

ESCOLA POLITÉCNICA
PROGRAMA DE PÓS GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO
MESTRADO EM CIÊNCIA DA COMPUTAÇÃO

EDUARDO GHIDINI

REPRESENTAÇÃO DE NARRATIVAS INTERATIVAS POR MEIO DE
VISUALIZAÇÕES COM DADOS EXTRAÍDOS DE REDES SOCIAIS

Porto Alegre
2017

PÓS-GRADUAÇÃO - *STRICTO SENSU*



Pontifícia Universidade Católica
do Rio Grande do Sul

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
FACULDADE DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO**

**REPRESENTAÇÃO DE
NARRATIVAS INTERATIVAS
POR MEIO DE VISUALIZAÇÕES
COM DADOS EXTRAÍDOS DE
REDES SOCIAIS**

EDUARDO GHIDINI

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Ciência da Computação na Pontifícia Universidade Católica do Rio Grande do Sul.

Orientador: Prof. Dra. Milene Selbach Silveira

**Porto Alegre
2017**

Ficha Catalográfica

G423r Ghidini, Eduardo

Representação de narrativas interativas por meio de visualizações com dados extraídos de redes sociais / Eduardo Ghidini . – 2017.

119.

Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência da Computação, PUCRS.

Orientadora: Profa. Dra. Milene Selbach Silveira.

1. IHC. 2. storytelling. 3. visualização de informação. 4. redes sociais. I. Silveira, Milene Selbach. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da PUCRS com os dados fornecidos pelo(a) autor(a).

Bibliotecária responsável: Salete Maria Sartori CRB-10/1363

Eduardo Ghidini

Representação de narrativas interativas por meio de visualizações com dados extraídos de redes sociais

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre em Ciência da Computação do Programa de Pós-Graduação em Ciência da Computação, Escola Politécnica da Pontifícia Universidade Católica do Rio Grande do Sul.

Aprovado em 29 de março de 2017.

BANCA EXAMINADORA:

Profa. Dra. Carla Maria Dal Sasso Freitas (UFRGS)

Prof. Dr. Márcio Sarroglia Pinho (PPGCC/PUCRS)

Profa. Dra. Milene Selbach Silveira(PPGCC/PUCRS – Orientador)

AGRADECIMENTOS

É impossível, no final desta etapa, não olhar pra trás e lembrar daqueles com quem pude contar, seja em momentos de descontração, seja nos momentos de dúvidas ou nos momentos que precisei de coragem de motivação para continuar nessa caminhada. Levo comigo o crescimento que tive nesses últimos dois anos, não apenas o conhecimento em uma área ou como pesquisador, mas, também, o crescimento pessoal e o amadurecimento que tive, sempre lembrando daqueles que estiveram no meu lado.

Primeiramente, agradeço a minha família, amigos e colegas, por dividirem momentos de lazer, de trabalho e por compreender os (muitos) necessários momentos de ausência e momentos de apreensão. Agradeço a uma colega em especial, Caroline Queiroz Santos, que, mesmo sem precisar, se disponibilizou, enquanto pode, a participar das minhas reuniões de orientação e ajudou a direcionar o trabalho.

Agradeço, também, a Faculdade de Informática da PUCRS, por permitir que eu fosse um aluno, desta vez, de pós-graduação. Agradeço, ainda, os participantes da pesquisa, que se disponibilizaram e foram cruciais para a conclusão da mesma.

Por fim, um agradecimento especial à minha orientadora Prof^o Dra. Milene Selbach Silveira, por ter acreditado em mim, por ter se disponibilizado, até mais do que necessário, e com quem tive o prazer de aprender muito.

REPRESENTAÇÃO DE NARRATIVAS INTERATIVAS POR MEIO DE VISUALIZAÇÕES COM DADOS EXTRAÍDOS DE REDES SOCIAIS

RESUMO

O crescimento de dados gerados com o uso da Internet, faz com que, também, aumente a dificuldade de fazer com que informações relevantes cheguem aos usuários. Para que isso aconteça, além do tratamento da informação, é necessário que esta informação seja apresentada de tal forma que o usuário consiga analisá-la e extrair informações úteis. Nos últimos anos, vem se explorando o uso de narrativas para a apresentação de grandes volumes de dados, considerando-se os benefícios desta forma de apresentação. Neste contexto, o presente trabalho objetiva analisar e propor formas de apresentar dados, em forma de narrativas, no contexto de redes sociais, a fim de apoiar a interação do usuário e sua análise dos dados. O trabalho apresenta um modelo para representação de narrativas por meio de visualizações, com dados extraídos de redes sociais, concebido a partir de estudos relacionados às áreas de narrativas e de visualização de informações. O modelo foi instanciado para que viabilizasse a sua análise por meio de estudos com usuários. Os estudos indicam que o modelo tem variadas aplicações, tanto no sentido de possibilitar a análise de diferentes tipos de dados, como, também, no que se refere ao uso por usuários com e sem interesse prévio em algum assunto específico. Através do modelo proposto e do resultado de sua análise, acredita-se que o presente trabalho possa apoiar a construção de narrativas interativas neste contexto, contemplando, não somente a construção da interface em si, mas, também, o mapeamento dos dados para a sua representação visual.

Palavras-Chave: IHC, *storytelling*, visualização de informação, redes sociais.

REPRESENTAÇÃO DE NARRATIVAS INTERATIVAS POR MEIO DE VISUALIZAÇÕES COM DADOS EXTRAÍDOS DE REDES SOCIAIS

ABSTRACT

The growth of data generated through the Internet increases the difficulty of gather relevant information for users. To minimize this, it is necessary to presented this information to the users in such a way that they can analyze it and extract useful information. In recent years, the use of narratives for the presentation of large volumes of data has been explored, considering its benefits. In this context, the present work aims to analyze and propose ways of presenting data, in narrative form, in the context of social networks, in order to support user interaction and data analysis. The work presents a model for the representation of narratives through visualizations, with data extracted from social networks, conceived from studies related to the areas of narratives and information visualization. The model was instantiated to enable its analysis through user studies. The studies indicated that the model has several applications, allowing the analysis of different types of data as well as its use by users with and without prior specific interests. We believe that the present work can support the construction of interactive narratives, through visualizations with data extracted from social networks, contemplating, not only the construction of the interface itself but, also, the mapping of the data to its visual representation.

Keywords: HCI, *storytelling*, information visualization, social network.

LISTA DE FIGURAS

Figura 2.1 – Representação da teoria Gestaltiana.	18
Figura 2.2 – Exemplos da lei de Gestalt. Fonte: [Ber06]	19
Figura 2.3 – Processo de percepção. Fonte: [Gol09].	20
Figura 2.4 – Variáveis visuais de Bertin: (a) tamanho; (b) orientação; (c) forma; (d) posicionamento; (e) matiz; (f) tonalidade da cor; (g) textura.	22
Figura 2.5 – Exemplos de utilização de cores em visualizações. Fonte: [Est06]. . .	22
Figura 2.6 – Exemplo de utilização de tamanho. Fonte: [Wor06].	23
Figura 2.7 – Mudança da orientação. Fonte: [War12].	23
Figura 2.8 – Apresentação por (a) cores e (b) padrões de orientação. Fonte: [Pen06].	23
Figura 2.9 – Representação de dados através de tabela.	24
Figura 2.10 – Exemplos de Gráficos para visualizações de Padrão.	25
Figura 2.11 – Exemplos de Gráficos para visualizações de Padrão.	25
Figura 2.12 – Exemplo de gráficos que representam relações.	26
Figura 2.13 – Gráfico de coordenadas paralelas. Fonte: [And06].	26
Figura 2.14 – Exemplo de mapa de calor. Fonte: [Kri06].	27
Figura 2.15 – Exemplo de utilização de <i>Treemaps</i> . Fonte: [NCo06].	27
Figura 2.16 – Exemplos de representações com grafos (a) e (b). Fonte: [Bos06b] [Bos06a].	28
Figura 2.17 – Exemplo de uso de mapas. Fontes: [Car06].	29
Figura 3.1 – Processo para transformação de dados em narrativas. Fonte: [LRIC15].	32
Figura 3.2 – Gêneros de Narrativa. Fonte: [SH10].	34
Figura 3.3 – Exemplos de <i>Magazine Style</i> . Fonte: [Ros06].	35
Figura 3.4 – Exemplos de <i>Annotated Chart</i> . Fonte: [Tim06b].	35
Figura 3.5 – Exemplo de <i>Partitioned Poster</i> . Fonte: [Gua06b].	35
Figura 3.6 – Exemplos de <i>Flow Chart</i> [(a) e (b)]. Fonte: [Tim06a] [IGS06].]	36
Figura 3.7 – Exemplo de <i>Comic Com</i> . Fonte: [BKH+16].	36
Figura 3.8 – Exemplo de <i>Slideshow</i> . Fonte: [Gua06a].	37
Figura 3.9 – Modelos Híbridos para narrativas. Fonte: [SH10].	38
Figura 3.10 – Exemplo de utilização da estrutura <i>Martini Glass</i> . Fonte: [Gua06a]. .	39
Figura 3.11 – Exemplo de utilização da estrutura <i>Interactive Slideshow</i> . Fonte: [BAMQ06].	40
Figura 3.12 – Exemplo de utilização da estrutura <i>Drill Down Story</i> . Fonte: [Pos09].	40

Figura 4.1 – Visualizações do CitySensing. Fonte: [BDVA+15].	47
Figura 4.2 – Exemplo do uso de StoryFlow. Fonte: [LWW+13].	48
Figura 4.3 – Visualizações do CodeTimeline. Fonte: [KS12].	48
Figura 5.1 – Exemplo de utilização do modelo.	51
Figura 5.2 – Exemplo de elementos a serem extraídos do modelo.	53
Figura 5.3 – Representação Gráfica do Gênero <i>Partitioned Poster</i> . Fonte: [GP01].	55
Figura 6.1 – Visão geral do Sistema desenvolvido. Na imagem: A - Mapa iterativo com a rota do evento; B - Gráfico de linhas com a relação entre data e volume de postagens; C - Tabela com os <i>tweets</i> ; D- Gráfico de Barras com volume de postagem por cidades; E - Gráfico de barras com as <i>hashtags</i> mais utilizadas; F- Gráfico de barras com os usuários mais aparentes; G - <i>Word cloud</i> com os usuários mais aparentes e; H - <i>Word cloud</i> com as <i>hashtags</i> mais aparentes.	67
Figura 6.2 – (a) Mapa com as rotas entre os estados e (b) Exemplo de visão do mapa detalhando o estado do Rio Grande do Sul.	68
Figura 6.3 – (a) Gráfico de Linha com a relação dos <i>tweets</i> x Data e (b) Filtro e detalhamento do Gráfico de Linha.	68
Figura 6.4 – (a) Gráfico de barras com as cidades com maior número de <i>tweets</i> e detalhamento e opção de filtro; (b) Gráfico de barras com as <i>hashtags</i> mais utilizadas e (c) Gráfico de barras com os usuários mais aparentes.	69
Figura 6.5 – (a) <i>Word cloud</i> com os usuários e (b) <i>Word cloud</i> com <i>hashtags</i>	69
Figura 6.6 – Tabela com os <i>tweets</i>	70
Figura 6.7 – Opções de Filtro.	70
Figura 7.1 – Representação gráfica da etapa de estudo com usuários.	73
Figura 7.2 – Representação gráfica da etapa de estudo com usuários após a execução do teste piloto.	74
Figura 7.3 – Gráfico de linhas após as alterações.	82
Figura 7.4 – Tabela com os <i>tweets</i> após as alterações.	82
Figura 7.5 – Tabela com os <i>tweets</i> após as alterações.	82
Figura 7.6 – Mapa com os detalhes e opção de limpeza das visualizações, após as alterações.	83
Figura 7.7 – Visão geral do sistema após as alterações.	84

