apesar do esforço realizado para evitar notícias duplicadas, ainda assim, por vezes, são apresentadas algumas, nas quais o título difere ligeiramente. Basta uma vírgula, umas aspas ou qualquer outra ligeira

diferença para que estas notícias já não sejam detetadas como duplicadas pelo código, podendo levar à sua repetição, apesar desta situação estar mais contida.

Filtragem de *tags*

Quando o utilizador faz uso do botão de *Tags*, após seleção de um político, são apresentadas as *tags* por ordem decrescente, ou seja, da mais utilizada para a menos utilizada. Note-se que o pedido para as *tags* é diferente do pedido para uma palavra pesquisada. Enquanto, o pedido por palavras devolve as notícias que mencionam a palavra pesquisada, o pedido de *tags* devolve as notícias que os jornalistas – a pessoa que publica a notícia – associaram a essa *tag*.

Reparou-se, no entanto, que o resultado do pedido efetuado devolvia sempre as mesmas *tags*, para os diferentes políticos:

'Partidos e movimentos', 'Diplomacia', 'Economia', 'Radar', 'Portugal', 'Governo', 'Desporto', 'Media', 'Agricultura e Pescas', 'Ambiente', 'Transportes', 'Acidentes e Desastes', 'Parlamento', 'Constituicao', 'Religiao', 'Noticias', 'Sporting', 'Benfica', 'Porto', 'Ministros', 'Bloco de esquerda', 'Legislativas', 'Manuela ferreira leite', 'Paulo portas', 'Jose socrates', 'Cavaco silva', 'Pedro passos coelho', 'Jeronimo de sousa', 'Francisco louca', 'Antonio jose seguro', 'Manuela moura guedes', 'Alberto joao jardim', 'Antonio costa', 'Baptista bastos', 'Jardim goncalves', 'Alberto martins', 'Antonio vitorino', 'Rui pereira', 'Jorge coelho', 'Antonio gutermes', 'Psd', 'Cds/pp', 'Cdu', 'Ugt', 'Ps', 'Cgtp', 'Conselho de ministros', 'Presidente da republica', 'Republica', 'Primeiro ministro', 'Guerras e Conflitos', 'Oposicao', 'Historia', 'Acidentes e Desastes', 'Luis amado', 'Artes', 'Orcamento', 'Rtp', 'Tvi', 'Sic', 'Lusa', 'Angola', 'Futebol', 'Congresso', 'Chefes de estado', 'Caca', 'Autoridades locais', 'Madeira', 'Cemiterio', 'Acidentes de transporte', 'Policia', 'Paulo rangel', 'Rui rio', 'Miguel Relvas', 'Candidatos eleitorais', 'Sondagens'.

Esta situação acontece, em virtude de todas as personalidades representadas estarem ligadas à política e, portanto, partilharem de notícias com os mesmos temas.

Para evitar que todos os políticos devolvessem o mesmo resultado, foi realizada uma recolha das palavras comuns e, posteriormente, estas foram armazenadas num *array* 'filtro' para possibilitar a filtragem destas.

A triagem destas *tags* processa-se da seguinte forma:

- percorre o *array* em que estão guardadas as *Tags*
 - cria um novo *array* 'TagsArray'
 - percorre o filtro
 - se a *Tag* é diferente de todas as que estão no filtro
 - então, adiciona *tag* ao *array* 'TagsArray'

Após a triagem de todas as *tags*, é passado o novo *array* – 'TagsArray' – à função que vai escrever estas palavras no espaço dedicado às *tags*. A aplicação do filtro revelou que o político Hermínio Martinho deixava de apresentar *tags*, pelo que, para este político, foi dispensada a função que filtra as *tags*.

Fatores socioeconómicos

A visualização dos fatores socioeconómicos é apresentada, logo, na página principal, ao lado da visualização das eleições. Esta visualização faz uso dos dados armazenados em ficheiros CSV que são carregados e armazenados em *arrays*, enquanto a introdução é apresentada. Esta visualização é construída em duas fases: a primeira preocupa-se com a construção da base da visualização – linhas e anos que fazem parte da representação cronológica – e, a segunda, preocupa-se com o desenho dos pontos e linhas que representam os fatores socioeconómicos. Na primeira fase é necessário calcular o número de anos representados inicialmente – desde 1960 a 1976 –, o que totaliza cerca de dezasseis anos. Como tal, por cada ano, é calculada a posição de cada linha e, seguidamente, esta é adicionada.

- por cada ano:
 - calcula a posição *xUm*
 - calcula posição *xDois*
 - calcula posição *yUm*
 - calcula posição *yDois*
 - adiciona uma linha à div 'socioEconomic' em que:
 - *x1* é igual a *xUm*
 - *y1* é igual a *yUm*
 - *x2* é igual a *xDois*
 - *y2* é igual a *yDois*
 - contorno tem espessura 1
 - e a cor do contorno é cinzento claro;

Logo de seguida é adicionado o ano que está a ser desenhado. Contudo, neste caso, não basta apenas adicionar o texto à *div* certa. Para que o texto acompanhe o círculo, é necessário definir o arco em que o texto assenta. Esta ação é realizada, ainda antes de se começar a percorrer os anos, uma vez que apenas precisamos de um único arco:

- define as propriedades do arco:
 - o raio interior é igual ao raio Máximo que define o círculo criado pelas linhas
 - o raio exterior é igual ao raio interior
 - o ângulo inicial é zero
 - e o ângulo final é 2π

Seguidamente é criado o arco com as propriedades definidas:

- adiciona um caminho *svg* à div 'socioEconomic' em que:
 - o caminho é definido pelo arco
 - o contorno é transparente
 - e a posição encontra-se no centro da div

Depois deste processo, todas as condições estão reunidas para adicionar os anos à *div*. Assim, logo a seguir à adição das linhas, segue-se a adição dos anos:

- adiciona texto div 'socioEconomic' em que:
 - o tamanho da fonte é 10pt
 - a fonte é "Droid Serif"
 - a cor da letra é cinzento claro
 - adiciona um caminho para o texto em que:
 - é definido pelo arco
 - começando num dados ângulo

- o texto é centrado
- o texto é o ano



Figura 5.4.13

Resultado do código explicado anteriormente.

Na segunda fase, são desenhados os pontos e as linhas para cada tema. A função, que executa estes comandos, recebe a lista de números – para cada ponto –, o máximo, o mínimo e o tema do fator socioeconómico a desenhar. Recebe, ainda, a quantidade de anos a ser representada e o ano mais antigo.

Note-se que a lista de números recebida, apenas contém os anos necessários à representação, ou seja, inicialmente são recebidos apenas dezasseis valores. No entanto, quando o tempo cronológico da representação é aumentado, a quantidade de números da lista também aumenta.

Para que seja possível calcular as posições de cada ponto é necessário, primeiro, fazer o mapeamento dos valores. Consequentemente, à semelhança do que foi explicado para as escalas das notícias, é dado um domínio – *domain* – compreendido entre o máximo e o mínimo do fator socioeconómico em desenho e uma variação – *range* – que compreende valores entre o raio interior e o raio exterior do arco base, que define a área desta visualização.

A aplicação desta escala aos valores da lista, diferentes de zero, permite-nos o posicionamento dos pontos dentro da área da visualização. São, então, calculadas as posições e adicionadas a um novo *array* 'lineData'. No entanto, antes do desenho dos pontos, deve ser desenhada a linha que une todos estes pontos.

Uma vez que queremos uma linha que liga mais de dois pontos, devemos definir as propriedades da linha para podermos criar um caminho – *path* – definido pelas linhas que ligam os diferentes pontos.

- define as propriedades da linha:
 - cujo x corresponde ao x guardado em 'lineData'
 - cujo y corresponde ao y guardado em 'lineData'
 - e a interpolação é 'cardinal'

Ao contrário do que acontece com a interpolação linear- figura 5.4.14 – a interpolação cardinal permite-nos controlar a forma da linha para que esta assuma uma forma arredondada do interseção dos pontos – como é mostrado na figura 5.4.15.



Figura 5.4.14 Interpolação linear.

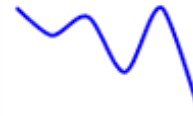


Figura 5.4.15 Interpolação cardinal.

Após a definição das propriedades da linha, podemos então criar o caminho que ligará todos os pontos.

- adiciona um caminho svg à 'socioEconomic' em que:
 - o caminho é definido pelas linhas do tema
 - a cor do contorno corresponde à cor do tema
 - a cor de preenchimento é nula
 - e a espessura do contorno é 1.5;

Apesar da espessura do contorno ser apenas 1.5 pixels, é criada uma área clicável, para uma interação com o rato facilitada, cujo contorno tem uma espessura maior. Isto também, acontece com os pontos. Finalmente, podemos desenhar os pontos:

- por cada posição do array 'lineData':
 - adiciona um círculo à 'socioEconomic' em que:
 - cujo x corresponde ao x guardado em 'lineData'
 - cujo y corresponde ao y guardado em 'lineData'
 - a cor do preenchimento corresponde à cor do tema
 - o raio é 1.85

Quando esta visualização é alterada, face à utilização do controlo temporal, todo este processo é repetido, diferindo apenas no número de anos a representar, pelo que tem de se renovar o gráfico, calculando novamente o intervalo entre cada ano.

Ajuda

A ajuda é disponibilizada quando o utilizador seleciona o botão de ajuda. Contudo, se o utilizador ainda não explorou a visualização ao ponto de ter selecionado um político, a ajuda só é aplicada à página principal, não sendo dada informação sobre a visualização das notícias.

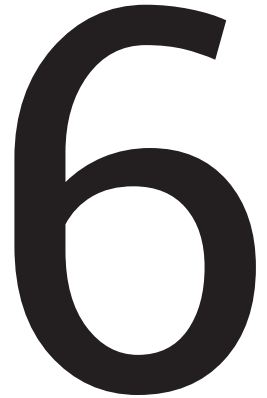
Face ao objetivo da ajuda, que implica não só informar sobre as funcionalidades disponíveis, como também, a localização das diferentes informações às quais, os utilizadores, podem ter acesso, era imperativo encontrar uma solução para este problema.

Este problema começa no facto de nenhum político estar selecionado, por isso, foi necessário pesquisar uma maneira de simular o click num político a partir de código. Descobriu-se, então, que o *d3* dispõe de uma função – *d3Click* – que soluciona perfeitamente o problema, ao simular o *click* num elemento à nossa escolha.

Ao simular o *click* num político é revelada a visualização das notícias, permitindo explicar as suas funcionalidades. Da mesma forma, são reveladas mensagens de ajuda.

Avaliação

Este capítulo serve para documentar e analisar a avaliação da solução desenvolvida, retirando consequências para a proposta de design das visualizações. Os testes de usabilidade foram efetuados na primeira semana de agosto – que teve início no dia 3 de agosto e finalizou no dia 9 de agosto – do corrente ano. Este pretendia recolher dados sobre a eficiência da visualização, tanto a nível de comunicação, como a nível de navegação e interação.



6.1. Metodologia

A metodologia utilizada para a avaliação da usabilidade da visualização de informação, foi a realização de testes informais de usabilidade com utilizadores reais. Este teste foi preparado para ser realizado sem supervisão, contudo, foram observados onze testes para facilitar a compreensão de possíveis problemas de usabilidade. A observação dos testes é especialmente importante, pois é a forma mais efetiva de entender o que funciona, o que não funciona e o que causa problemas aos utilizadores²⁴. Durante os testes, foi pedido ao utilizador que “pense alto”, isto é, que exponha o seu raciocínio, que expresse os seus sentimentos, emoções e dúvidas, enquanto realiza as tarefas pedidas. É, portanto, utilizado o protocolo “*Think-Aloud*”.

Os testes foram gravados através da captura de vídeo no ecrã e de áudio, registando o pensamento do utilizador, possibilitando completar ou confirmar as notas retiradas durante os testes.

6.1. 1. Requisitos

Os participantes dos testes de usabilidade com observação, utilizaram um computador ASUS com o sistema operativo *Windows 7* com o programa *Rylstim Screen Recorder* instalado para gravação do ecrã durante os testes. O computador tem instalada a VPN do SAPO que se manterá ligada durante a realização dos testes para a busca dos dados necessários às visualizações de informação. O computador contém um microfone incorporado para gravar a conversa do utilizador. Apesar do computador ter instalado dois *browsers* diferentes – *Google Chrome* e *Internet Explorer* – é pedido ao utilizador que faça utilização do *Google Chrome*.

Os participantes dos testes de usabilidades sem observação – funcionários do SAPO – utilizarão um computador que tenham disponível com características desconhecidas. A estes, não será pedida a gravação do teste. Serão pedidas as dificuldades e tempos de execução das tarefas, bem como a avaliação da gravidade das dificuldades encontradas. Deverão ter instalada a VPN do SAPO para o bom funcionamento do site e aceder ao site utilizando o *Google Chrome*.

6.1. 2. Sessão

As sessões de observação de testes de usabilidade terão uma duração de, aproximadamente, uma hora por participante.

Antes do utilizador começar a primeira fase do teste, são explicadas as diferentes fases que se seguem e é pedida a autorização para a gravação do teste, evidenciando o facto de que apenas o ecrã e a voz são gravados. O teste de usabilidade é composto por três partes: questionário sobre o perfil do utilizador (5 minutos), execução das tarefas pedidas (40 minutos) e, por último, resumo das dificuldades sentidas, opiniões e sugestões (15 minutos).

Depois de concluída a primeira fase do teste pelo utilizador, é pedido ao mesmo que inicie a gravação do teste e execute as tarefas pedidas, obtendo a informação no site da visualização de informação – já aberto numa *tab* ao lado.

Nesta segunda fase, as tarefas serão cronometradas pelo observador e serão anotados os comportamentos do utilizador. A cada tarefa é pedido ao utilizador que identifique as dificuldades encontradas e lhes atribua uma classificação consoante a sua gravidade.

Ao concluir a segunda fase do teste, a gravação é parada e é pedido ao utilizador que resuma as suas dificuldades durante a utilização da visualização e sugira soluções para a solução dos problemas encontrados. É nesta fase que o utilizador poderá dar todas as suas opiniões sobre o site, incluindo opinião sobre a utilidade da visualização e melhorias que gostaria de ver aplicadas.

6.2. Participantes

Para a realização dos testes de usabilidade com observação, o objetivo seria recrutar no mínimo cinco utilizadores. De acordo com Nielsen, cinco utilizadores é o mínimo suficiente para identificar mais de 75% dos problemas de usabilidade do produto em teste²⁰.

Seria também importante recolher opiniões de funcionários do SAPO, mas para facilitar a realização dos testes a estas pessoas que se encontravam a trabalhar, não foi pedida gravação nem foi possível observar os testes. Assim, os participantes em questão puderam realizar o teste quando o tempo lhes permitiu.

Apesar de poder ser interessante planejar o recrutamento dos participantes por perfis, a altura em que os testes foram realizados, bem como o tempo planeado para esta tarefa, não permitiu um recrutamento rígido baseado em perfis. Poderia ter sido interessante recrutar jornalistas e politólogos, contudo, o facto de não ter sido possível testar com estes perfis nesta primeira fase de testes, não invalida a sua participação em testes de usabilidade futuros.

Este subcapítulo analisa e resume os perfis dos participantes do teste de usabilidade, bem como os dados relativos à experiência e conhecimentos dos participantes em questão. Esta análise não terá em conta a diferenciação dos participantes observados e dos participantes não observados.

6.2.1. Perfis de utilizador

Durante a fase de testes, foram realizados um total de 14 testes com utilizadores de diferentes áreas de ocupação e com idades entre os dezassete e os cinquenta anos. Seguem-se os resultados dos dados de perfil recolhidos.

1. Idades

- 20 ou menos: 1
- entre 21 e 30: 7
- entre 31 e 40: 3
- entre 41 e 50: 3

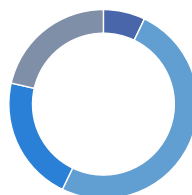


Gráfico 1

Distribuição das idades.

2. Frequência na utilização do PC

- Nunca: 0
- Raramente: 1
- Medianamente: 0
- Frequentemente: 2
- Muito frequentemente: 11

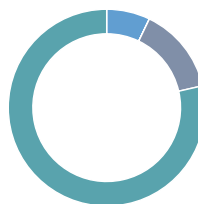


Gráfico 2
Distribuição da frequência no uso de PC.

3. Frequência na utilização de Tablet

- Nunca: 4
- Raramente: 2
- Medianamente: 3
- Frequentemente: 3
- Muito frequentemente: 2

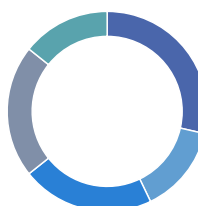


Gráfico 3
Distribuição da frequência no uso de Tablet.

4. Frequência na utilização de Smartphone

- Nunca: 0
- Raramente: 1
- Medianamente: 1
- Frequentemente: 1
- Muito frequentemente: 11

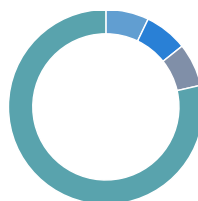


Gráfico 4
Distribuição da frequência no uso de Smartphone.

5. Frequência na utilização de um browser

- Nunca: 0
- Raramente: 0
- Medianamente: 1
- Frequentemente: 3
- Muito frequentemente: 10

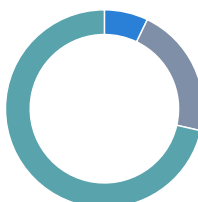


Gráfico 5
Distribuição da frequência no uso de um browser.

6. Frequência na visualização de dados ou análise de dados

- Nunca: 2
- Raramente: 2
- Medianamente: 3
- Frequentemente: 2
- Muito frequentemente: 5

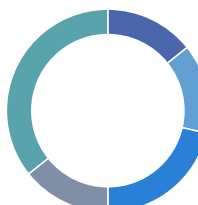


Gráfico 6
Distribuição da frequência na visualização de dados.

7. Frequência de atenção em assuntos políticos

- Nunca: 0
- Raramente: 5
- Medianamente: 2
- Frequentemente: 1
- Muito frequentemente: 3
- Sem resposta: 3

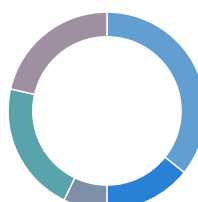


Gráfico 7
Distribuição da frequência de atenção em assuntos políticos.

8. Costumam votar?

- Sim: 9
- Não: 2
- Sem resposta: 3

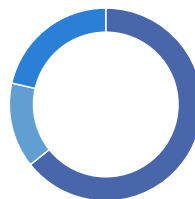


Gráfico 8
Distribuição de eleitores entre a totalidade de participantes.

9. Profissão

- Engenheiro Informático/ Programador/ Técnico de informática: 4
- Innovation and Business Development Manager : 1
- Estudante do Ensino Secundário: 1
- Estudante do Mestrado em Design e Multimédia: 2
- Designer : 2
- Bióloga : 1
- Distribuidor dos CTT: 1
- Professor: 2

6.3. Questões para a pesquisa

As questões para a pesquisa são pontos de foco para os testes. Elas descrevem as principais dúvidas a que há necessidade de responder. Para esta primeira fase de testes as dúvidas centram-se em perceber se:

- A informação dada na página inicial chega para introduzir e clarificar os assuntos da visualização?
- O utilizador percebe o que são publicações?
- O utilizador percebe que os dados das publicações estão, inicialmente, organizados por Trimestre?
- O utilizador entende que os botões de pesquisa e das características das notícias são apenas utilizados na visualização das publicações relativas ao político selecionado?
- O utilizador percebe que pode controlar o tempo representado no gráfico dos fatores socioeconómicos?
- O utilizador fica desconfortável com o *scroll* horizontal?

6.4 Tarefas

Observar a interação do utilizador com a aplicação em teste é o modo mais eficiente de perceber as razões que causam problemas aos utilizadores. Para que esta observação seja realizada com sucesso, é necessário pedir aos utilizadores que façam alguma ação ou tarefa. Só assim, será possível observar os passos realizados pelo utilizador para atingir o objetivo da tarefa e identificar as dificuldades durante a sua realização.

Neste subcapítulo o leitor poderá conhecer as tarefas realizadas pelos participantes do teste, através da descrição da tarefa que contém:

Cenário da tarefa: a tarefa que é dada ao utilizador.

Ponto de início: onde é esperado que os participantes comecem a tarefa.

Ponto de fim: onde os participantes devem acabar a tarefa.

Condição de sucesso: quais são as condições que definem o sucesso da tarefa.

Passos esperados: Os passos que são esperados que o participante realize para atingir o objetivo com sucesso.

As tarefas pedidas no teste de usabilidade, efetuado para esta primeira ronda de testes, estão organizadas por visualização e por grau de dificuldade. Segue-se a lista das tarefas:

TAREFA 1

Esta tarefa tem como primeiro objetivo averiguar se os participantes percebem o que fazer para ultrapassar a página de introdução. Seguidamente, esta tarefa verifica se os participantes conseguem interpretar a visualização eleitoral e perceber que os “Primeiros-Ministros”, ou seja, os políticos mais votados, estão situados mais perto da *timeline*. Adicionalmente, é necessário perceber se o utilizador tem dificuldade na perceção e identificação dos anos que estão apenas parcialmente escritos.

Cenário da tarefa: Identifique o(s) ano(s) em que o José Sócrates foi eleito Primeiro-Ministro.

Ponto de início: Homepage (página de introdução).

Ponto de fim: Página da visualização eleitoral.

Condição de sucesso: Os participantes chegam à visualização das eleições e identificam os anos em que o José Sócrates foi mais votado.

Passos esperados: Os participantes carregam no botão “Explorar” dado na Homepage (página de introdução), seguidamente centram a sua atenção na visualização das eleições e procuram o político – José Sócrates – nos pássaros que estão mais próximos da *timeline*. Depois de identificarem o político José Sócrates, observam a *timeline* para conhecerem os anos em que aconteceram as eleições. Voltam a procurar por José Sócrates, nos pássaros que estão mais perto da *timeline* para verificar que o ano descoberto foi o único e encontram novamente o político, realizando os passos já conhecidos para a identificação do ano.

TAREFA 2

Esta tarefa tem como objetivo averiguar se os participantes conseguem encontrar a informação pedida e se é intuitiva a sua localização.

Cenário da tarefa: Identifique o partido com maior votação em 1979.

Ponto de início: Página da visualização eleitoral.

Ponto de fim: Página da visualização eleitoral.

Condição de sucesso: Os participantes percebem que a informação que procuram é relativa à eleição legislativa e, portanto, têm que passar o rato, ou clicar, no poleiro. Os participantes identificam a sigla partidária de acordo com a cor ou posição do pássaro maior.

Passos esperados: Já na página correta devido à tarefa anterior, os participantes procuram, na *timeline*, pelo ano 1979, passam o rato, ou clicam, no poleiro e, face à informação dada, identificam a sigla partidária de acordo com a cor ou posição do pássaro maior.

TAREFA 3

Esta tarefa tem como objetivo perceber se a informação dada ao

clicar no poleiro se torna confusa para os participantes, uma vez que não está alinhada com os pássaros.

Cenário da tarefa: Identifique o partido correspondente a cada cor.

Ponto de início: Página da visualização eleitoral.

Ponto de fim: Página da visualização eleitoral.