LISTA DE TABELAS

Tabela 5.1 – Modelo proposto com os sub-componentes de cada componente . . .	51
Tabela 5.2 – Tabela com a relação entre os elementos, as informações para apresentação e formas de representá-los.	57
Tabela 5.3 – Sub-Componentes que compõe os elementos visuais com as respectivas técnicas.	58
Tabela 5.4 – Sub-componentes que compõe os elementos estruturais com as respectivas técnicas.	58
Tabela 6.1 – Técnicas de Elementos Visuais utilizadas.	64
Tabela 6.2 – Técnicas de Elementos Estruturais utilizadas.	64
Tabela 7.1 – Perfil dos participantes da primeira etapa do estudo	74
Tabela 7.2 – Resultado da Execução das Tarefas Específicas da Etapa 1.	75
Tabela 7.3 – Estrutura de Codificação Para o Código Interface.	77
Tabela 7.4 – Estrutura de codificação para o código Interatividade - Etapa 1.	79
Tabela 7.5 – Estrutura de codificação para o código Dados - Etapa 1.	81
Tabela 7.6 – Perfil dos participantes da segunda etapa do estudo	85
Tabela 7.7 – Estrutura de codificação para o código Interface - Etapa 2.	87
Tabela 7.8 – Estrutura de codificação para o código Interatividade - Etapa 2.	89
Tabela 7.9 – Estrutura de codificação para o código Dados - Etapa 2.	89

SUMÁRIO

1	INTRODUÇÃO	13
1.1	OBJETIVOS	14
1.2	METODOLOGIA	15
1.3	ORGANIZAÇÃO DO VOLUME	15
2	VISUALIZAÇÃO DE INFORMAÇÃO	16
2.1	PERCEPÇÃO HUMANA NA VISUALIZAÇÃO DE INFORMAÇÃO	17
2.2	VARIÁVEIS VISUAIS	21
2.3	TÉCNICAS DE VISUALIZAÇÃO	23
3	NARRATIVAS COM VISUALIZAÇÃO DE INFORMAÇÕES	30
3.1	ELEMENTOS DE NARRATIVAS	30
3.2	PROCESSO DE CRIAÇÃO	32
3.3	ESPAÇO DE DESIGN	33
4	TRABALHOS RELACIONADOS	41
4.1	PROCESSOS E ATIVIDADES	41
4.2	NARRATIVAS COM VISUALIZAÇÕES	46
5	MODELO PARA REPRESENTAÇÃO DE NARRATIVAS INTERATIVAS POR MEIO DE VISUALIZAÇÕES COM DADOS EXTRAÍDOS DE REDES SOCIAIS	50
5.1	DADOS	52
5.2	ELEMENTOS DA NARRATIVA	52
5.3	INFORMAÇÕES	53
5.4	GÊNERO	55
5.5	VISUALIZAÇÕES	56
5.6	ELEMENTOS VISUAIS	56
5.7	ELEMENTOS ESTRUTURAIS	58
6	#SEGUEATOCHA: INSTANCIAMENTO DO MODELO PROPOSTO	59
6.1	IMPLEMENTAÇÃO DA INTERFACE	59
6.1.1	DADOS	59
6.1.2	ELEMENTOS DE NARRATIVA	61

6.1.3	INFORMAÇÕES	61
6.1.4	GÊNERO	62
6.1.5	VISUALIZAÇÕES	62
6.1.6	ELEMENTOS VISUAIS	63
6.1.7	ELEMENTOS ESTRUTURAIS	63
6.2	#SEGUEATOCHA: INTERFACE PARA A NARRATIVA PROPOSTA	64
7	#SEGUEATOCHA: ANÁLISE DO MODELO PROPOSTO	71
7.1	PROCEDIMENTO	71
7.2	METODOLOGIA DE ANÁLISE	72
7.3	EXECUÇÃO DA PRIMEIRA ETAPA DOS ESTUDOS	73
7.3.1	EXECUÇÃO DAS TAREFAS ESPECÍFICAS	75
7.3.2	INTERFACE	76
7.3.3	INTERATIVIDADE	79
7.3.4	DADOS	80
7.4	EXECUÇÃO DA SEGUNDA ETAPA DOS ESTUDOS	81
7.4.1	EXECUÇÃO DA ANÁLISE DE UM ESTADO	85
7.4.2	INTERFACE	86
7.4.3	INTERATIVIDADE	88
7.4.4	DADOS	89
7.5	EXPLORANDO A NARRATIVA	90
7.6	DISCUSSÃO	91
8	CONCLUSÃO	94
	REFERÊNCIAS	97
	APÊNDICE A – Termo de Consentimento Livre e Esclarecido	104
	APÊNDICE B – Protocolo de Análise Original	105
	APÊNDICE C – Protocolo de Análise Final - Etapa 1	107
	APÊNDICE D – Protocolo de Análise Final - Etapa 2	109
	APÊNDICE E – Estrutura de Codificação Para o Código Interface - ETAPA 1 ..	111
	APÊNDICE F – Estrutura de Codificação Para o Código Interatividade - ETAPA	1113

APÊNDICE G	– Estrutura de Codificação Para o Código Dados - ETAPA 1	114
APÊNDICE H	– Estrutura de Codificação Para o Código Interface - ETAPA 2 . .	115
APÊNDICE I	– Estrutura de Codificação Para o Código Interatividade - ETAPA 2	118
APÊNDICE J	– Estrutura de Codificação Para o Código Dados - ETAPA 2	119

1. INTRODUÇÃO

Nos tempos atuais, vive-se o que é conhecido, por muitos, como a sociedade da informação, isto porque a informação e o conhecimento passaram a se tornar recursos estratégicos e são os agentes transformadores da sociedade [CC09].

Craine [Cra00] afirma que nos últimos trinta anos, produziu-se mais dados do que em toda a história da civilização. Assim, é natural que, com o aumento de dados gerados, encontrar informações relevantes se torne uma tarefa mais difícil. Com isso surgiu a área de Visualização de Informações, que objetiva, a partir de recursos gráficos, auxiliar a análise e compreensão de um, geralmente grande, conjunto de dados.

Neste âmbito, uma importante fonte de geração de dados é a Internet. Os dados provenientes da Internet são das mais variadas mídias (textos, imagens, vídeos ou sons) e das mais diversas fontes [RND15]. Redes sociais, por exemplo, são responsáveis por gerar uma enorme quantidade de dados [FAC16, Ins16, Twi16], e foram exploradas no contexto de visualização de informação em diversos trabalhos [WLY⁺14, ZCW⁺14, HB05, HFM07]. A análise das interações entre pessoas em redes sociais também tem sido utilizada, por exemplo, no jornalismo, pelo governo e no mundo dos negócios, o que atentou Weigend [Wei09] para o que é chamado de “revolução de dados sociais”. Isso porque elas contêm informações que podem ser úteis, se apresentadas de forma que provejam mecanismos de análises apropriados aos usuários.

Embora a quantidade de informação seja muito grande, nem toda a informação gerada pode ser relevante para quem a está analisando, e, neste sentido, a frase de Godin [Ros10] é muito pertinente: “Nós não temos uma escassez de informações, temos uma escassez de atenção”. A visualização de dados auxilia no direcionamento desta atenção, permitindo, com isso, que o leitor entenda estes dados e crie conhecimento a partir deles [GP01].

A área de visualização vem avançando ao longo dos anos e novas formas de representar dados estão surgindo, como, por exemplo, por meio de histórias. O ato de contar histórias é uma das mais antigas formas de comunicação e a disciplina de visualização de informações vem se beneficiando disto. Resumidamente, “narrativas com visualizações” é o ato de organizar as representações visuais de tal forma que o leitor estabeleça sequenciamento das informações por meio de relações temporais ou causais, permitindo, com isso, a experiência de entender a visualização como uma história [SH10]. Nos últimos anos, diversos trabalhos emergiram por meio da intersecção da área de visualização de informações e de narrativas [WR12, SH14, EAB13, KM13, SH10]. Conforme exposto por diversos autores, apresentar os dados em forma de narrativas possui diversos benefícios [WR12, SH14, EAB13], como, por exemplo, estabelecer uma conexão emocional e apoiar a retenção de informação por parte do leitor [Fig14]. Segundo Figueiras [Fig14], deve-se

entender as preferências da audiência da história (leitores) para que características de narrativas sejam melhor incorporadas com visualizações de dados. Diante desse panorama, cada vez mais se vê a necessidade de propor e avaliar novas formas de apresentar dados e prover maneiras para que a audiência explore estes dados de maneira efetiva. Para isso, ainda conforme a autora, estudos empíricos são necessários para que esta área avance.

Outros autores atentam para as oportunidades de estudos, no que diz respeito a narrativas com visualizações, como Kosara e Mackinlay [KM13], que indicam a necessidade de estudos em narrativas, com foco em abordagens adotadas para criação de histórias, métodos para avaliação destas narrativas, estudos sobre percepção e memorização, formas de interação com a história e técnicas de visualização que melhor se enquadrem em narrativas com visualizações. Segel e Heer [SH10] acreditam que devam haver estudos direcionados às experiências dos usuários, quando estes estão visualizando e interagindo com narrativas visuais.

A partir deste contexto, emergiu a seguinte questão de pesquisa: *"Como interfaces que utilizam narrativas com diferentes técnicas de visualização podem apoiar a interação do usuário e a sua análise de dados obtidos a partir de redes sociais?"*

A fim de refletir sobre esta questão foram definidos o objetivo geral e os objetivos específicos deste trabalho, bem como a metodologia a ser seguida para sua realização. Estes e a forma como este volume está organizado são apresentados nas próximas seções.