Condição de sucesso: Os participantes sabem que devem passar o rato por cima de cada poleiro para identificarem as siglas partidárias e atribuí-las às cores dos pássaros.

Passos esperados: Já na página correta devido à tarefa anterior, os participantes já sabem onde devem ir buscar a informação que procuram, por isso passam o rato por cima do poleiro, lêem as siglas dadas e associam-nas à cor.

TAREFA 4

Esta tarefa tem como objetivo verificar se é fácil para os participantes encontrar a visualização das notícias sobre o político pedido e perceber se conseguem entender que a visualização dada organiza as notícias por trimestre. Pretende saber-se, também, se é fácil compreender que o trimestre com mais notícias corresponde à bola maior. Posteriormente, pretende-se averiguar se a data da bola em questão é visível e se esta se encontra num sítio intuitivo.

Cenário da tarefa: Identifique o trimestre em que Paulo Portas foi mais mencionado pelos *media*.

Ponto de início: Página da visualização eleitoral.

Ponto de fim: Página da visualização das publicações – a mesma página da visualização eleitoral, mas deslocada mais para a direita.

Condição de sucesso: Os participantes carregam no político desejado, percebem que as notícias estão organizadas por trimestre e identificam a bola maior. Passam o rato por cima da bola maior e identificam a data.

Passos esperados: Já na página correta devido à tarefa anterior, os participantes procuram o político desejado e selecionam-no. Seguidamente, é-lhes apresentada uma nova visualização. Lêem o título da visualização que os ajuda a perceber a organização das notícias por trimestres e percebem que apenas têm de identificar a bola maior. Colocam o rato por cima da bola e identificam a data no centro da visualização.

TAREFA 5

Esta tarefa tem como principal objetivo averiguar se o utilizador percebe que é possível comparar dois políticos e quais as notícias que pertencem a cada um. Pretende verificar-se, também, se o utilizador percebe a existência da funcionalidade que lhe permite alterar a granularidade com que as notícias são aglomeradas.

Cenário da tarefa: Identifique qual dos dois políticos foi mais noticiado nos últimos três anos: Pedro Passos Coelho ou José Sócrates.

Ponto de início: Página da visualização das publicações – a mesma página da visualização eleitoral, mas deslocada mais para a direita.

Ponto de fim: Página da visualização das publicações – a mesma página da visualização eleitoral, mas deslocada mais para a direita.

Condição de sucesso: Os participantes selecionam os dois políticos pedidos, alteram a granularidade das notícias para “Anual” e, por

comparação direta, identificam o Pedro Passos Coelho como o político mais noticiado nos últimos três anos.

Passos esperados: Os participantes desseleccionam o político Paulo Portas, que estava selecionado da tarefa anterior, e selecionam os dois políticos pretendidos. Seguidamente, selecionam a granularidade “Anual” e analisam os últimos três anos, ou seja, as últimas três bolas. Relacionam as cores das bolas com a cor dos políticos selecionados e concluem que as notícias de Pedro Passos Coelhos são mais numerosas do que as notícias de José Sócrates.

TAREFA 6

Pretende-se verificar, com esta tarefa, se a pesquisa de notícias que contém uma certa palavra é, ou não, intuitiva. Para que possamos entendê-la como uma tarefa intuitiva é necessário perceber se o botão de pesquisa é visível, se o icon é adequado à funcionalidade e se o participante entende que este botão apenas produz efeito na representação das notícias.

Cenário da tarefa: Identifique o semestre em que a palavra ‘submarinos’ foi mais empregue nas notícias sobre Durão Barroso.

Ponto de início: Página da visualização das publicações – a mesma página da visualização eleitoral, mas deslocada mais para a direita.

Ponto de fim: Página da visualização das publicações – a mesma página da visualização eleitoral, mas deslocada mais para a direita.

Condição de sucesso: Os participantes selecionam apenas o político pretendido – Durão Barroso – e selecionam a granularidade “semestral”. Seguidamente, introduzem a palavra pedida na caixa de pesquisa e identificam a data correspondente à bola que tem a fatia que corresponde à maior percentagem.

Passos esperados: Os participantes desseleccionam os políticos anteriormente selecionados e selecionam o Durão Barroso. Seguidamente, selecionam a granularidade “Semestral” e procuram a caixa de pesquisa. Carregam no botão de pesquisa, introduzem a palavra pedida, carregando na tecla *Enter* e voltam a observar a representação visual das notícias. Passam o rato pelas fatias coloridas que lhes dá as percentagens. Pela análise das percentagens que correspondem às maiores fatias, identificam aquela em que a percentagem é maior e com o rato por cima podem indicar a data que encontram no centro da visualização das notícias.

TAREFA 7

Esta tarefa tem como objetivo perceber se é fácil e intuitivo para o utilizador chegar às notícias dos políticos.

Cenário da tarefa: Identifique o título da notícia mais recente de Francisco Sá Carneiro.

Ponto de início: Página da visualização das publicações – a mesma página da visualização eleitoral, mas deslocada mais para a direita.

Ponto de fim: Página com a visualização das publicações e notícias.

Condição de sucesso: Os participantes selecionam o político pretendido – Francisco Sá Carneiro –, identificam a última bola da espiral e clicam nela. Identificam o título da primeira notícia dada.

Passos esperados: Os participantes desseleccionam o político anteriormente selecionado e selecionam o político pretendido. Seguidamente, procuram a última bola que corresponde às notícias

mais recentes. Carregam na bola para ver as notícias e identificam o título da primeira notícia.

TAREFA 8

Esta tarefa tem como objetivo averiguar se os participantes conseguem identificar qual das palavras dadas é a mais utilizada, ou seja, perceber que a ordem das palavras é decrescente e não aleatória.

Cenário da tarefa: Indique a palavra-chave mais utilizada para descrever as notícias relativas a Manuela Ferreira Leite.

Ponto de início: Página com a visualização das publicações e notícias escritas.

Ponto de fim: Página com a visualização das publicações.

Condição de sucesso: Os participantes selecionam o político pretendido, identificam o botão que lhes retorna as palavras que categorizam as notícias e identificam a mais utilizada.

Passos esperados: Os participantes desseleccionam o político anteriormente selecionado e selecionam o político pretendido. Seguidamente carregam no botão das *Tags/Características* das notícias e identificam a primeira palavra como a palavra que é mais utilizada.

TAREFA 9

Esta tarefa tem como objetivo prioritário verificar se é fácil para o utilizador encontrar as legendas que lhe indicam o tema de cada linha na visualização dos fatores socioeconómicos. Menos prioritário é perceber se o utilizador dá conta que, inicialmente, apenas estão representados fatores desde 1960 a 1976 e avaliar a sua dificuldade na identificação do ‘botão’ que lhe permite controlar o período de tempo, representado na visualização dos fatores socioeconómicos.

Cenário da tarefa: Identifique os diferentes temas representados na visualização dos fatores socioeconómicos (desde 1960 a 2015).

Ponto de início: Página principal com a visualização das eleições e dos fatores socioeconómicos.

Ponto de fim: Página principal com a visualização das eleições e dos fatores socioeconómicos.

Condição de sucesso: Os participantes passam o rato por cima das linhas e indicam o tema sem dificuldade.

Passos esperados: Os participantes desseleccionam o político anteriormente selecionado, voltando para a página que mostra a visualização eleitoral à direita e a visualização dos fatores socioeconómicos à esquerda. Seguidamente, passam o rato por cima das linhas e identificam os temas de cada linha.

Os participantes menos distraídos percebem que o período representado não é o período pedido e, por isso, irão procurar uma forma de aumentar o período visualizado. Conduzirão o rato até à *timeline* e clicarão ou arrastarão o controlo temporal da visualização. Uma vez com o período certo representado, passarão por cima das linhas que ainda não conhecem e identificam o tema a elas associado.

TAREFA 10

Esta tarefa tem como objetivo averiguar se o utilizador consegue associar os números da visualização dos fatores socioeconómicos a

datas e perceber que o período representado inicialmente tem início no ano de 1960 e final do ano 1976. Adicionalmente, pretende-se avaliar a dificuldade do participante em saber como alterar o período temporal representado.

Cenário da tarefa: Identifique o ano em que o número de nascimentos atingiu o mínimo relativamente a todos os outros anos (1960 a 2015).

Ponto de início: Página principal com a visualização das eleições e dos fatores socioeconómicos.

Ponto de fim: Página principal com a visualização das eleições e dos fatores socioeconómicos.

Condição de sucesso: Os participantes aumentam o período temporal representado, selecionam a linha da Natalidade e identificam o ano em que o mínimo atingiu o mínimo.

Passos esperados: Os participantes clicam ou arrastam o rato na *timeline*, aumentando o período temporal representado. Seguidamente, procuram a linha correspondente à natalidade e selecionam-na para uma visualização individualizada que facilitará a leitura deste fator. Identificam o ano em que a linha atingiu o mínimo, ou seja, se encontra mais próxima do centro.

TAREFA 11

Esta tarefa tem como objetivo apurar se é intuitivo para o participante como obter os valores de um tema em cada ano.

Cenário da tarefa: Identifique o número de indivíduos empregados em 2014.

Ponto de início: Página principal com a visualização das eleições e dos fatores socioeconómicos.

Ponto de fim: Página principal com a visualização das eleições e dos fatores socioeconómicos.

Condição de sucesso: Os participantes passam o rato pelo ponto situado em 2014 da linha correspondente ao tema pedido e identificam o valor situado no centro da visualização.

Passos esperados: Os participantes procuram a linha referente à população empregada e clicam nela para individualizar a leitura. Identificam o ano 2014 e passam o rato pelo ponto que se situa nesse ano, identificam o valor no centro da visualização.

TAREFA 12

Esta tarefa tem como objetivo averiguar se é intuitivo o relacionamento da visualização dos fatores socioeconómicos com a visualização das eleições.

Cenário da tarefa: Identifique o(s) político(s) no cargo de Primeiro-Ministro durante os anos em que a dívida pública teve maior crescimento.

Ponto de início: Página principal com a visualização das eleições e dos fatores socioeconómicos.

Ponto de fim: Página principal com a visualização das eleições e dos fatores socioeconómicos.

Condição de sucesso: Os participantes selecionam a linha que diz respeito à Dívida Pública, identificam o intervalo dos anos em que lhes parece haver um maior aumento da dívida e identificam o nome do político no cargo de Primeiro-Ministro.

Passos esperados: Os participantes procuram a linha referente à Dívida Pública e clicam nela para individualizar a leitura. Identificam o intervalo dos anos em que lhes parece haver um maior aumento da dívida e deslocam o rato para a visualização eleitoral, passando o rato por cima do pássaro maior nos anos identificados, de modo a obter o nome do político.

TAREFA 13

Esta tarefa tem como objetivo averiguar se o ícone do botão pedido é intuitivo, ou seja, se o ícone utilizado no botão é percecionado como o botão que lhes permite ler sobre o projeto.

Cenário da tarefa: Identifique o botão que lhe permite ler sobre o projeto.

Ponto de início: Página principal com as visualizações eleitorais e dos fatores socioeconómicos.

Ponto de fim: Página com a informação do projeto.

Condição de sucesso: Os participantes selecionam o botão de informação.

Passos esperados: Os participantes procuram o botão com “i” de informação e carregam nele.

TAREFA 14

Esta tarefa tem como objetivo averiguar se o ícone do botão pedido é intuitivo, ou seja, se o ícone utilizado é percecionado como o botão com funcionalidade de ajuda.

Cenário da tarefa: Identifique a funcionalidade com ponto de interrogação.

Ponto de início: Página com a informação sobre o projeto.

Ponto de fim: Página com a informação de ajuda.

Condição de sucesso: Os participantes reconhecem que a funcionalidade é ajudar.

Passos esperados: Os participantes respondem antes de verificar a funcionalidade do botão. Uma vez que se encontram mais abaixo na página, onde encontram a informação no projeto, sobem a página com o *scroll* vertical da página. Identificam o botão em questão, carregam nele e confirmam a sua resposta.

6.5 Resultados

Este subcapítulo tem como objetivo documentar os erros e, ou, problemas encontrados durante a realização dos testes de utilizador. Serão documentadas as conclusões observadas relativamente às questões em foco e os resultados para as tarefas realizadas. Citações dos participantes poderão ser adicionadas à descrição do problema, quando relevantes, para melhor contextualizar e descrever as opiniões e sentimentos do utilizador.

6.5.1 Resultados para as perguntas em questão

1. A informação dada na página inicial chega para introduzir e clarificar os assuntos da visualização?

Dos participantes observados notou-se que uma grande maioria

tomou atenção às informações dadas inicialmente. Essa informação inicial guiou a maioria dos participantes diretamente para a visualização das eleições e possibilitou uma melhor e mais rápida interpretação. Contudo, dois participantes não tiveram tanta facilidade em interpretar a visualização eleitoral na sua totalidade. Uma das dificuldades encontradas residiu em perceber a linha cronológica - *“Não é imediatamente perceptível no gráfico a forma como está organizada a informação em termos cronológicos.”* (P14) - e a outra foi a interpretação dos elementos da visualização - *“Nos passarinhos, como não tinha nenhum texto, não percebi o que eram.”* (P3)

Apesar de se mostrar ser uma informação útil para a compreensão da visualização das eleições, uma minoria ficou confusa ou “intrigada” com a razão para a existência de duas visualizações diferentes logo na primeira página - *“(…) Depois de clicar no “Explorar” não percebi a página que me apareceu. O único texto que tinha à vista era o do gráfico do lado esquerdo que diz “Fatores socioeconómicos”. Aproveito para referir que no gráfico da esquerda [Fatores socioeconómicos] não percebo o seu significado.”* (P3) , *“Estou intrigado com o gráfico da direita”* (P8).

Apesar de ter sido indicado na introdução a existência de uma visualização de publicações nos *media* referentes aos políticos, ao carregar nos políticos os participantes não estavam à espera que aparecesse um novo gráfico. Alguns participantes mostraram sinais de que esperavam de que a informação das notícias estivesse visível, como acontece com as outras duas visualizações - *“Não tive dificuldade em identificar os assuntos [de cada uma das visualizações]. Só não tive conhecimento imediato que existia uma secção que tinha informação das notícias relativas a cada político.”* (P11)

2. O utilizador percebe o que são publicações?

Apesar de ter pouca gravidade, uma minoria dos utilizadores mostrou-se baralhada com o título da visualização das notícias na primeira leitura. - *“Não sei o que são publicações trimestrais.”*(P4)

Alguns participantes, chegaram mesmo a discordar do título colocando novas hipóteses. - *“ ‘Publicações de’ não, ‘publicações sobre’ político X”* (P14)

3. O utilizador percebe que os dados das publicações estão, inicialmente, organizados por Trimestre?

Durante a realização da tarefa cinco, observou-se que uma minoria razoável dos participantes confundiram a granularidade das publicações, pensando que a informação estava representada por anos, ou seja, cada bola corresponderia a um ano. Os participantes recuperaram deste erro ao lerem o título da visualização ou pela consulta das datas relativas a cada bola.

“ [Dificuldade em] Relacionar as bolas com a unidade temporal. Primeiro pensei que cada bola correspondia a um ano, no entanto como a pergunta anterior a bola maior correspondia a um trimestre cheguei à conclusão que um ano teria 4 bolas” (P6)

“Identificar a periodicidade de análise das publicações.” (P9)

4. O utilizador entende que os botões de pesquisa e das características das notícias são apenas utilizados na visualização das publicações relativas ao político selecionado?

Observou-se que a maioria dos participantes que mostraram a intenção de utilizar uma caixa de pesquisa para filtrar políticos ou partidos, não perceberam pela opacidade dos botões, principalmente do botão de pesquisa, que estes se encontravam inativos. Muitos dos participantes só sentiram necessidade de utilizar a caixa de pesquisa para executar a tarefa seis, altura em que já existia um político selecionado, portanto, altura em que os botões já estavam ativos. Observaram-se dois casos em que, não percebendo a inatividade inicial do botão de pesquisa, os participantes tentaram fazer uso desta funcionalidade sem sucesso. Nestes dois casos, ambos os participantes perceberam mais tarde, ao ler a tarefa seis, qual a funcionalidade do botão de pesquisa. Existiu apenas um participante que por ter falhado na utilização do botão de pesquisa no início do teste, pensou que este botão continuava inativo quando lhe foi apresentada a tarefa seis, tendo demorado na conclusão da tarefa.

“Procurei pelo Paulo Portas primeiro no menu superior, mas as opções não estavam ativas.” (P4)

“Não há como filtrar pássaros por nome.” (P1)

“Queria pesquisar, mas o botão não funcionava.[...]” (P11)

5. O utilizador percebe que pode controlar o tempo representado na visualização dos fatores socioeconómicos?

Uma minoria dos participantes encontrou o controlo temporal por acaso durante a exploração do site, ao realizar as tarefas do teste de usabilidade. No entanto, a maioria dos participantes não sentiu qualquer necessidade de aumentar o período temporal visualizado, até chegarem à sexta ou sétima tarefa do teste de usabilidade. Apenas a indicação que era necessária informação relativa a um ano que não estava a ser mostrado, ativou a necessidade de estender o período temporal. Face a esta necessidade, a maioria dos participantes guiou o rato para a *timeline*, clicando nela ou arrastando a barra mais escura. Assim, conclui-se que não é evidente que existe a possibilidade de aumentar o período temporal da visualização dos fatores socioeconómicos.

“Visualizei facilmente os temas mas só posteriormente reparei que a informação ia até ao ano de 76.

Como na barra estava sombreado até esse ano tentei prolongar o espaço temporal arrastando a barra até ao ano de 2015” (P6)

“ [A dificuldade foi] assimilar que podia controlar a janela de tempo que aparecia no gráfico mais à esquerda.” (P11)

“Fiquei sem perceber se há dados para depois de 1976...” (P1)

“Chamou-me a atenção o facto de esta pergunta referir 2015. Tive de tentar descobrir como mudava as datas. Apercebi-me então da cor diferente que aparece na barra do poleiro que dava para mudar. A dificuldade foi perceber que conseguia mudar a data aí.” (P3)

“Não me apercebi logo que só tinha um intervalo de tempo e não estava a ver toda a barra cronológica.” (P10)

“AHHHHH. Voltei a esta pergunta depois de fazer a anterior porque percebi, depois de ler a ajuda (?) que carregar no poleiro vertical muda o intervalo dos gráficos. Eh pah...não estava mesmo a ver... ok.” (P1)

6.O utilizador fica desconfortável com o *scroll* horizontal?

Uma maioria dos participantes não sentiu necessidade de utilizar o *scroll* horizontal, uma vez que, existe um deslocamento do ecrã ao seleccionar ou desseleccionar um político e ao seleccionar uma bola das notícias, aparecendo os títulos de notícias desse período. Não deixaram de existir, no entanto, participantes que sentiram a necessidade de mover a visualização e que tiveram de fazer uso do *scroll* horizontal. Neste caso, nem todos os participantes acharam esse facto desconfortável, mas concordam que não é muito intuitivo.

“Navegação horizontal através de arrastamento. Não foi problema porque depois encontrei o *scroll*.” (P6)

“Em relação a ser uma navegação horizontal, não tive qualquer problema.”(P10)

“Visto ser uma navegação diferente do habitual não me ocorreu à primeira fazer *scroll* para baixo para ver se conseguia fazer *scroll* na horizontal. Não senti que fosse grave mas não é extremamente intuitivo.” (P13)

6.5.2 Resultados e problemas das tarefas

Nesta secção serão analisados os resultados das tarefas realizadas pelos participantes durante o teste de usabilidade. Estas serão classificadas de acordo com a condição de sucesso de cada tarefa, descrita no capítulo 6.5. Em simultâneo, será feita uma descrição individualizada dos problemas observados e registados pelos participantes.

Seguidamente são apresentadas duas tabelas: a primeira – Tabela 6.6.2.1 – classifica o sucesso ou insucesso da tarefa por participante, de agora em diante nomeado por ‘P’ e a segunda – Tabela 6.6.2.2 – regista a classificação, atribuída pelos participantes, aos problemas encontrados durante a realização de cada tarefa. Será primeiramente, apresentada uma conclusão geral e, seguidamente, a análise individual de cada uma das tarefas.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T1	≠	≠	≠	•	•	•	≠	•	•	≠	≠	≠	•	•	8/14
T2	•	•	•	-	•	•	•	•	•	•	•	•	•	•	13/14
T3	•	•	•	•	•	•	•	•	•	•	•	•	•	•	14/14
T4	•	•	•	•	•	•	•	•	•	•	•	•	•	•	14/14
T5	•	⊖	⊖	•	≠	≠	≠	•	≠	•	≠	•	•	≠	8/14
T6	•	-	•	•	-	•	-	•	•	•	•	•	•	•	11/14
T7	•	•	•	•	•	•	•	•	•	•	•	•	•	•	14/14
T8	•	-	•	•	•	-	-	•	•	•	•	-	•	•	10/14
T9	•	•	•	•	•	•	•	•	•	•	•	•	•	•	14/14
T10	•	-	•	•	-	•	•	•	•	•	•	•	•	•	12/14
T11	•	-	•	•	•	•	•	•	•	•	•	•	•	•	13/14
T12	•	-	•	•	•	•	•	•	•	•	•	•	•	•	13/14
T13	•	•	•	•	•	•	•	•	•	•	•	•	•	•	14/14
T14	•	•	•	•	•	•	•	•	•	•	•	•	•	•	14/14

Tabela 6.1 Classificação do sucesso das tarefas realizadas pelos participantes.

Legenda:

- Tarefa executada com sucesso.

- Tarefa executada sem sucesso. Desistência.

☉ Teste não observado. Pela resposta deduz-se que a tarefa foi realizada com sucesso.

≠ Tarefa não cumpre com requisitos para o sucesso, mas é executada com sucesso parcial, face ao sucesso na obtenção da informação pedida no cenário.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T1	3	3	5	2	-	1	-	3	-	1	4	-	2	4	28/70
T2	2	4	4	5	2	1	-	-	4	3	-	2	-	2	29/70
T3	-	2	3	2	-	-	-	-	-	-	-	1	-	1	9/70
T4	-	3	5	2	-	3	-	-	-	-	4	2	2	4	25/70
T5	-	-	4	-	4	2	-	-	3	1	4	-	2	3	23/70
T6	2	5	4	-	2	3	1	4	4	2	3	2	3	3	38/70
T7	3	4	3	-	-	4	-	3	-	4	3	1	3	4	32/70
T8	1	4	4	5	-	5	5	2	-	4	3	2	1	2	38/70
T9	2	-	1	-	1	2	-	-	-	-	4	3	-	1	14/70
T10	4	4	2	-	5	1	1	-	4	1	3	5	1	2	33/70
T11	-	3	1	1	1	-	-	-	-	1	-	3	-	-	10/70
T12	4	3	1	-	-	1	-	-	-	-	4	-	2	1	16/70
T13	-	-	1	-	-	-	-	-	-	2	-	-	-	3	6/70
T14	-	-	1	-	-	-	-	-	-	-	-	-	-	-	1/70

Tabela 6.2 Classificação dos problemas encontrados pelos participantes.

Legenda:

- 1 Causa desconforto.

- 2 Causa muito desconforto.

- 3 Causa dúvidas.

- 4 Quase crítico para a realização da tarefa.