1.1 Objetivos

O principal objetivo deste trabalho é analisar como a utilização de técnicas de visualização, como apoio para a interação do usuário, com histórias contadas por meio de dados, pode auxiliar a audiência a analisar dados extraídos de redes sociais.

Para atingir tal objetivo, alguns objetivos específicos foram elencados:

- Realizar uma pesquisa aprofundada sobre técnicas de visualização e narrativas, bem como formas de utilizá-las em conjunto.
- Propor uma forma de extrair elementos de narrativas, a partir de postagens de redes sociais.
- Desenvolver uma forma de representação de dados que una narrativas e técnicas de visualização, com dados extraídos de redes sociais.
- Analisar as potencialidades desta forma de representação de dados, no que diz respeito a interatividade e ao apoio a análise de dados.

1.2 Metodologia

Para que os objetivos deste trabalho fossem alcançados, foi realizada, primeiramente, uma revisão bibliográfica na área de visualização de informações, narrativas e como a intersecção destas duas áreas têm sido trabalhada na comunidade científica. Com isso, foi possível verificar as técnicas referente ao *design* deste tipo de visualização, além de formas de representar visualizações para redes sociais. Em paralelo, estudou-se formas de mapear a estrutura de um *tweet* para os elementos de narrativa.

A partir dos estudos realizados, foi construído um modelo para visualização de informações extraídas de redes sociais, estruturadas em formas de narrativas, identificando seus componentes, bem como as técnicas relativas a cada um destes componentes. Após sua concepção, o modelo foi instanciado com dados obtidos do Twitter.

Por fim, o modelo foi analisado por meio de estudos com usuários. Estes estudos foram planejados, executados, e seus resultados foram discutidos com vistas à questão de pesquisa e aos objetivos que norteiam este trabalho.

1.3 Organização do Volume

O capítulos seguintes do volume abordam a fundamentação teórica do trabalho, apresentando, primeiramente, uma visão geral da área de Visualização de Informação (capítulo 2), com questões relacionadas à percepção (seção 2.1), variáveis visuais (2.2) e técnicas de visualização (2.3). Após, é apresentado o tópico de *storytelling*¹ em visualização de informação (capítulo 3), apresentando, além dos conceitos iniciais, elementos de narrativa (seção 3.1), o processo de criação de narrativas, na área de visualização de informação (seção 3.2) e o espaço de *design*, apresentando técnicas de interface para a criação de narrativas (seção 3.3).

Após, são apresentados alguns trabalhos relacionados (capítulo 4). Na seção 4.1 são apresentados trabalhos que mapearam processos e atividades para criação de narrativas na área de visualização. Após, na seção 4.2, são apresentados trabalhos que criaram narrativas com visualização.

Já no capítulo 5 é apresentado o modelo proposto, bem como cada componente existente no mesmo. No capítulo 6, é mostrada uma instância do modelo que foi utilizada para fins de sua análise. A execução da etapa da análise, bem como a discussão a respeito dos estudos realizados, é apresentada no capítulo 7. Por fim, no capítulo 8, são apresentadas a conclusão do trabalho, as limitações e possíveis trabalhos futuros.

¹os termos *storytelling* e narrativas são utilizados indistintamente ao longo do volume

2. VISUALIZAÇÃO DE INFORMAÇÃO

Para Kirk [Kir12], a Visualização de Informação (VI) é "a representação e a apresentação de dados que exploram as nossas habilidades de percepção visual, a fim de aumentar a cognição". Neste contexto, representação de dados diz respeito a forma visual em que os dados serão retratados (representados), podendo ser, por exemplo, a partir de linhas, gráficos, elipses ou mapas. A apresentação dos dados preocupa-se com questões relacionadas com o *design* (como as cores e a disposição dos conteúdos), com as opções e formas de interação que a interface oferece e com as informações adicionais providas na interface, que podem servir de auxílio para compreensão dos dados. Já a habilidade de percepção visual está relacionada com o entendimento científico de como nosso cérebro processa as informações de uma maneira mais efetiva, utilizando conhecimentos relacionados, como, por exemplo, ao raciocínio espacial e ao reconhecimento de padrões. Aumentar a cognição significa maximizar a eficiência e a eficácia no processo que transforma a informação recebida em conhecimento [Kir12].

O processo para criar representações visuais de dados, segundo Duce (1991) apud Campos et al. [COT97], consiste em três fases: acesso aos Dados, mapeamento e renderização. Acesso aos dados é a fase onde organiza-se e estrutura-se os dados (geralmente em um banco de dados), o mapeamento consiste em converter estes dados em uma representação gráfica, com suas respectivas propriedades (por exemplo, cores e texturas), e, por fim, a fase de renderização diz respeito a criação e exposição desta representação. Campos et al. [COT97] complementaram este processo, adicionando mais três fases: abstração, interação e estado de visualização. A abstração consiste em uma preparação semântica dos dados (como, por exemplo, uma filtragem ou preparação dos dados), sendo essa uma atividade anterior a fase de mapeamento. A interação diz respeito as opções de manipulação e exploração dos dados que serão providas para o usuário e a fase de estado de visualização preocupa-se em manter as informações relacionadas com as operações realizadas pelo usuário, para que seja mantido o estado do sistema.

Steele e Illinsky [SI10] consideram uma boa visualização aquela que é original, informativa, eficiente e que possui uma boa aparência estética. A originalidade mostra ao usuário uma nova perspectiva através de uma forma de apresentação inovadora dos dados, que pode proporcionar uma melhor forma de entendimento. Uma visualização eficiente é aquela que, além de possuir objetivos, mensagens e perspectivas claras, garante o acesso direto à informação e não apresenta informações irrelevantes que possam vir a confundir o usuário. Uma das características mais importantes, e que pode ser a chave do sucesso (ou falha) de uma visualização, é que ela deve ser informativa, pois, como citado pelos autores (e anteriormente nesta seção), o objetivo principal de qualquer visualização é prover acesso à informação para adquirir conhecimento – segundo os autores: "Uma visualização que não

atingiu este objetivo falhou". Por fim, a utilização adequada de elementos gráficos (como formas, cores, linhas e tipografia), além de garantir um nível estético satisfatório, ajudará o usuário a melhor entender e fazer conclusões a respeito dos dados apresentados.

Algumas das principais vantagens da VI, segundo Ware [War12], é que ela, além de permitir a compreensão de grandes volumes de dados, também permite a percepção de características que não são detectadas apenas com dados em sua forma bruta. Além disso, visualizações facilitam o entendimento de características, tanto com um grande volume de dados quanto com um volume de dados pequeno, e, com isso, facilitam a formulação de hipóteses. Por fim, visualizações podem auxiliar na identificação de problemas relacionados aos dados, sendo possível, através de uma visualização apropriada, detectar possíveis anomalias nos dados.

A partir das definições apresentadas, pode-se concluir que a área de VI tem, como um de seus principais objetivos, propor representações de dados que facilitem a sua compreensão, gerando, com isso, conhecimento para o indivíduo que está analisando esta visualização. As seções seguintes preocupam-se em prover um melhor detalhamento de alguns conceitos que envolvem a área de VI. Primeiramente, serão apresentados alguns fatores humanos que influenciam o usuário na percepção e no entendimento de visualizações. Após, será apresentado o detalhamento de algumas das variáveis visuais que compõem uma visualização e, então, algumas das principais técnicas de visualização utilizadas atualmente.

2.1 Percepção Humana na Visualização de Informação

A percepção humana se dá a partir dos sentidos. Estes são constantemente estimulados pois, a todo momento, os acontecimentos ao redor de uma pessoa podem mudar sua concepção em relação aos fatos. Os estímulos percebidos pelos sentidos geram impulsos elétricos que, através do sistema nervoso, geram sensações no cérebro, mudando, com isso, a maneira pela qual as pessoas percebem aquilo que está ao seu redor [BS06]. Seguindo a linha de raciocínio de que a percepção se dá a partir da interação com o ambiente, Sekuler et al. [BS06] afirmam que o mundo real e o que percebemos não são o mesmo, mas que estão fortemente relacionados. A relação se dá pelo fato de que o mundo percebido é proveniente do mundo real, mas é distinto deste pois podem haver deturpações em relação a realidade, ao passo que é possível que haja um aumento ou diminuição das informações captadas pelo observador.

A VI baseia-se, principalmente, no sentido humano que tem a maior capacidade de capturar e processar estímulos, a visão. A partir dela é possível construir relações entre espaço, luz, cor, forma e contraste [Kan03]. Segundo Ware [War12], o sistema visual humano, além de ser um investigador de padrões com grande poder e acuidade, também

forma "um poderoso centro de processamento paralelo que é capaz de fornecer um canal de banda larga para processos cognitivos humanos".

O uso de visualização de informações é influenciado, principalmente, por dois conjuntos de fatores, que são o subjetivo (que compreende, por exemplo, personalidade, experiência e habilidades) e o contextual (como motivação, expectativas e estresse) [Ber12]. Apesar das dificuldades de considerar todos estes fatores em uma VI, tendo em vista que são características inerentes de cada indivíduo, a forma pela qual os usuários compreendem visualizações de dados podem ser limitadas por processos perceptivos e cognitivos comuns (ou muito similares).

Uma das mais conhecidas teorias relacionadas à percepção, e que é adotada em diversas áreas de conhecimento, é a teoria da Gestalt [Ell99]. Essa teoria defende a ideia de que o todo é compreendido de maneira diferente que a soma de suas partes. Um exemplo deste fenômeno seria a interpretação de um carro. Segundo a visão Gestaltiana, a interpretação do objeto carro não é a mesma que a interpretação do conjunto de todas as partes que o compõem (como o motor, pneus, chassi, sistema de embreagem, etc.), pois o produto "Carro", tem características próprias. A Figura 2.1 mostra este princípio, na qual o quadrado resultante é um objeto diferente do que simplesmente a soma de todos os quadrados que o entornam.



Figura 2.1 – Representação da teoria Gestaltiana.

A teoria da Gestalt afirma que nosso cérebro tende a criar associações entre formas e cores para obter o maior número de significados possíveis. Para isso, baseia-se em alguns princípios [Ell99], que são:

- Proximidade: Objetos próximos tendem a ser percebidos como grupos, em relação aos objetos mais distantes. Na Figura 2.2 (a) cada conjunto de círculos agrupados tende a ser percebido como um grupo.
- Similaridade: Elementos que possuem características similares ou iguais (como cores ou formas) tendem a ser agrupados. Cabe ressaltar que o agrupamento por proximidade geralmente é o fator dominante em relação ao agrupamento por similaridade. Este princípio pode ser observado na Figura 2.2 (b), que mostra que os quadrados e os círculos tendem a ser percebidos como grupos distintos.