- 5 Crítico para a realização da tarefa.

- Não teve qualquer problema.

Estes primeiros testes foram muito importantes, não só para identificar problemas de usabilidade, mas, também, erros de código. Numa primeira análise generalizada, através das duas tabelas, podemos verificar que as tarefas em que os participantes tiveram menos dificuldade foram na T3, T9, T11, T12, T13 e, T14. Todas estas tarefas foram executadas com sucesso por todos os participantes, com exceção do P3 que não conseguiu concluir as tarefas T9, T11 e T12 com sucesso. Com estes resultados podemos concluir que, na maioria dos casos, à medida que o utilizador explora a visualização e toma conhecimento das funcionalidades e localização da informação, deixa de encontrar tantos problemas. Contudo, é necessário ter em atenção que o tempo de exploração necessário para amenizar problemas poderá ser demais para um utilizador normal. Pelo que deve ser encontrada uma solução para otimizar o tempo necessário para o utilizador perceber as funcionalidades e informações disponíveis.

As tarefas que levantaram mais problemas aos utilizadores foram a T6, T7, T8 e T10. A T6 e a T8 são as tarefas que tiveram mais desistências ou insucesso na execução. Estas são, também, as tarefas que punham à prova a usabilidade das funcionalidades mais interessantes da visualização das notícias – pesquisa de notícias por palavras e por tags. Conforme verificado, é necessário solucionar os problemas encontrados para permitir ao utilizador usufruir mais facilmente destas funcionalidades.

A T7, no entanto, é uma das tarefas que levantou mais problemas, tal como as tarefas anteriores, mas ao contrário destas, todos os participantes sem exceção conseguiram realizá-la com sucesso. O que acontece durante a execução desta tarefa é que os utilizadores detetam um problema de código – bugg – que os impede de clicar facilmente nos círculos das notícias, sendo que o raciocínio dos participantes é acertado, a funcionalidade testada é intuitiva, mas o código falha em servir os seus propósitos.

Podemos observar que a T1 e T5 foram as tarefas que os participantes conseguiram executar mas, com sucesso parcial, visto que na primeira identificam apenas um ano e, na segunda, não fazem uso do controlo da granularidade, perdendo mais tempo a fazer contas.

TAREFA 1

Identifique o(s) ano(s) em que o José Sócrates foi eleito Primeiro-Ministro.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T1	≠	≠	≠	•	•	•	≠	•	•	≠	≠	≠	•	•	8/14

Tabela 6.3 Classificação do sucesso da tarefa um (T1) realizada pelos participantes.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T1	3	3	5	2	-	1	-	3	-	1	4	-	2	4	28

Tabela 6.4 Classificação dos problemas encontrados durante a realização da tarefa um (T1) pelos participantes.

Todos os participantes conseguem ultrapassar a página de introdução e encontram o político José Sócrates na visualização das eleições, contudo, apenas pouco mais do que metade dos participantes concluem esta tarefa com total sucesso. Este resultado deve-se ao facto dos restantes participantes pararem de procurar o político em questão assim que o encontram pela primeira vez, ignorando o facto de que poderia existir mais um ano em que José Sócrates foi eleito para PM.

Um dos participantes – P3 – teve bastante dificuldade em perceber a visualização das eleições inicialmente, ultrapassando este problema apenas devido à exploração da visualização e comparando com os conhecimentos já adquiridos.

“Nos passarinhos como não tinha nenhum texto não percebi o que eram. Passando o rato por cima apercebi-me que estavam a aparecer nomes mas não estava a associar pois apareciam-me de outros políticos [para além do pedido]. Percebo a ideia de o pássaro maior ser o candidato eleito, mas não foi claro, pois só depois de passar o rato por cima e de ver os nomes é que me apercebi o que eram. Não associei as cores dos pássaros à cor política, só depois de estar a escrever este texto, numa segunda análise é que percebi quando vi os nomes com mais atenção.”

Observou-se, também, que alguns participantes – P6, P8 e P14 – tiveram alguma dificuldade em interpretar imediatamente a linha temporal, devido à forma como os anos são indicados (ex.: 76).

“Como o ano está registado só em metade do número tentei confirmar,

se era mesmo o ano que estava no meio e, na falta de outro elemento presumi que era.” (P6)

Para o P2, também, não foi claro que os políticos eleitos eram os pássaros que se encontravam mais perto da barra vertical, ou seja, estão ordenados de forma decrescente.

“Não é claro se Sócrates ganhou em 2002 ou 2005 porque em 2002 os pássaros têm praticamente o mesmo tamanho.”

O P5 apresentou um comportamento bastante interessante ao guiar o rato diretamente para os anos mais recentes e descobrindo muito rapidamente onde encontrava o político desejado. Este comportamento deveu-se à aplicação dos conhecimentos de política, facilitando a execução da tarefa com total sucesso.

Três dos participantes – P1, P8 e P11 – tentam, sem sucesso, fazer uso do botão de pesquisa nesta pergunta inicial para encontrarem rapidamente o político desejado. P1 e P11 classificam este problema como um problema relevante. *“Quería pesquisar, mas o botão não funcionava. Para quem não está muito dentro da política tem que se andar a ver os pássaros todos até encontrar o que queremos.”*

TAREFA 2

Identifique o partido com maior votação em 1979.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T2	•	•	•	-	•	•	•	•	•	•	•	•	•	•	13/14

Tabela 6.5 Classificação do sucesso da tarefa dois (T2) realizada pelos participantes.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Tota
T2	2	4	4	5	2	1	-	-	4	3	-	2	-	2	29

Tabela 6.6 Classificação dos problemas encontrados durante a realização da tarefa dois (T2) pelos participantes.

Praticamente todos os participantes finalizam esta tarefa com sucesso, uns com mais facilidade do que outros, apesar de concordarem que a localização da informação não se encontra num local intuitivo. O comportamento observado foi muito similar para todos os participantes: dirigem o rato para o ano pedido – 79 – e passam o rato por cima do político com maior votação. Com esta ação apenas conseguem descobrir o nome do político, pelo que, a maioria dos participantes seleciona o político. Quando se deparam com a visualização das notícias percebem que a informação desejada não se encontra aí e, portanto, voltam a desselecionar o político. Sem saberem onde procurar, voltam a colocar o rato sobre o passarinho maior e, quase por acidente, ao deslocar o rato descobrem que a informação se encontra no poleiro.

O P4 foi o único que desistiu da tarefa sem a concluir. O participante age tal como os outros participantes, dirigindo-se ao ano correto e identificando o político com maior votação, mas não encontra a informação pretendida. Dirige-se então ao botão de ajuda mas, também, não repara na informação que lhe é dada relativamente às

eleições/ poleiros. *“Encontrei o candidato com mais votos, mas não encontrei a qual partido pertence. Percebi qual era o botão de ajuda e utilizei-o, mas não era esse o local da informação pretendida.”*

O problema mais identificado foi, portanto, o facto de não saberem onde procurar a informação e de, esta, não se encontrar num sítio intuitivo. Os participantes esclarecem, também, que seria mais intuitivo se a informação aparece-se quando se clica ou passa o rato por cima do rato, ou seja, o impulso reside em procurar a informação no pássaro maior.

TAREFA 3

Identifique o partido correspondente a cada cor.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T3	•	•	•	•	•	•	•	•	•	•	•	•	•	•	14/14

Tabela 6.7 Classificação do sucesso da tarefa três (T3) realizada pelos participantes.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T2	2	4	4	5	2	1	-	-	4	3	-	2	-	2	29

Tabela 6.8 Classificação dos problemas encontrados durante a realização da tarefa três (T3) pelos participantes.

Face à tarefa pedida anteriormente, desta vez os participantes souberam imediatamente como executar esta tarefa. *“Dada a tarefa anterior, já compreendi que a cor do pássaro é a cor do partido e que o hover do poleiro dá os dados partidários, logo foi fácil.”* (P1)
 O resultado verificado é, assim, que todos concluem a tarefa com sucesso.

O problema identificado centrou-se na existência de cores muito parecidas, dificultando a distinção. *“Já agora, em 79, 83 e 85 há dois vermelhos, só olhando com muita atenção é que dá para perceber que são ligeiramente diferentes.”*

TAREFA 4

Identifique o trimestre em que Paulo Portas foi mais mencionado pelos *media*.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T4	•	•	•	•	•	•	•	•	•	•	•	•	•	•	14/14

Tabela 6.9 Classificação do sucesso da tarefa quatro (T4) realizada pelos participantes.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T4	-	3	5	2	-	3	-	-	-	-	4	2	2	4	25

Tabela 6.10 Classificação dos problemas encontrados durante a realização da tarefa quatro (T4) pelos participantes.

A maioria dos participantes observados – P4, P8, P9, P10, P12, P13, P14 – deparou-se como a visualização das notícias mesmo antes de chegar a esta quarta tarefa. Uma vez fora de contexto, alguns dos participantes ficaram surpresos com o aparecimento da nova visualização – a visualização das notícias. O P4 não percebeu imediatamente o que eram publicações (informação retirada do título da visualização) – *“Ok...Não sei o que são publicações trimestrais”* (P4). A informação tornou-se mais clara, para o P4, quando leu esta quarta tarefa associando as publicações às notícias. – *“Ahhh!...Acho que já percebi a parte dos trimestres e das publicações... tem a ver com as notícias... já faz mais sentido.”*

Para a maioria destes participantes que se depararam com a visualização das publicações mais cedo, a realização desta tarefa tornou-se mais fácil, uma vez que já sabiam onde ir buscar a informação.

Observou-se que os participantes que ainda não tinha descoberto a localização da visualização das notícias – P6, P7, P11 –, tiveram mais dificuldades na realização da tarefa, uma vez que não sabiam onde localizar a informação. – *“[Dificuldade em] Saber onde ir buscar a informação uma vez que estava escondida no pássaro correspondente.”* (P6)

“[Dificuldade em] Descobrir que existia um novo gráfico mais à direita. Algumas coisas não têm legendas, o que ajudava a entender melhor os gráficos. Só depois de clicar na ajuda é que descobri que dava para selecionar os políticos e obter informação relativa a este.” (P11)

Notou-se que o P14, apesar de ter tido contato com esta visualização anteriormente, também teve dificuldades na realização desta tarefa face ao cruzamento de informações. – *“[Dificuldade em] Separar a informação consultada anteriormente com a informação que pretendia consultar.”* (P14)

Curiosamente, verificou-se que tanto o P14, como o P6, P7 e P11 começaram a sua busca das notícias na visualização dos fatores socioeconómicos – a única visualização que estava à vista – esperando encontrar um tema que se relacionasse com as publicações.

Existiram ainda dois casos – P4 e P3 – em que os utilizadores tentaram fazer uso da pesquisa, sem sucesso, para encontrarem a informação pretendida.

“Procurei pelo Paulo Portas primeiro no menu superior, mas as opções não estavam ativas. Por isso procurei-o no mapa pela cor do partido, cheguei rapidamente ao pretendido. Foi fácil de perceber pelo tamanho dos círculos a quantidade de publicações.” (P4)

“Tinha-me focado nos ícones de topo, não me apercebi que podia clicar em cada um dos nomes. Sem querer descobri. Cliquei e o gráfico dos pássaros passaram para o lado esquerdo e aparece o novo gráfico do lado direito. Vendo que clicando nos nomes, a informação ia mudando descobrir como responder.”(P3)

Grande parte dos utilizadores detetaram dificuldades em manter o rato em cima dos círculos, ou notícias.

“Dificuldade em manter o rato sobre o texto para ver as datas.”(P13)

Este problema deve-se ao facto de o texto, que se encontra no centro da bola, interferir com a área de interação com o rato. Quando este problema apareceu, foi necessário ajudar o utilizador devido a tratar-se de um bugg e de o instinto do participante estar correto.

À parte das dificuldades encontradas, esta foi uma das poucas tarefas que todos os participantes conseguiram completar com sucesso.

TAREFA 5

Identifique qual dos dois políticos foi mais noticiado nos últimos três anos: Pedro Passos Coelho ou José Sócrates.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T5	•	e	e	•	≠	≠	≠	•	≠	•	≠	•	•	≠	8/14

Tabela 6.11 Classificação do sucesso da tarefa cinco (T5) realizada pelos participantes.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T5	-	-	4	-	4	2	-	-	3	1	4	-	2	3	23

Tabela 6.12 Classificação dos problemas encontrados durante a realização da tarefa cinco (T5) pelos participantes.

Pelo que se pode observar na tabela 6.11 esta tarefa foi uma das que todos os participantes conseguiram completar, ou seja, responder acertadamente, mas apenas cerca de metade conseguiram atingir os requisitos para o sucesso.

Todos os participantes descobrem que conseguem seleccionar e comparar dois políticos, no entanto, nem todos reparam nos botões que permitem alterar a granularidade.

A situação que decorre com o P5, P6, P7, P9, P11 e P14 – ou seja, maioria dos participantes observados – é que, para obterem a informação pedida, vão verificar todos os círculos com informação trimestral que correspondem aos últimos três anos – cerca de doze círculos. Por acaso, apenas num dos doze trimestres, o José Sócrates tem maior menção que o Pedro Passos Coelho, o que facilita a conclusão de que o Pedro Passos Coelho é o mais mencionado ao longo dos três anos.

A dificuldade mais identificada foi a perceção da granularidade com que as notícias estavam a ser organizadas – P6, P9 e P14 – sendo que a informação trimestral foi, inicialmente, confundida com informação anual. A recuperação deste erro foi rápida, uma vez que todos já tinham anteriormente analisado a informação trimestral e existia um título informando que eram as publicações trimestrais.

O P3 mostrou dificuldades na atribuição das notícias a cada um dos políticos dizendo que *“Reparei que apareciam dois círculos, um preto e um laranja mas não sabia a quem correspondiam. Não fazia sentido serem da cor do partido porque o laranja era o do José Sócrates. Era para ter vindo responder que não fazia ideia a quem correspondiam as cores, mas depois apercebi-me que estavam identificados do lado*

esquerdo no sítio onde seleccionei os pássaros. Quando cliquei neles não liguei à cor com que ficaram.”

O P5 e o P10 ficaram surpreendidos pela possibilidade de comparar os políticos, não estando à espera desta funcionalidade.

“Não estava à espera que desse para comparar, sendo que iria buscar os resultados individualmente e depois comparava. Apercebendo-me que essa feature existe, foi muito mais rápido de realizar a tarefa.” (P10)

“A informação de que pode ser possível seleccionar 2 políticos deveria encontrar-se algures na página.” (P5)

Existiu, também, alguma dificuldade na localização da visualização das notícias por parte do P11. – *“Foi difícil perceber que só depois de se clicar num político é que o gráfico das notícias aparece.”*

TAREFA 6

Identifique o semestre em que a palavra ‘submarinos’ foi mais empregue nas notícias sobre Durão Barroso.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T6	•	-	•	•	•	•	-	•	•	•	•	•	•	•	12/14

Tabela 6.13 Classificação do sucesso da tarefa seis (T6) realizada pelos participantes.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T6	2	5	4	-	2	3	1	4	4	2	3	2	3	3	38

Tabela 6.14 Classificação dos problemas encontrados durante a realização da tarefa seis (T6) pelos participantes.

A maioria dos participantes observados conseguiu concluir a tarefa pedida, no entanto, esta foi uma das tarefas que se mostrou mais problemática, mesmo para aqueles que a conseguiram concluir.

Observa-se pelas tabelas que um dos participantes que não conseguiu realizar a tarefa, classificou as suas dificuldades com um valor muito baixo, pelo que, esta tarefa deveria ter um valor total ainda mais alto.

Os principais problemas identificados residem na dificuldade em conseguir extrair a informação relativa ao número de notícias que mencionam “submarinos”, do gráfico apresentado.

“O gráfico deverá ser otimizado para se visualizar melhor a percentagem obtidos.” (P10)

Também é concluído que os botões de pesquisa deveriam ser maiores para chamarem mais a atenção. *“[Dificuldade em] Encontrar a opção que permitiria procurar por palavras, os ícones têm uma dimensão reduzida.”(P9)*

Verificou-se, também, através da observação, que uma minoria dos participantes teve alguma dificuldade em identificar a funcionalidade que lhe permitiria alterar a granularidade dos dados relativos às

notícias. “*Dificuldade em identificar a possibilidade de colocar o gráfico representando semestres.*” (P6)

TAREFA 7

Identifique o título da notícia mais recente de Francisco Sá Carneiro.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T7	•	•	•	•	•	•	•	•	•	•	•	•	•	•	14/14

Tabela 6.15 Classificação do sucesso da tarefa sete (T7) realizada pelos participantes.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T7	3	4	3	-	-	4	-	3	-	4	3	1	3	4	32

Tabela 6.16 Classificação dos problemas encontrados durante a realização da tarefa sete (T7) pelos participantes.

Esta tarefa foi uma das mais intuitivas para os participantes e todos conseguiram concluí-la com sucesso.

O problema identificado, cuja relevância é maior, registou a dificuldade dos participantes em clicar nos círculos para aceder às notícias. Este deve-se a um erro de código e, por isso, quando o raciocínio e intuição do participante estava certo, houve necessidade de interferir no teste para avisar que deveriam deslocar o rato um pouco, de modo a que não houvesse interferência com o texto.

“[Dificuldade] Em aceder às notícias, uma vez que carregando no número [texto] e na bola as notícias não apareciam.”

Um dos participantes (P14) mostrou, também, dificuldades em identificar o círculo de notícias mais recentes devido à pouca visibilidade da linha cronológica das notícias. “*Não é visível a linha cronológica tracejada no gráfico.*”

TAREFA 8

Identifique o título da notícia mais recente de Francisco Sá Carneiro.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T8	•	-	•	•	•	-	-	•	•	•	•	-	•	•	9/14

Tabela 6.17 Classificação do sucesso da tarefa oito (T8) realizada pelos participantes.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T8	1	4	4	5	-	5	5	2	-	4	3	2	1	2	38

Tabela 6.18 Classificação dos problemas encontrados durante a realização da tarefa oito (T8) pelos participantes.

A maioria dos participantes conseguiu realizar esta tarefa com sucesso, no entanto, quatro participantes não conseguiram chegar a uma conclusão. Esta foi uma das tarefas que levantou mais dúvidas e problemas, mesmo para os participantes que chegaram a uma conclusão acertada.

Dos participantes observados que não conseguiram realizar a tarefa – P6, P7 e P12 –, o primeiro participante (P6) chegou ao botão necessário mas não conseguiu perceber que a ordem das *tags* dadas se apresentavam por ordem decrescente. O segundo (P7) não conseguiu chegar ao botão desejado, pois tinha feito um *scroll* vertical que deixou os botões fora da parte visível do ecrã. E o terceiro (P12) chegou ao botão certo, mas não percebeu que a informação dada era aquela de que estava à procura, continuando a sua procura nas notícias dadas, clicando nos círculos.

Os principais problemas apontados pelos participantes foram: reconhecerem a ordem decrescente das palavras dadas e em encontrar o botão desejado. “*Não é claro que as palavras-chave estejam ordenadas por frequência mas assumo que sim, já que intuitivamente os 10s tópicos são populares para ela e não estão por ordem alfabética.*” (P1)

TAREFA 9

Identifique os diferentes temas representados na visualização dos fatores socioeconómicos (desde 1960 a 2015).

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T9	•	•	•	•	•	•	•	•	•	•	•	•	•	•	14/14

Tabela 6.19 Classificação do sucesso da tarefa nove (T9) realizada pelos participantes.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T9	2	-	1	-	1	2	-	-	-	-	4	3	-	1	14

Tabela 6.20 Classificação dos problemas encontrados durante a realização da tarefa nove (T9) pelos participantes.

Todos os participantes observados conseguiram concluir esta tarefa com sucesso, uma vez que, nos requisitos da tarefa não é obrigatória a ação de alterar o período temporal.

Observou-se que uma minoria dos participantes tomou atenção ao que era pedido entre parêntesis, pelo que apenas identificaram, sem problemas, os temas representados desde 1960 até 1976.

Os principais problemas identificados pelos participantes centram-se no facto das linhas da visualização serem muito finas, em perceber que poderiam controlar o período temporal do gráfico e no facto de terem de utilizar o *scroll* horizontal (P14) para verem a visualização dos fatores socioeconómicos. Esta utilização do *scroll* horizontal apenas foi feita por uma minoria dos utilizadores, pois todos os outros perceberam que ao desseleccionar os políticos voltavam ao ecrã inicial.

TAREFA 10

Identifique o ano em que o número de nascimentos atingiu o mínimo, relativamente a todos os outros anos (1960 a 2015).

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T10	•	-	•	•	-	•	•	•	•	•	•	•	•	•	12/14

Tabela 6.21 Classificação do sucesso da tarefa dez (T10) realizada pelos participantes.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T10	4	4	2	-	5	1	1	-	4	1	3	5	1	2	33

Tabela 6.22 Classificação dos problemas encontrados durante a realização da tarefa dez (T10) pelos participantes.

A maioria dos participantes conseguiu realizar esta tarefa com sucesso, sendo que apenas dois participantes desistiram da tarefa.

Observou-se que durante a sua realização, os participantes tiveram a necessidade de reler novamente a tarefa para perceberem que o período temporal da visualização não abrangia os anos pedidos na descrição da tarefa. Assim, a maioria dos participantes, após primeira tentativa de alterar o período temporal na própria visualização dos fatores socioeconómicos, levaram o rato até ao único sítio que tinha anos e carregaram ou arrastaram a barra que controla o período temporal da visualização.

“Chamou-me a atenção o facto de esta pergunta referir 2015. Tive de tentar descobrir como mudava as datas.

Apercebi-me então da cor diferente que aparece na barra do poleiro que dava para mudar.

A dificuldade foi perceber que conseguia mudar a data aí.”(P3)

Observou-se, também, que maioria dos participantes não sentiram necessidade de isolar/ selecionar o tema em análise, seguindo a linha correspondente com alguma dificuldade.

TAREFA 11

Identifique o número de indivíduos empregados em 2014.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T11	•	-	•	•	•	•	•	•	•	•	•	•	•	•	13/14

Tabela 6.23 Classificação do sucesso da tarefa onze (T11) realizada pelos participantes.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T11	-	3	1	1	1	-	-	-	-	1	-	3	-	-	10

Tabela 6.24 Classificação dos problemas encontrados durante a realização da tarefa onze (T11) pelos participantes.

Face à tarefa realizada anteriormente, a maioria dos participantes não teve grandes dificuldades em concluir esta tarefa com sucesso. Apenas um participante, continuou sem perceber como alterar o controlo temporal da visualização dos fatores socioeconómicos.

Os problemas anotados pelos participantes focam-se, assim, na dificuldade em encontrar o tema pretendido e em acertar nas linhas dos temas.

TAREFA 12

Identifique o(s) político(s) no cargo de Primeiro-Ministro durante os anos em que a dívida pública teve maior crescimento.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T12	•	-	•	•	•	•	•	•	•	•	•	•	•	•	13/14

Tabela 6.25 Classificação do sucesso da tarefa doze (T12) realizada pelos participantes.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T12	4	3	1	-	-	1	-	-	-	-	4	-	2	1	16

Tabela 6.26 Classificação dos problemas encontrados durante a realização da tarefa doze (T12) pelos participantes.