- **Pregnância (Simplicidade):** Os objetos tendem a ser percebidos em sua forma mais simples possível. Quanto mais simples, mais facilmente ela será assimilada. A Figura 2.2 (c) apresenta um exemplo deste fenômeno, em que o símbolo das olimpíadas tende a ser visto como cinco círculos e não como uma figura única e mais complexa.
- **Fechamento:** Objetos que possuem componentes faltantes em sua representação, quando possível, são complementados pelo nosso cérebro. A Figura 2.2 (d) apresenta um exemplo deste princípio.
- **Continuidade:** Direções e ligamentos contínuos entre os objetos tendem a ser melhor percebidos por causa da tendência do nosso cérebro em procurar padrões ou fluxos ao longo de uma mesma direção. A Figura 2.2 (e) exemplifica este princípio, na qual as linhas tendem a ser vistas como se estivessem em um fluxo contínuo.
- **Figura/Fundo:** A percepção pode ser dividida em duas partes, que são a figura e o fundo. Neste contexto, a figura representa o objeto a ser procurado, ou seja, aquilo que desejamos focar a atenção e o fundo é o contexto em que a figura está inserida. O elemento só é percebido quando há a separação entre estas duas partes. Isso pode ser percebido na Figura 2.2 (f) na qual a imagem mostra tanto uma leoa, quanto um gorila e uma árvore.

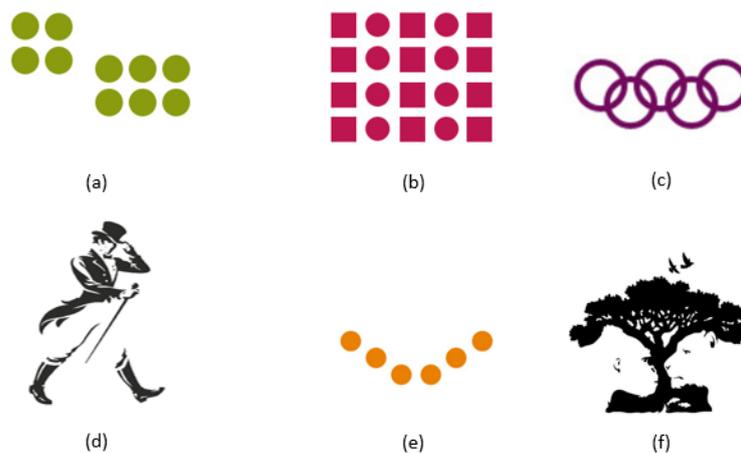


Figura 2.2 – Exemplos da lei de Gestalt. Fonte: [Ber06]

A despeito da importância dos princípios da Gestalt para o entendimento da percepção humana, estes não contemplam todas as etapas do processo da percepção, por se tratar de um processo bastante complexo e interativo. Nesse sentido, Goldstein [Gol09] apresenta um modelo de percepção (Figura 2.3) que é dividido em três grandes fases chamadas de estímulo, percepção, e experiência e ação. Cada fase possui um sub-processo, que contém mais três fases. Na fase de estímulo, o estímulo ambiental compreende todas as sensações possíveis de serem percebidas pelo indivíduo em um ambiente (independente

do foco de atenção). A atenção ao estímulo acontece quando o indivíduo focaliza a atenção em um ponto específico. A fase de estímulos nos receptores diz respeito a excitação dos sentidos para a captação destes estímulos (como, por exemplo, a observação de um objeto a qual resultará em uma imagem na retina).



Figura 2.3 – Processo de percepção. Fonte: [Gol09].

Tudo aquilo que percebemos baseia-se em sinais elétricos. Por isso a fase chamada de eletricidade compreende as etapas de transformação de estímulos em formas que sejam possíveis que o cérebro humano interprete. A transdução é a transformação de sinais químicos ou físicos em impulsos elétricos – isso ocorre, por exemplo, quando a luz visível (fóton) é transformada em impulso elétrico pelo nosso sistema visual. A transmissão ocorre no momento em que um neurônio transmite o pulso elétrico para outro, e, a comunicação entre um conjunto de neurônios, é denominada processamento.

Por fim, a etapa de experiência e ação começa pela percepção, que ocorre quando o indivíduo toma consciência do estímulo. A identificação do objeto caracteriza a fase de reconhecimento e, finalmente, as ações motoras, como movimentos da cabeça ou dos olhos caracterizam a última etapa do processo, chamada de ação. O autor [Gol09] ainda faz duas considerações importantes sobre este processo. Primeiramente, embora as etapas de percepção e reconhecimento sejam facilmente entendidas como uma única etapa, doenças, como a agnosia visual, mostram que são de fato fases distintas em nosso cérebro. Além disso, o autor cita que este é um processo interativo: isso porque a ação do indivíduo pode provocar (e geralmente provoca) o início de um novo ciclo deste processo.

Ware [War12] propõe um modelo de percepção visual simplificado, possuindo três etapas. Na primeira etapa, que, segundo o modelo, é um processamento do tipo *bottom-up* (quando a análise acontece a partir dos receptores sensoriais), acontece, por meio de um rápido processamento paralelo dos neurônios, a extração de propriedades em baixo nível

da imagem (como orientação, cor, textura e padrões de movimento). Este processamento paralelo ocorre através de uma ação involuntária do indivíduo. Na segunda etapa, o campo visual é dividido em regiões e padrões simples, como por algum contorno contínuo, por cores ou pela mesma textura. Nesta etapa ocorre interações entre processos *bottom-up* e *top-down* (quando o processamento das informações é guiado por processos mentais baseados em nossas experiências e expectativas). As principais características desta etapa são:

- processamento em série lento;
- envolvimento entre a memória operacional e a memória de longo prazo;
- diferentes caminhos para o reconhecimento de objetos e movimento visualmente guiado.

Por fim, na última etapa, denominada processamento sequencial dirigido, são construídas consultas visuais, que são respondidas por meio de estratégias visuais de busca, orientando, com isso, os objetos para os quais o usuário direcionará a atenção.

2.2 Variáveis Visuais

A transformação dos dados brutos em representações visuais pode ser uma tarefa complexa. Quando se objetiva representar os dados visualmente, devem ser considerados diversos aspectos, como, por exemplo, as variáveis visuais. Basicamente, as variáveis visuais são as diferentes características de um símbolo que podem ser percebidas.

Um dos conjuntos de variáveis mais conhecidos é o proposto por Bertin [Ber11]: tamanho, posicionamento, forma, orientação, matiz da cor, tonalidade da cor e textura. Neste contexto, o tamanho diz respeito ao espaço ocupado por um símbolo dentro de uma área e, o posicionamento, a sua localização. A forma é relativa às características físicas do símbolo. A orientação descreve a direção do símbolo, e, a textura, o aspecto da sua superfície. Por fim, as características relacionadas às cores, matiz e tonalidade, caracterizam, respectivamente, a distinção das cores em seus aspectos mais primários (como a distinção entre o vermelho e o azul) e a quantidade de luz que está presente na cor (determinando, com isso, a sua escala tonal). A Figura 2.4 apresenta a representação gráfica destas variáveis.

O uso das cores em visualização de dados não é apenas uma questão de estética, mas um importante instrumento para construção efetiva de visualizações. Segundo Lidwell et al. [LHB10], as cores podem ser utilizadas para atrair a atenção para algo, agrupar elementos, indicar significados, além de melhorar a apresentação estética dos elementos. O uso inadequado de cores pode atrapalhar e confundir o usuário [Sto06]. Ware [War12]

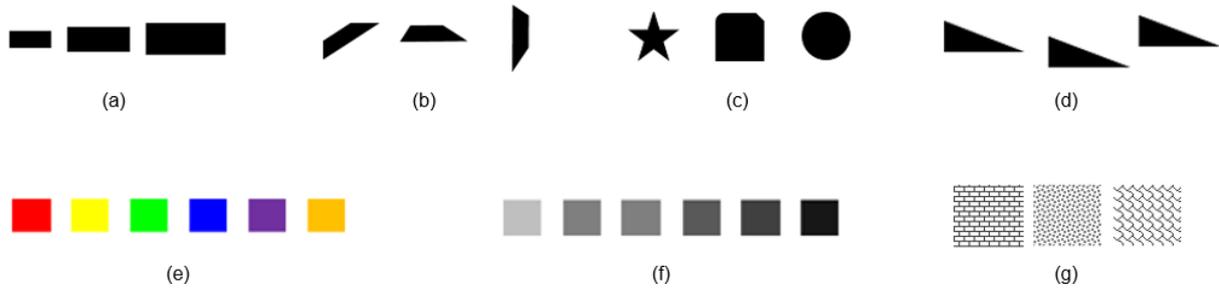


Figura 2.4 – Variáveis visuais de Bertin: (a) tamanho; (b) orientação; (c) forma; (d) posicionamento; (e) matiz; (f) tonalidade da cor; (g) textura.

afirma que devem ser utilizadas cores distintas apenas quando houver diferentes significados entre elas. Além disso, o autor afirma que não devem ser utilizadas mais do que 12 cores, pois pode ser que isto cause confusão no usuário. A Figura 2.5 apresenta um exemplo de como as cores podem ser exploradas em visualizações de informação.

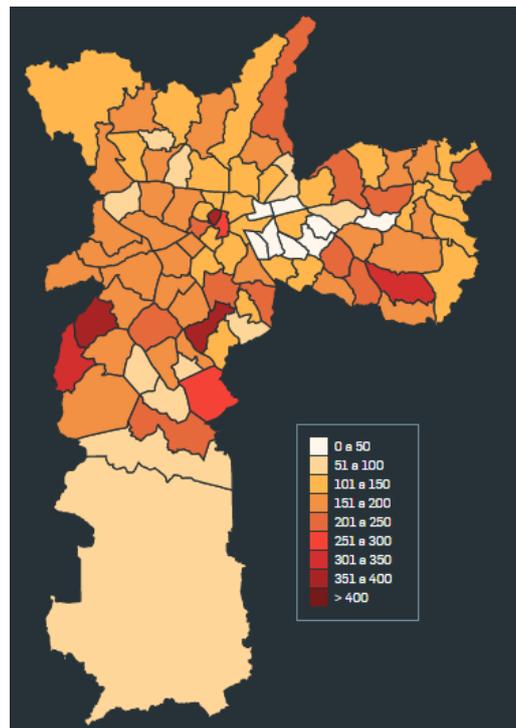


Figura 2.5 – Exemplos de utilização de cores em visualizações. Fonte: [Est06].