Também nesta tarefa, os participantes não encontraram problemas de maior. Existe um participante que continua sem conseguir identificar a barra que lhe permite controlar o período temporal e, por isso, não consegue concluir a tarefa.

Os problemas descritos focam-se maioritariamente no desconforto em ter de procurar manualmente as linhas dos temas pretendidos. “É chato ter de procurar os temas um a um.”(P13)

Outro problema consistiu “em ler corretamente a informação derivada da sobreposição de linhas e cores”(P6), face ao facto de o participante não ter percebido que poderia seleccionar o tema pretendido para individualizar a informação.

TAREFA 13

Identifique o botão que lhe permite ler sobre o projeto.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T13	•	•	•	•	•	•	•	•	•	•	•	•	•	•	14/14

Tabela 6.27 Classificação do sucesso da tarefa treze (T13) realizada pelos participantes.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T13	•	•	•	•	•	•	•	•	•	•	•	•	•	•	14/14

Tabela 6.28 Classificação dos problemas encontrados durante a realização da tarefa treze (T13) pelos participantes.

Todos os participantes conseguiram concluir esta tarefa com sucesso, não anotando problemas de maior. Os problemas encontrados residiram, essencialmente, na dimensão dos botões.

O P10 identificou, também, um problema com o botão das informações. Ao carregar no botão, este recolheu a informação, pois o participante já havia carregado neste botão anteriormente, evidenciando que terão de ser encontradas soluções para este problema. *“Quando clicado no botão certo não aparece à primeira vez.”*

TAREFA 14

Identifique o botão que lhe permite ler sobre o projeto.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T14	•	•	•	•	•	•	•	•	•	•	•	•	•	•	14/14

Tabela 6.29 Classificação do sucesso da tarefa catorze (T14) realizada pelos participantes.

Tarefas	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	Total
T14	-	-	1	-	-	-	-	-	-	-	-	-	-	-	1

Tabela 6.30 Classificação dos problemas encontrados durante a realização da tarefa catorze (T13) pelos participantes.

Todos os participantes observados conseguiram concluir esta tarefa com sucesso, respondendo com toda a certeza que a função seria “Ajuda”.

O problema apontado continua a residir na dimensão diminuta dos botões. *“Mais uma vez, botões pequenos.”* (P3)

6.6. Recomendações de Melhoramentos

Neste subcapítulo será descrita uma lista de melhorias recomendadas. Estas recomendações têm como objetivo solucionar ou minimizar os problemas de usabilidade encontrados durante a realização de testes, descritas anteriormente.

As melhorias serão aplicadas ao projeto consoante a importância e o tempo necessário para as resolver.

Mostrar *layer* de ajuda na primeira visita.

Pelo menos três dos participantes sugeriram que a ajuda fosse dada inicialmente, na primeira visita. A informação de todas as funcionalidades e da localização das visualizações seria uma grande vantagem para uma utilização que tira partido de todas as funcionalidades da visualização.

“Se calhar era útil abrir a ajuda logo na 1a visita, já que dá informação muito útil e a visualização é bastante original e beneficia dessa leitura.”(P1)

“O layer que aparece a indicar cada um dos pontos deveria aparecer na primeira visita.” (P3)

“Se calhar seria melhor se a ajuda fosse aberta logo no início.”(P11)

Permitir flexibilidade na alteração do período temporal.

Verificou-se que, conforme exploravam o controlo temporal da visualização dos fatores socioeconómicos, muitos dos participantes tentavam alterar o intervalo inferior.

“Uma vez que o poleiro vertical permite filtrar anos, talvez fosse interessante facilitar a filtragem de dados socioeconómicos pelo intervalo de tempo em que o dado governo esteve em funções.”(P1)

Filtrar políticos.

Para quem não tinha conhecimentos de política, este tipo de pesquisa tornou-se um processo mais lento e aborrecido, pelo que, foi sugerida a possibilidade de filtrar os políticos ou de existir uma lista selecionável de todos os nomes dos políticos representados.

Legendas para os partidos.

Apesar de algumas das cores serem facilmente associadas aos seus partidos, como é o caso do PS, PPD/PSD e CDS-PP, outras não dizem nada sobre o partido que representam. Assim, surgiu a sugestão de existir uma legenda visível com as cores ou até mesmo a possibilidade de filtrar os partidos.

Loading feedback.

Observou-se durante os testes que, por vezes, ao alterar a granularidade das notícias, a atualização do gráfico pode ser um pouco demorada. Para que o utilizador saiba que algo está a acontecer em função da sua ação deverá ser dado *feedback*.

Legendas tooltip para mouse over.

Alguns dos participantes sugeriram que fosse dada a informação das funcionalidades quando se passa o rato pelos botões.

Botões para fechar em vez de segundo click.

Durante as observações dos testes de usabilidade verificou-se que o segundo click para fechar as caixas relativas a cada botão é pouco intuitivo. Os participantes chegam a esta conclusão ao perceberem que não tem um botão para fechar/dispensar as *layers*. Contudo, ficam um pouco perdidos durante uns segundos.

Legendas selecionáveis para os temas.

Procurar pelos temas desejados na visualização dos fatores socioeconómicos foi, também, uma das tarefas mais morosa e aborrecida para os participantes. Assim, é sugerida a solução deste problema com legendas selecionáveis.

Setas laterais para a navegação horizontal.

Apesar de muitos participantes terem percebido como navegar pela visualização ao selecionar e desselecionar políticos, uma minoria dos participantes teve que utilizar o *scroll* horizontal que se encontra ao fundo da página. A posição em que se encontra faz com que nem sempre este esteja visível, dependendo da dimensão do ecrã, pelo que é sugerida a aplicação do *scroll* horizontal através de setas laterais.

Alteração do estilo ou local dos botões das notícias.

Muitos dos participantes dos testes de usabilidade tiveram alguma dificuldade em descobrir os botões de pesquisa. Também não foi perceptível, para os participantes, que estes botões se encontravam inicialmente inativos. Assim, sugere-se o aumento dos botões de pesquisa, o aumento da opacidade inicial e, eventualmente, uma separação dos botões gerais e dos botões das notícias (se assim se mostrar necessário).

Reparação do problema na seleção dos círculos das notícias.

Durante a realização dos testes, muitos participantes tiveram dificuldades em manter o rato sobre os círculos ou, até mesmo, clicar. Este problema deve-se ao facto de ser apresentada uma legenda textual quando o rato está em cima do círculo, que interfere com a área clicável do círculo. Para este problema será usada uma função que puxa as áreas clicáveis para a frente do texto, assim que este é adicionado à página.

Destaque dos temas na passagem do rato.

Vários participantes, ao realizarem a tarefa oito, tentaram seguir a linha correspondente ao tema pedido sem a individualização da mesma. Conclui-se que, este facto, deve-se à inexistência de destaque da linha quando o rato passa por cima da mesma. Espera-se com esta alteração que o utilizador perceba imediatamente que pode seleccionar a linha pretendida para uma visualização mais fácil e detalhada da informação relativa ao tema a consultar.

Conclusão

7

É incontestável que, hoje em dia, nos tornámos peritos colecionadores de informação. A evolução da internet e das tecnologias está na origem deste novo hábito que adquirimos, sendo cada vez mais fácil aceder à informação pretendida e, facilmente, guardá-la num dispositivo de reduzida dimensão física. Eventualmente, esta informação deixa de ser utilizada e, ao contrário do que acontece com a informação em suporte físico, que ocupa espaço e vai-se acumulando, a informação guardada em suporte digital fica, muitas vezes, esquecida. A área da visualização da informação traz um novo propósito a estas coleções de dados – esquecidas ou não – permitindo uma análise fácil dos mesmos e extrair novos conhecimentos, novas conclusões.

Este projeto pretende atingir esse mesmo objetivo – permitir a extração de novo conhecimento – através da visualização interativa de dados. Aborda, então, o campo da visualização de forma teórica e de forma prática.

A investigação para o estado da arte culminou numa extensão do conhecimento sobre a evolução das formas gráficas aplicadas à visualização e dos princípios de design que originam os melhores gráficos. A investigação estendeu-se, ainda, ao levantamento das primeiras técnicas interativas face à evolução tecnológica. Foram ainda recolhidos, explicados e analisados semióticamente alguns projetos relacionados com este projeto.

Após esta etapa inicial de pesquisa, deu-se início ao desenvolvimento prático do projeto, já nas instalações do SAPO. Após a comunicação do tema decidido para o projeto – Eleições –, decisão essa que coube ao SAPOLabs, foi necessária a definição dos dados que seriam interessantes tratar. Assim, começaram a ser levantadas questões não só eleitorais mas também relativas ao carácter dos políticos e à variação de fatores socioeconómicos.

Os primeiros esboços visuais dos dados eleitorais seguiram um caminho muito óbvio, tendo sido difícil o desvio para um caminho diferente. Contudo a utilização de um conceito visual baseado na expressão popular “Eles querem é poleiro” veio trazer uma nova dimensão ao projeto, influenciando a forma gráfica e transmitindo uma nova mensagem. O cuidado com a fácil leitura dos dados foi uma preocupação constante ao longo do projeto, bem como a aplicação de princípios de maximização de informação face à mancha gráfica utilizada. A implementação, o esboço visual de ideias e a conceção da interface gráfica foram processos que se influenciaram mutuamente e que, portanto, percorreram um caminho conjunto com durações semelhantes.

Ao contrário do que era pensado para a implementação, foi aconselhada uma nova linguagem tendo sido utilizado o *d3.js* ao invés do já conhecido *Processing.js*. Apesar da ambientação à nova linguagem ter sido uma tarefa com uma

duração razoavelmente demorada, concluiu-se que, à medida que a aprendizagem avançava, a utilização desta biblioteca apresentava diversas vantagens à execução do projeto. O desenvolvimento prático foi um processo que teve muitos avanços e recuos face a problemas emergentes de comunicação, de design ou de código. Os contratempos também tiveram lugar neste projeto, como a emergência de tarefas não planeadas, especialmente a necessidade de recolher dados políticos relativos aos nomes dos Primeiros-Ministros e dos Líderes de cada partido nos anos de eleições. Esta foi a tarefa que mais influenciou a forma visual do projeto, pois nem sempre foi fácil encontrar a informação pretendida e, esta, quando encontrada, nem sempre se dispunha num formato útil à extração de dados através de código, estendendo a duração desta tarefa face à necessidade de recolha manual dos dados.

O acesso aos dados em Portugal começou a caminhar recentemente em direção à Open Data, mas a pesquisa, recolha e tratamento de dados ainda é um dos passos mais demorados num projeto de visualização de informação, face aos problemas descritos anteriormente.

A realização de testes de usabilidade foi, também, uma tarefa que não correspondeu ao planeado face a limitações de tempo, resultante da execução de tarefas não planeadas. Consequentemente, durante a realização deste projeto, apenas uma fase de testes foi realizada pela aluna.

Concluiu-se com este projeto, que a escolha de uma forma visual para a representação dos dados pretendidos é, apenas, possível quando se conhecem os dados. Desta forma, a primeira tarefa na construção de uma visualização de informação, após definição dos dados que se pretendem tratar, deve ser a extração dos mesmos e respetiva conversão e análise. Só assim podem ser percebidos e escolhidos, os níveis dos dados que interessam representar. A aplicação desta técnica de filtragem aos dados das notícias, foi especialmente útil, uma vez que, cada pedido devolve muita informação prescindível. Durante a fase de ambientação ao *d3*, o tratamento e análise dos dados da Máquina do Tempo, tornou possível pôr de parte abordagens que evidenciavam a falta de dados. Isto veio comprovar a importância deste processo inicial, de aquisição, conversão e exploração dos dados.