Outra variável, o tamanho, é uma das mais comumente utilizadas em visualização de informação [SI10]. O tamanho, quando utilizado para diferenciar objetos, é facilmente reconhecido, além de ser muito utilizado para dar noção de proporção e contrastar informações. A Figura 2.6 mostra um exemplo da utilização de tamanho para visualizar proporções (produções científicas no mundo).

Já a orientação é a variação do posicionamento horizontal e vertical de um símbolo gráfico, dando o aspecto do sentido da imagem. Deve-se ter claro que a alteração

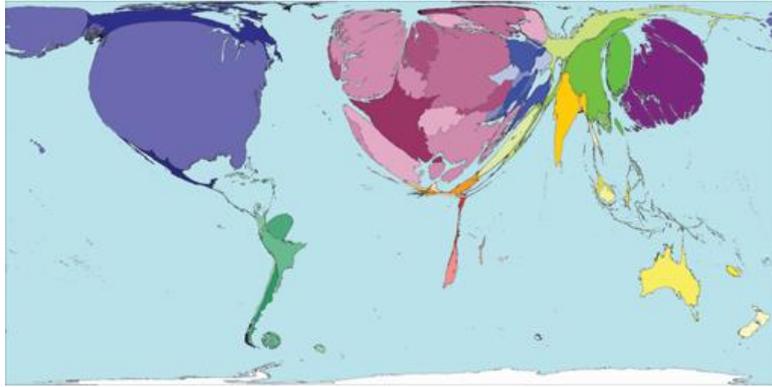


Figura 2.6 – Exemplo de utilização de tamanho. Fonte: [Wor06].

da orientação de um símbolo gráfico pode mudar a forma que este será percebido, como pode ser observado na Figura 2.7. A utilização de padrões de orientações pode ser uma alternativa em relação ao uso de cores, conforme mostrado na Figura 2.8.

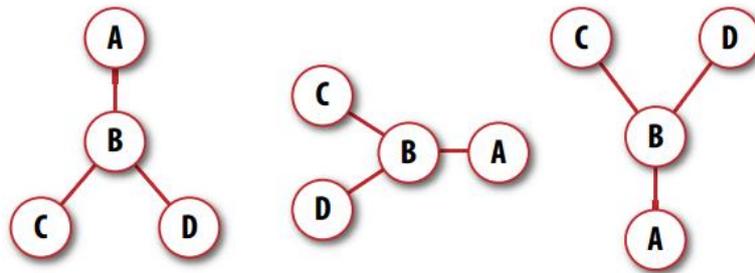


Figura 2.7 – Mudança da orientação. Fonte: [War12].

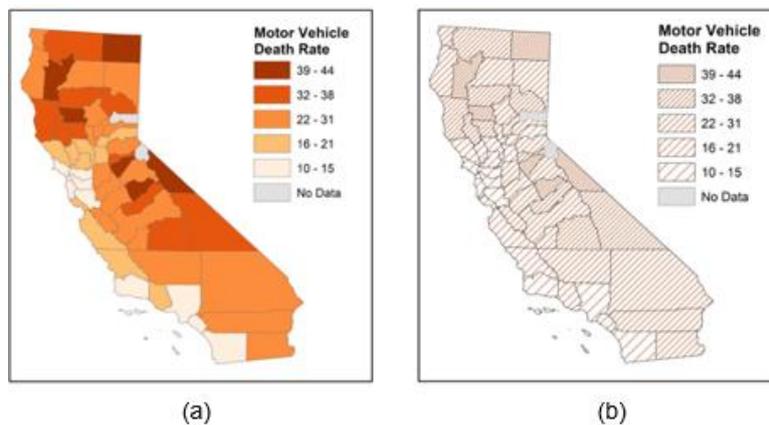


Figura 2.8 – Apresentação por (a) cores e (b) padrões de orientação. Fonte: [Pen06].

2.3 Técnicas de Visualização

A escolha da técnica de visualização a ser utilizada é uma das principais e mais importantes tarefas no processo de visualização. Isto porque a escolha adequada da vi-

sualização permite que grandes volumes de dados se tornem compreensíveis e inteligíveis aos usuários. Esta etapa deve levar em conta, principalmente, o tipo da informação a ser apresentado e as tarefas que serão realizadas pelo usuário [FCLC01].

Devido a grande variedade de técnicas de visualização existentes, o foco desta seção está nas técnicas que são mais comumente utilizadas. Como base, será utilizado o trabalho realizado por Khan e Khan [KK11], que mapearam e listaram diversas técnicas de visualização, e, também, outras técnicas encontradas na literatura.

Uma das técnicas de visualização mais convencional são as tabelas. Através delas é possível facilmente entender e interpretar os dados nelas contidos. Nesta estrutura, os dados são representados em linhas e é possível representar as relações através das colunas. A Figura 2.9 apresenta um exemplo de representação de dados em tabela.

Aparelho	Potência (KW)	Tempo de uso diário (horas)
Ar condicionado	1,5	8
Chuveiro elétrico	3,3	1/3
Freezer	0,2	10
Geladeira	0,35	10
Lâmpadas	0,10	6

Figura 2.9 – Representação de dados através de tabela.

Outra forma de representar dados muito utilizada são os gráficos. Dentre os tipos de gráficos citados pelos autores estão os que foram categorizados por Yau [Yau11] como sendo para visualizar padrões, proporções ou relações.

Quanto aos gráficos existentes para visualizar padrões, podem ser citados os gráficos de colunas, de dispersão e de linha (Figura 2.10). Estes gráficos apresentam dois eixos (x e y) e, por meio destes, é possível verificar tendências (é comum que no eixo x seja utilizada a variável tempo) e fazer comparações. No gráfico de colunas apresentado na Figura 2.10 (a), por exemplo, poderia ser representado o número de vendas de uma loja ao longo dos meses de um ano, sendo possível, por meio dele, fazer uma comparação das vendas através dos meses. Outra forma de verificar tendências e fazer comparações é por meio do gráfico de linhas apresentado na Figura 2.10 (c), que mostra a quantidade de itens vendidos de dois produtos através dos meses de julho e outubro.

Os gráficos que melhor permitem visualizar proporções são os gráficos de setores (também conhecidos como gráficos de pizza) e os gráficos de área. Os gráficos de proporção objetivam, principalmente, visualizar a distribuição das partes de um valor em relação às outras, que, juntas, formam um todo [Yau11]. Estes dois gráficos são bem distintos em suas representações: o gráfico de setores possui uma forma circular em que as proporções são visualizadas através do recorte do círculo. Os gráficos de áreas, por sua vez, possuem valores nos eixos x e y; geralmente, o eixo x representa dados temporais e o eixo

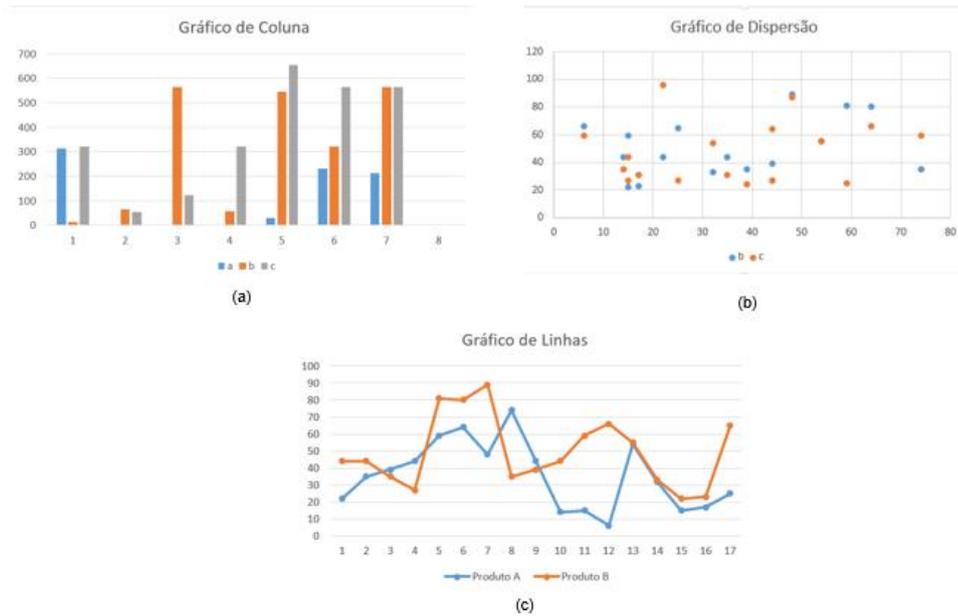


Figura 2.10 – Exemplos de Gráficos para visualizações de Padrão.

y apresenta alguma variável categórica. A Figura 2.11 apresenta exemplos de gráficos de proporção.

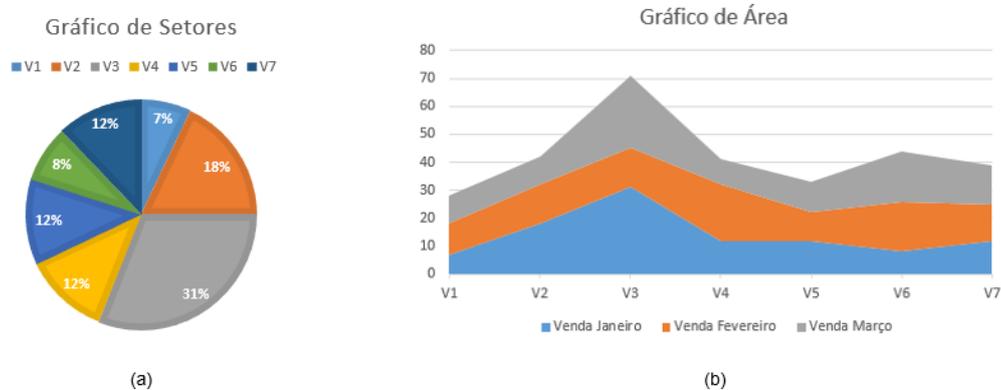


Figura 2.11 – Exemplos de Gráficos para visualizações de Padrão.

Gráficos de bolhas, ou histogramas, permitem visualizar relações não triviais, como, por exemplo, a relação entre a escolaridade e a criminalidade de um país, ou o tamanho das televisões ao longo dos anos. Os gráficos de bolha permitem a visualização de três variáveis, pois estes são compostos por um eixo x, um eixo y e, além disso, é possível visualizar a proporção de um valor no gráfico através do tamanho da bolha. Os histogramas são, visualmente, muito similares aos gráficos de colunas, com a diferença que, no eixo x, os valores dos intervalos são contínuos, e, no gráfico de colunas, os valores são discretos. A Figura 2.12 apresenta exemplos de representação dos gráficos de relação.

Os gráficos de coordenadas paralelas permitem visualizar dados multidimensionais. As dimensões, neste contexto, referem-se aos atributos que o gráfico possui. Estes

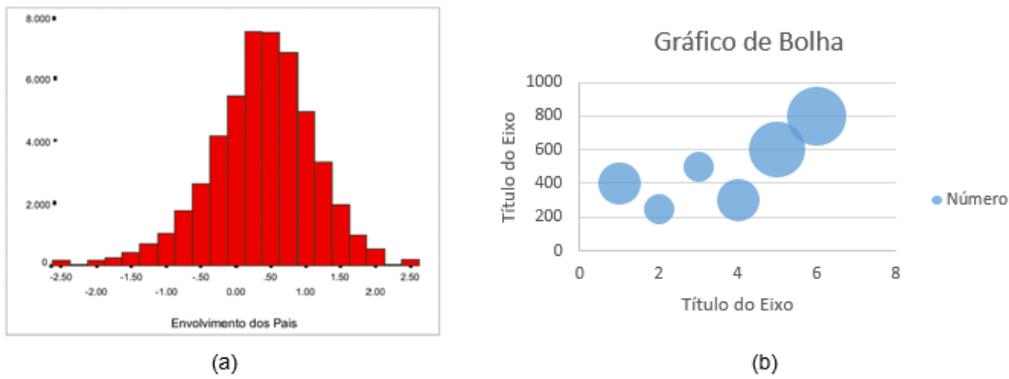


Figura 2.12 – Exemplo de gráficos que representam relações.

atributos são representados através de linhas verticais valoradas (geralmente quanto mais acima, maior é o seu valor) ao longo do gráfico. Cada linha representa uma unidade de registro que liga os atributos de acordo com o valor. Com isto, é possível identificar, além das relações, as dependências entre os atributos. A Figura 2.13 mostra um exemplo de uso de coordenadas paralelas.

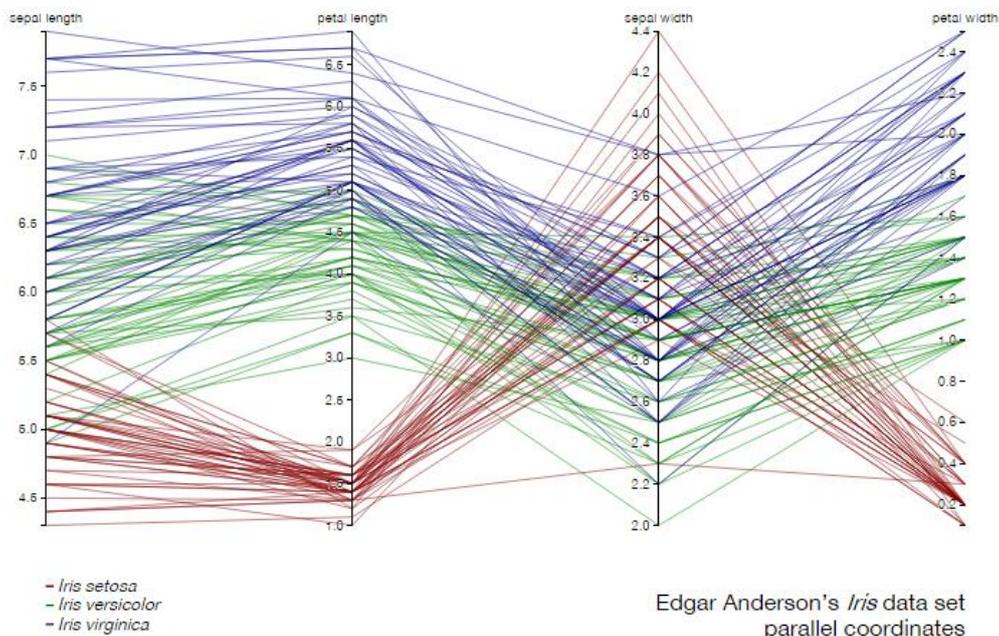


Figura 2.13 – Gráfico de coordenadas paralelas. Fonte: [And06].

Representados por retângulos, os mapas de calor (*heat maps*) apresentam os valores individuais de um conjunto de dados dispostos em uma matriz. Esta técnica utiliza cores para representar as grandezas dos valores. É comum que os valores mais altos possuam as cores mais escuras e os valores mais baixos possuam as cores mais claras. Com o mapa de calor é possível contrastar informações e comparar rapidamente uma grande quantidade de itens. A Figura 2.14 mostra um exemplo de uso de mapa de calor que apresenta a densidade de aniversariantes por dia e mês.

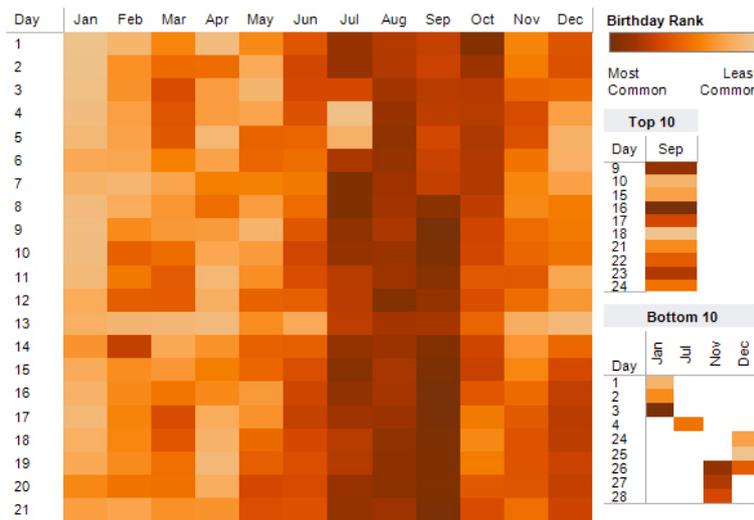


Figura 2.14 – Exemplo de mapa de calor. Fonte: [Kri06].

Outra forma de visualização muito utilizada é a baseada em grafos ou árvores. Elas são direcionadas, basicamente, à visualização de hierarquias e conexões. Dentre as diversas técnicas existentes, podem ser citadas, como exemplo, a *Treemappings* [JS91], o Browser Hiperbólico (*Hyperbolic Browser*) [HMM00] e a *Hierarchical Edge Bundles* [Hol06]. A *treemapping*, criada por Johnson e Shneiderman [JS91], é uma proposta de representação dos dados através de retângulos que possuem suas áreas baseadas em suas proporções dentro de um domínio. Cada retângulo possui seus “filhos” representados por retângulos que também possuem sua área proporcional à sua dimensão e são dispostas dentro do retângulo “pai”. É comum utilizar cores para diferenciar as categorias neste tipo de visualização. A Figura 2.15 mostra um exemplo desta técnica de visualização.



Figura 2.15 – Exemplo de utilização de *Treemaps*. Fonte: [NCo06].

Já na técnica Browser Hiperbólico [LR96] temos mapas conceituais baseados na técnica *Fisheye* (olho de peixe), que consiste em mostrar um nó principal no centro e os nós filhos em seu entorno. Isto faz com que seja possível observar um determinado ponto específico dos dados e ainda manter a noção do todo. Já a técnica *Hierarchical Edge*

Bundles mostra os dados em um círculo. Seus nodos ficam nas bordas deste círculo e as relações entre os nodos são mostradas em seu centro. Na Figura 2.16 são apresentados um exemplo do *Browser Hiperbólico* (a) e do *Hierarchical Edge Bundles* (b).

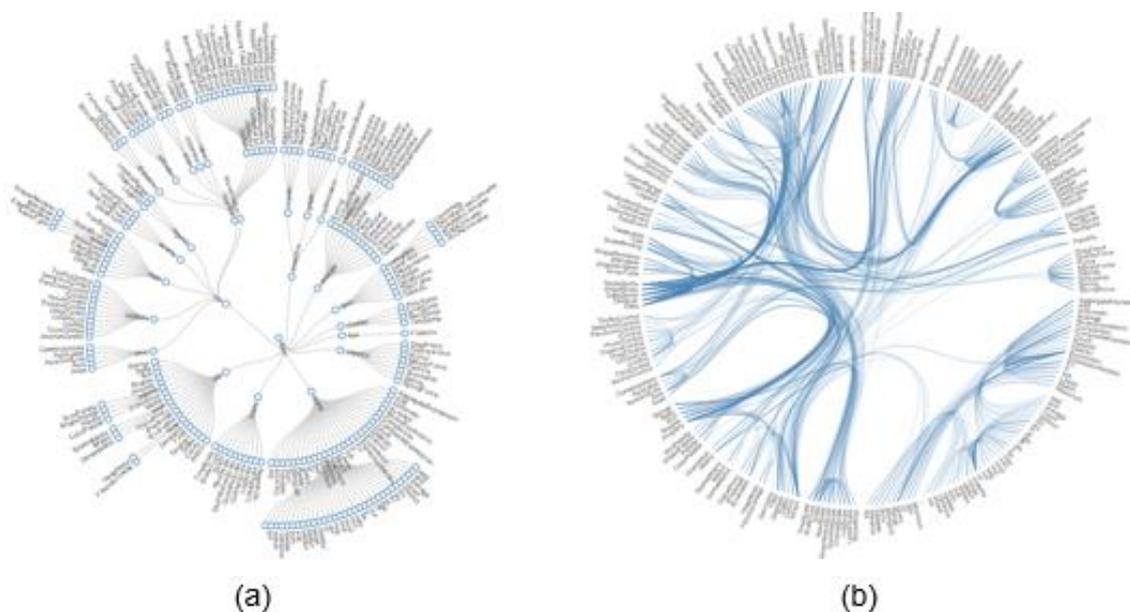


Figura 2.16 – Exemplos de representações com grafos (a) e (b). Fonte: [Bos06b] [Bos06a].

Representação de dados por mapas é outra forma bastante comum de visualizar dados. Segundo Yau [Yau11], “os mapas são uma subcategoria da visualização que têm o benefício de ser incrivelmente intuitivos”. O autor ainda afirma que, por serem uma representação do mundo físico com a qual, geralmente, temos contato desde cedo, o seu uso é de fácil entendimento para a maior parte das pessoas. As possibilidades que mapas oferecem são inúmeras. É possível, através deles, estabelecer relações, destacar regiões, visualizar proporções de diversos domínios de dados. Geralmente, utilizam-se outras características visuais para dar o foco desejado em mapas, como, por exemplo, formas geométricas (como bolhas ou pontos), cores para destacar regiões ou ícones que tenham significado em algum domínio (como quando se deseja destacar estabelecimentos ou ocorrências em uma área de uma cidade) [Nöl07]. A Figura 2.17 apresenta exemplos de visualização de dados em mapas.