Face à grande divulgação e atual utilização de gráficos de barras, gráficos tarte, entre outros mais comuns, a inovação na forma gráfica para representar dados, torna-se uma tarefa complexa. A tendência é fazer uso de uma forma gráfica comum, garantindo uma boa interpretação. A utilização de um conceito forte pode ser um dos passos que ajuda nesta inovação, contudo, a forma gráfica provou ser muito influenciada pelos dados e pelas limitações computacionais de quem constrói a visualização. Desta forma, com os mesmos dados é possível obter visualizações diferentes, cada uma com as suas vantagens e desvantagens. Porém, as melhores visualizações, aplicam regras já comprovadas eficientes, pelo que deve ter-se em atenção quando estas regras se aplicam ao

trabalho desenvolvido.

Constatou-se, que é também necessário conhecer o tema que se trata, a fim de se perceber o que cada nível dos dados significa. Neste caso, foi necessário conhecer o sistema do governo português, permitindo a relação com os fatores socioeconómicos e ajudando na definição dos dados.

Finalmente, os participantes dos testes de usabilidade classificam este projeto como sendo útil e bastante interessante – anexo D.

Trabalho futuro

Pretende-se divulgar esta visualização para que os utilizadores interessados possam fazer uma exploração dos dados tratados, especialmente nesta época em que se aproximam as eleições legislativas do ano corrente. Para que esta divulgação aconteça, devem ainda ser melhorados alguns problemas de usabilidade encontrados. Depois destes melhoramentos, o projeto deve ser aprovado pelas equipas Segurança e Usabilidade do SAPO.

Bibliografia

¹⁸Alexandre, D., & Tavares, J. (2007). Fatores da Percepção Visual Humana na Visualização de Dados. *CMNE 2007-Congresso de Métodos ...* Retrieved from [http://luizrodrigues.com/artigos-tcc/Information Visualization/Percep%E7%E3o humana na infovis.pdf](http://luizrodrigues.com/artigos-tcc/Information%20Visualization/Percep%E7%E3o%20humana%20na%20infovis.pdf)

⁹Backus, J., Beeber, R., Best, S., Goldberg, R., Haibt, L., Herrick, H., ... Nutt, R. (1957). The FORTRAN automatic coding system. In *Western Joint Computer Conference*. Retrieved from <http://dl.acm.org/citation.cfm?id=1455599>

⁸Bertin, J. (2011). *Semiology of Graphics: diagrams, networks, maps* (First Edit.). Esri Press.

¹Chen, C. (2006). *Information visualization: Beyond the horizon* (Second Edi.). Springer. Retrieved from [http://www.crisismanagement.com.cn/templates/blue/down_list/llzt_dsj/Information Visualization.pdf](http://www.crisismanagement.com.cn/templates/blue/down_list/llzt_dsj/Information%20Visualization.pdf)

¹⁰Chernoff, H. (1973). The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association*, 68(342), 361–368. doi:10.1080/01621459.1973.10482434

²¹Comissão Nacional de Eleições. (n.d.). Retrieved April 10, 2015, from eleicoes.cne.pt

²²Comissão Nacional de Eleições - Assembleia da República. (n.d.). Retrieved April 10, 2015, from <http://cne.pt/content/assembleia-da-republica>

⁵Cruz, P. (2010). *Boundaries in information visualization - towards information aesthetics*. FCTUC.

¹⁵Elmqvist, N., Moere, A., Jetter, H., Cernea, D., Jankun-Kelly, T., & Reiterer, H. (2011). Fluid interaction for information visualization. *Information ...*, (Section 2), 1–16. Retrieved from <http://ivi.sagepub.com/content/early/2011/08/19/1473871611413180.abstract>

²Friendly, M. (2006). *A Brief History of Data Visualization* (p. 43). Toronto. Retrieved from <http://www.datavis.ca/papers/hbook.pdf>

²⁵Fry, B. (2008). *Visualizing data* (Second Edi). O'Reilly Media. Retrieved from <http://dl.acm.org/citation.cfm?id=529269>

¹²Furnas, G. W. (1986). Generalized fisheye views. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '86*, 16–23. doi:10.1145/22627.22342

²⁷Havre, S., Hetzler, B., & Nowell, L. (2000). ThemeRiver: visualizing theme changes over time. *IEEE Symposium on Information Visualization 2000. INFOVIS 2000. Proceedings*. <http://doi.org/10.1109/INFOVIS.2000.885098>

²⁶Jog, N., & Shneiderman, B. (1995). Starfield visualization with interactive smooth zooming. *Visual Database Systems* 3, 1–10. Retrieved from http://link.springer.com/chapter/10.1007/978-0-387-34905-3_1

- ⁴Meirelles, I. (2013). *Design for Information: An introduction to the histories, theories, and best practises behind effective information visulizations*. Rockport Publishers.
- ¹⁴Murray, S. (2012). *Interactive Data Visualization for the Web* (p. 250). O'Reilly Media. Retrieved from <https://www.ebooks-it.net/ebook/interactive-data-visualization-for-the-web>
- ²⁴Nielsen, J. (2014). Turn User Goals into Task Scenarios for Usability Testing. Retrieved August 12, 2015, from <http://www.nngroup.com/articles/task-scenarios-usability-testing/>
- ²⁰Nielsen, J. (2000). Why you only need to test with 5 users. Retrieved January 25, 2015, from <http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>
- ²³Nielsen, J. (2012). Usability 101: Introduction to Usability. Retrieved August 11, 2015, from <http://www.nngroup.com/articles/usability-101-introduction-to-usability/>
- ¹⁷Shneiderman, B. (1983). Direct manipulation: A step beyond programming languages, 57–69. Retrieved from <http://dl.acm.org/citation.cfm?id=810991>
- ¹⁶Shneiderman, B. (1996). The eyes have it: a task by data type taxonomy for information visualizations. *Proceedings 1996 IEEE Symposium on Visual Languages*, 336–343. doi:10.1109/VL.1996.545307
- ¹¹Spinelli, Joseph G., Zhou, Y. (2004). Mapping Quality of Life with Chernoff Faces. In 2004 *EdUC Proceedings* (pp. 1–7). Retrieved from <http://proceedings.esri.com/library/userconf/educ04/papers/pap5000.pdf>
- ⁶TRANSPORT FOR LONDON. (n.d.). Harry Beck's Tube map - Transport for London. Retrieved December 01, 2014, from <https://www.tfl.gov.uk/corporate/about-tfl/culture-and-heritage/art-and-design/harry-becks-tube-map>
- ³Tuffte, E. R. (1995). *The Visual Display of Quantitive Information*. United States: Graphics Press.
- ¹⁹Ware, C. (2004). *Information Visualization: Perception for Design* (Second.). Morgan Kaufmann Publishers.
- ¹³Yi, J., Kang, Y. ah, Stasko, J., & Jacko, J. (2007). Toward a deeper understanding of the role of interaction in information visualization. *Visualization and Computer Graphics, IEEE Transactions*, 3(06), 1224 – 1231. doi:10.1109/TVCG.2007.70515

Bibliografia de imagens

ⁿ Aisch, G. (2011). 10 years Wikipedia Deutsche Welle. Retrieved January 26, 2015, from <http://visualdata.dw.de/en/wikipedia/>

^o Cable, D. (2013). The Racial Dot Map One Dot Per Person Weldon Cooper Center for Public Service. Retrieved January 26, 2015, from <http://www.coopercenter.org/demographics/Racial-Dot-Map>

[§] BBC News London. (n.d.). Retrieved January 23, 2015, from <http://www.bbc.com/news/uk-england-london-20943525>

^p Cruz, P., Costa, C., & Machado, P. *An Ecosystem of Corporate Politicians* Rangel, xCoax, Porto (2014), 146–156

^d Cruz, P. (2010). *Boundaries in information visualization - towards information aesthetics*. FCTUC.

ⁱ Cruz, P. (2012). Data lenses Visualization et al. Retrieved January 20, 2015, from <http://pmcruz.com/information-visualization/data-lenses>

^a Friendly, M. (2006). *A Brief History of Data Visualization* (p. 43). Toronto. Retrieved from <http://www.datavis.ca/papers/hbook.pdf>

^k Jussi Ängeslevä, & Cooper, R. (2005). Last Clock for iPad keeps you in factual time, human time and remote time. Retrieved from <http://www.creativeapplications.net/i-os/last-clock-ipad/>

^q Lúsica. (2014). Retrieved December 20, 2014, from <http://www.lasige.di.fc.ul.pt/webtools/fpsapo/music/lusica/about.html>

^m Mathew Bloch, Lee Byron, S. C. and A. C. (2008). The Ebb and Flow of Movies - Box Office Receipts 1986 — 2008 - Interactive Graphic - NYTimes. Retrieved January 26, 2015, from http://www.nytimes.com/interactive/2008/02/23/movies/20080223_REVENUE_GRAPHIC.html

^c Meirelles, I. (2013). *Design for Information: An introduction to the histories, theories, and best practises behind effective information visulizations*. Rockport Publishers.

^h Spinelli, Joseph G., Zhou, Y. (2004). Mapping Quality of Life with Chernoff Faces. In 2004 *EdUC Proceedings* (pp. 1–7). Retrieved from <http://proceedings.esri.com/library/userconf/educ04/papers/pap5000.pdf>

^e Sociology-Economics. (n.d.). Retrieved January 23, 2015, from http://libweb5.princeton.edu/visual_materials/maps/websites/thematic-maps/quantitative/sociology-economics/sociology-economics.html

^r SAPO Labs – Grande Área – estatísticas do mundial de futebol desde 1930 (n.d.). Retrieved January 20, 2015, from <http://labs.sapo.pt/2014/06/grande-area-estatisticas-do-mundial-de-futebol-desde-1930/> Teixeira, J. (2014b).

^sSAPO Labs – Grande Área – um repositório interativo de informação desportiva(n.d.). Retrieved January 20, 2015, from <http://labs.sapo.pt/2014/04/grande-area-um-repositorio-interativo-de-informacao-desportiva/>

^vSapo Labs, Universidade de Aveiro, Universidade da Beira Interior, & LED. (2013). Um país como nós - Autárquicas. Retrieved August 10, 2015, from <http://umpaiscomonos.labs.sapo.pt/Autarquicas/ResultadosDoPais/>

^tFerreira, R. (2013). TeK Notícias Internet A Máquina do Tempo do SAPO tem 25 anos de histórias em notícias. Retrieved January 20, 2015, from http://tek.sapo.pt/noticias/internet/a_maquina_do_tempo_do_sapo_tem_25_anos_de_his_1351165.html

^fTelling a story with data - Deloitte University Press. (n.d.). Retrieved January 23, 2015, from <http://dupress.com/articles/telling-a-story-with-data/>

^zThe Guardian. (2015). Election 2015 interactive: what did the opinion polls say about your seat? Retrieved March 2, 2015, from <http://www.theguardian.com/politics/ng-interactive/2015/apr/20/election-2015-constituency-map?src=nav>

^xThe Guardian. (2015). Election 2015 The Guardian poll projection. Retrieved March 2, 2015, from <http://www.theguardian.com/politics/ng-interactive/2015/feb/27/guardian-poll-projection>

^uThe New York Times (2008). Election Results 2008. Retrieved March 23, 2015, from <http://elections.nytimes.com/2008/results/president/map.html>

^bTufte, E. R. (1995). *The Visual Display of Quantitative Information* (Fouteenth.). United States: Graphics Press.

^lWattenberg, M. (2005). *Baby Name Wizard*. Retrieved January 26, 2015, from <http://www.babynamewizard.com/voyager#>

^jYi, J., Kang, Y. ah, Stasko, J., & Jacko, J. (2007). Toward a deeper understanding of the role of interaction in information visualization. *Visualization and Computer Graphics, IEEE Transactions*, 3(06), 1224 – 1231. doi:10.1109/TVCG.2007.70515



UNIVERSIDADE DE COIMBRA