A técnica de visualização utilizada é um fator determinante para a extração de conhecimento efetivo através de dados. Esta seção objetivou apresentar uma visão geral das técnicas existentes, bem como apresentar possibilidades de seu uso. Existem outras formas de apresentar dados visualmente, diferente das técnicas que foram apresentadas. Uma destas formas de representar dados é através de histórias, o que será apresentado no próximo capítulo.



Figura 2.17 – Exemplo de uso de mapas. Fontes: [Car06].

3. NARRATIVAS COM VISUALIZAÇÃO DE INFORMAÇÕES

Narrativas, no que se refere a visualização de informação, estão associadas ao termo *Storytelling*, que remete à expressão da língua inglesa “*tell a story*” (“contar uma história”). Uma história significa, resumidamente, uma sequência ordenada de fatos, que podem possuir palavras, imagens, visualizações, vídeos ou quaisquer combinações destes [KM13].

Massarolo [Mas13] afirma que as histórias sempre existiram e que é a partir delas que a humanidade transmite seu legado através das gerações. Segundo o autor, o ato de contar histórias remonta à própria história da humanidade, quando narradores sentados ao redor de fogueiras trocavam experiências entre si sobre os embates travados com as forças desconhecidas da natureza. A função do narrador era de ativar crenças e valores, fornecer sentidos aos acontecimentos através do relato de histórias. Esses relatos não se limitavam ao universo das experiências individuais, servindo para formular hipóteses e indagações sobre o nosso próprio destino.

Com o avanço da tecnologia, foi possível utilizar recursos computacionais para contar histórias por meio de sistemas interativos. Isso permitiu a expansão desta disciplina a outros contextos, como, por exemplo, o de visualização de informação. A arte, tecnologia e computação convergiram de tal forma que, hoje, é possível contar histórias através de dados. Para Gershon e Page, *storytelling* é uma forma muito efetiva de representar os dados pois permite que visualizações revelem informações tão efetiva e intuitivamente como se o espectador estivesse, por exemplo, assistindo a um filme [GP01]. Como os conceitos envolvidos nesta área são, até então, emergentes no campo de visualização de informação, serão apresentados, neste capítulo, elementos de narrativas, que são, resumidamente, as partes que compõem uma história; o processo de criação de histórias, no campo de visualização de informação; e, o espaço de *design*, que são os elementos presentes na interface, que, em conjunto, estruturam a visualização em forma de narrativa.

3.1 Elementos de Narrativas

Existem alguns elementos que compõem uma narrativa sem os quais ela não poderia existir. Segundo Pinna [Pin06], estes elementos são tema, enredo, tempo e narrador. No que tange ao tema, estão relacionados ainda os elementos assunto e mensagem. A diferença entre eles é que o tema é sobre o que a história trata, o assunto diz respeito a maneira pela qual o tema é abordado, e a mensagem compreende a conclusão ou o ensinamento que a história proporciona ao final da narrativa.

O conjunto de acontecimentos ordenados, os quais são protagonizados pelos personagens, é chamado de enredo. A evolução do enredo se dá a partir da organização dos eventos, que acontecem através do tempo e em algum lugar. Quando se conta uma história, é crucial que a causalidade dos eventos (influência que os eventos anteriores exercem sobre eventos posteriores) seja facilmente identificada pelos leitores [KM13]. O lugar compreende dois conceitos distintos: o primeiro é quando se refere a um lugar físico (como algum país, alguma cidade ou alguma casa, por exemplo) e é chamado de espaço; já o segundo, que é chamado de ambiente, objetiva situar os personagens nas condições as quais eles vivem (como características psicológicas, socioeconômicas, morais ou religiosas) [Gan04].

Por fim, o narrador é o elemento estruturador da história [Gan04] e, sem ele, não há como haver história. A história pode ter duas perspectivas diferentes no que se refere ao narrador: quando a história é contada em terceira pessoa ou quando a história é contada em primeira pessoa. Quando a história é contada em terceira pessoa, o narrador tem total conhecimento dos fatos e está em todos os lugares que envolvem a história (onisciência e onipresença da narrativa) e, geralmente, possui uma visão mais imparcial dos fatos. Quando a história é contada por um de seus personagens, a visão e os locais ficam limitados aos acontecimentos nos quais o narrador está presente.

No campo de visualização de informação, existem alguns conjuntos de elementos utilizados na literatura. Cruz e Machado [CM11a], baseando-se nos elementos de fábulas proposto por Bal [Bal09], apresentam uma série de elementos inerentes a narrativas de dados. Os elementos propostos por Bal são eventos, atores, tempo e localização. Embora Bal [Bal09] tenha elencado como elemento de uma fábula a localização, no trabalho proposto por Cruz e Machado [CM11a] este elemento não foi mapeado.

Eventos são definidos como a transição de um estado para outro, causados por atores [Bal09]. Atores, nesse sentido, são os personagens envolvidos e que, a partir das suas ações, geram as mudanças nas histórias. O tempo é o elemento que ordena e serializa a história. A ordem, neste contexto, é a ordem cronológica dos fatos. Por fim, a localização diz respeito ao local no qual os eventos ocorrem.

Outra forma de identificar os elementos de uma narrativa é através do *framework* 5W1H [YHYO01], que é um acrônimo para as palavras *where*, *why*, *what*, *when* e *who* (onde, porque, o que, quando e quem) e da palavra *how* (como). A sua proposta é que, respondendo estas perguntas, seja possível identificar todo o contexto de uma narrativa, como os participantes, a localização, o propósito, o momento, o evento em si e quais foram as circunstâncias que permitiram que o fato ocorra. Por se tratar de um *framework* teórico, este pode ser utilizados em diversos contextos (como, por exemplo, no jornalismo e no *marketing*).

Estes elementos são importantes em diversos níveis, pois, com eles, é possível identificar representações visuais que são baseadas em narrativas e, além disso, identificar

quais informações serão necessárias obter, a partir dos dados, para que tais narrativas sejam construídas.

3.2 Processo de Criação

O processo de criação de histórias, no contexto de visualização de informação, é uma tarefa complexa. Embora processos de criação de visualizações já tenham sido mapeados previamente, existem especificidades na criação de narrativas com visualizações que demandam um processo específico para este fim. Com este objetivo, a presente seção visa apresentar tal processo e uma visão geral de suas etapas.

A criação de histórias, em visualização de dados, requer a definição do contexto, seleção das informações, seleção da técnica de visualização e a escolha da ordem de apresentação dos dados para apresentar a narrativa [HDR⁺13]. Lee et al. [LRIC15] propõem um processo para transformar dados em histórias. Neste processo, é possível perceber as atividades envolvidas para explorar os dados, criar e contar as histórias. A representação visual do processo pode ser vista na Figura 3.1.

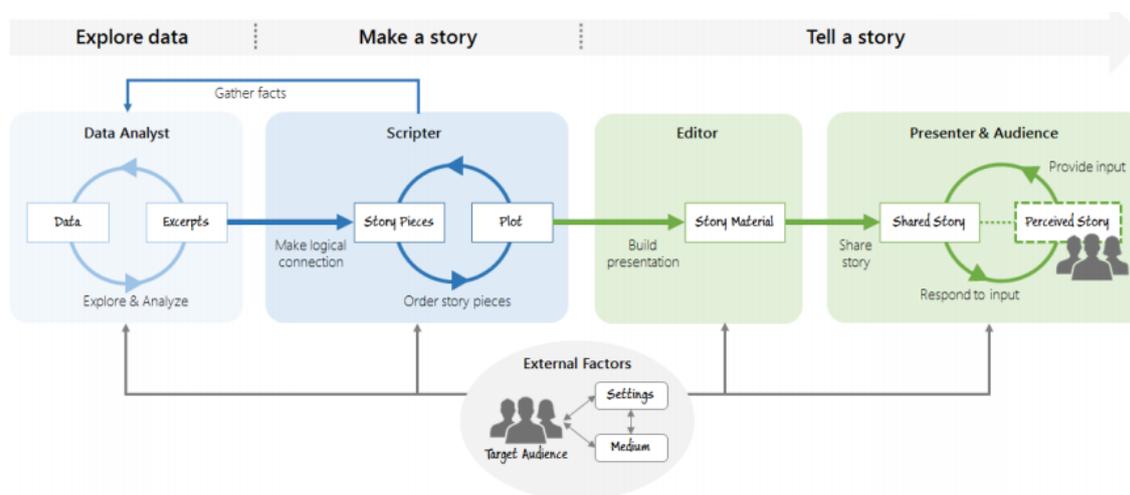


Figura 3.1 – Processo para transformação de dados em narrativas. Fonte: [LRIC15].

A exploração inicial dos dados tem por objetivo entendê-los para, com isso, criar fragmentos (resumos) de uma base de dados. Estes fragmentos podem compreender, por exemplo, um subconjunto de dados sequenciados ou dados procedentes de *insights* obtidos por meio da análise e podem ou não ser selecionados para uma representação visual já definida previamente.

A partir dos fragmentos obtidos na fase de análise de dados, cria-se a narrativa propriamente dita. Para isso, atividades envolvidas são ordenar os dados, estabelecer uma conexão lógica, desenvolver o fluxo dos dados, formular a mensagem a ser passada e criar o desfecho. Ao final desta etapa do processo, espera-se que seja possível decidir qual será

o enredo da história e como será a relação dos fragmentos (como, por exemplo, através do tempo, causa-efeito, padrões, entre outros).

Contar a história é a etapa em que, de fato, constrói-se a história, compartilha-se a história e recebe-se o *feedback* dos leitores. A construção da história (edição) consiste em desenvolver a representação visual e as interações da narrativa. É nesta etapa que se definem, além da forma de representação que será utilizada, os componentes gráficos que farão parte da representação visual, como textos de ajudas, animações, caminhos que são possíveis de serem percorridos, entre outros. A última etapa inicia-se a partir do momento em que a representação visual da história é compartilhada com alguém. A percepção da história é o que os leitores (ou a audiência) compreendem a partir da narrativa visual apresentada.

Ainda que este processo tenha sido estabelecido especificamente para criar narrativas de dados, ele é muito similar ao processo de criação de visualização (capítulo 2). Em ambos os processos, há o tratamento e a exploração dos dados disponíveis, a definição e o mapeamento dos dados para uma representação visual, a disponibilização para o usuário e, por fim, a percepção do usuário a respeito da representação visual proposta. A diferença deste para o processo de visualização é justamente o fato de que, este, contém etapas que só se aplicam ao contexto de narrativas.

3.3 Espaço de Design

Visualizações criadas com a intenção de contar histórias para sua audiência, necessitam incorporar características que lhe proporcionem um "tom de história". Na literatura, foram encontrados dois trabalhos [SLRS16, SH10] que, através de análise de um conjunto de visualizações, obtidas, principalmente, de portais de notícias, mapearam tais características. Através da identificação destas técnicas foi construído o que Segel e Heer chamaram de espaço de *design* [SH10].

Na presente seção, será apresentada uma visão geral do espaço de *design* proposto por Segel e Heer. Ainda, será apresentado o resultado da análise de uma série de trabalhos acadêmicos que utilizaram características de *storytelling* para apresentar dados, cujo o objetivo foi o de verificar as técnicas empregadas em trabalhos científicos, no que se refere ao *design* de histórias, no campo de visualização de informação¹. A análise dos trabalhos acadêmicos, bem como a análise do trabalho feito por Stolper et al. [SLRS16], re-

¹Os trabalhos foram obtidos por meio de uma revisão sistemática da literatura, aplicada em algumas das principais bibliotecas digitais (ACM digital library, IEEEExplore digital library, Elsevier ScienceDirect, Wiley and Google Scholar). Foram incluídos trabalhos que abrangessem - no resumo, título ou palavras-chaves - os termos (em inglês) "*story*", "*narrative*" ou "*storytelling*" e as palavras "*visualization*" ou "*visual*", e que estivessem na língua inglesa. A partir desta busca inicial, selecionou-se 71 trabalhos que continham estes critérios. Por fim, a partir da análise destes 71 trabalhos, selecionou-se 25 que apresentavam implementação de uma visualização com narrativa.

sultaram em uma extensão do espaço de *design* proposto por Segel e Heer, que, também, será apresentado nesta seção.

Segel e Heer [SH10], ao analisarem 58 visualizações que utilizam elementos de narrativas, propuseram um *framework*, cuja análise foi feita baseando-se em trabalhos relacionados a narrativas em filmes, literatura e arte. Desta análise, extraiu-se tal *framework*, que possui três categorias, que são os **gêneros**, os **elementos visuais** e os **elementos estruturais**.

Gêneros de narrativas são formas básicas de apresentar a história. Foram identificados sete gêneros distintos [SH10], que são *magazine style*, *annotated chart*, *partitioned pôster*, *flow chart*, *comic strip*, *slide show* e *film/vídeo/animation* (Figura 3.2). Estes gêneros variam em relação as suas características, como a quantidade de quadros/cenas e em relação a organização dos elementos visuais. O gênero *Magazine Style* é uma forma de representação de histórias cujo os dados são apresentados em um único quadro, como pode ser observado na Figura 3.3, que mostra a movimentação da bola durante jogos de futebol, vista de cima.

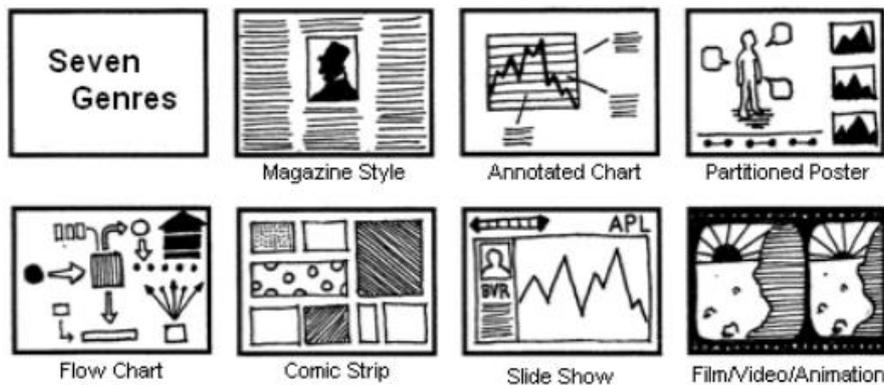


Figura 3.2 – Gêneros de Narrativa. Fonte: [SH10].

Annotated Chart apresenta algum gráfico (como, por exemplo, gráfico de linha, mapas, entre outros) com informações adicionais, contextualizadas, com o objetivo de prover ajuda ou explicações sobre algum ponto do gráfico. Esta forma de narrativa, geralmente, apresenta um único quadro. A Figura 3.4 apresenta um exemplo de *Annotated Chart*, que apresenta, com o uso de gráficos de linhas, a trajetória dos vinte e cinco atletas de baseball que mais marcaram pontuações (*Home Runs*), no que diz respeito a estas marcações. Já o gênero *Partitioned Poster* apresenta um pôster ou imagem particionada, cujas partes apresentam múltiplas visões da história. Este fenômeno pode ser visível ao analisar a Figura 3.5, que mostra os indicados ao Oscar de melhor fotografia. Nesta visualização, é possível selecionar cada indicado (parte) e visualizar informações relacionadas, como, por exemplo, outras indicações, *links* relacionados, imagens, entre outros) em todas as partes.

Flow Charts apresentam fluxos em uma forma visual, possuindo um ponto inicial e um ponto final. Neste tipo de visualização, é possível que haja diferentes “caminhos” a



Figura 3.3 – Exemplos de *Magazine Style*. Fonte: [Ros06].

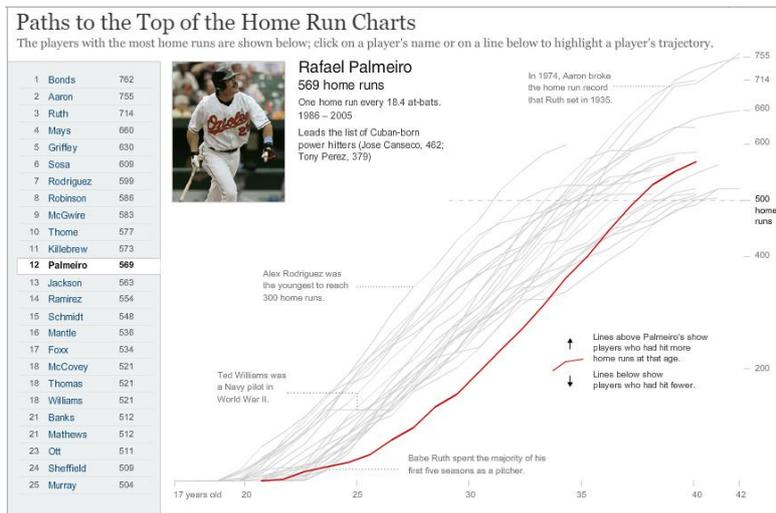


Figura 3.4 – Exemplos de *Annotated Chart*. Fonte: [Tim06b].



Figura 3.5 – Exemplo de *Partitioned Poster*. Fonte: [Gua06b].

serem seguidos. Embora o exemplo apresentado na Figura 3.6 (b) seja estático, é possível criar visualizações interativas, com fluxos únicos ou alternativos (Figura 3.6 (a)), cabendo ao leitor decidir o caminho que ele deseja seguir.

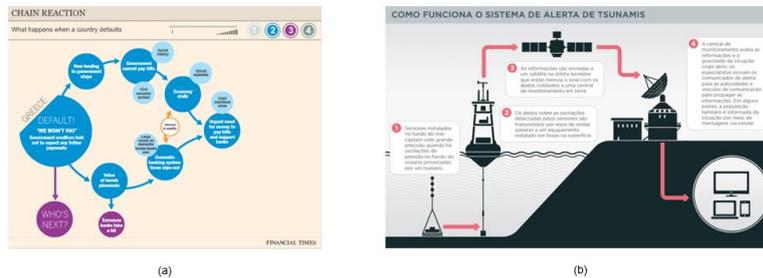


Figura 3.6 – Exemplos de *Flow Chart* [(a) e (b)]. Fonte: [Tim06a] [IGS06].]

As narrativas representadas pelo gênero *Comic Strip* (Figura 3.7) possuem fluxo contínuo, direcionado e com múltiplos quadros/cenas, que possuem relação causal direta entre as cenas [BKH⁺16]. *Slideshow* é uma forma de apresentação de dados de forma sequencial. A história, neste gênero, possui múltiplos quadros, representados por slides. Na Figura 3.8 é apresentado um exemplo deste gênero; nele, é possível observar, no canto inferior esquerdo, um botão cujo objetivo é avançar para o próximo slide (neste exemplo, não é disponibilizada a opção para que o leitor volte ao slide anterior). Por fim, o gênero *Film/Video/Animation* narra a história de forma totalmente direcionada e pré-estabelecida. Este gênero é muito utilizado em televisão e comerciais.

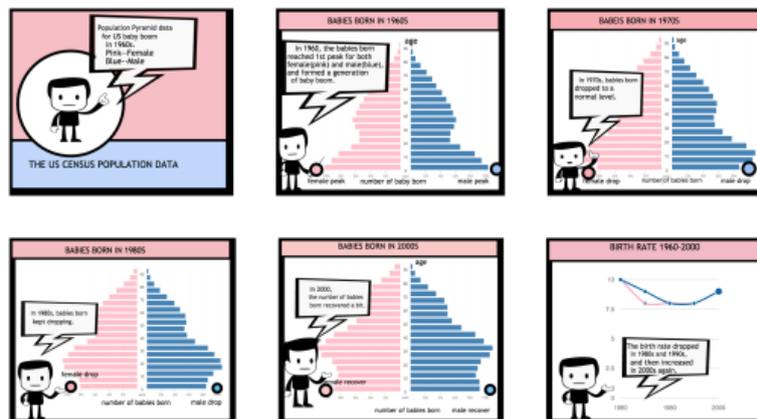


Figura 3.7 – Exemplo de *Comic Com*. Fonte: [BKH⁺16].

Quanto a categoria identificada como **elementos visuais**, esta compreende os recursos visuais que auxiliam no uso da narrativa e está dividida em três sub-categorias: (i) estruturas visuais, (ii) realce e (iii) orientação de transições [SH10]. Estruturas visuais referem-se aos mecanismos presentes que fazem com que o leitor compreenda toda a narrativa, além de identificar em que parte da história ele se encontra. Para isso, podem ser utilizados recursos como *splash screen*, padrões visuais consistentes ou barras de progresso/tempo. Foi identificada uma nova técnica, que é por nível. Nesta técnica a história



Figura 3.8 – Exemplo de *Slideshow*. Fonte: [Gua06a].

é mostrada sob demanda, através do caminho que o usuário decidiu traçar, sendo sempre comunicado sobre o nível atual da história em que ele se encontra. Geralmente, é mostrada uma visão geral da história para o usuário e, então, o usuário é livre para escolher o caminho que deseja seguir (histórias não-lineares).

O realce objetiva direcionar a atenção do leitor para algum ponto específico da visualização, podendo este realce ser por cor, áudio, tamanho, foco em partes específicas, *zoom*, entre outros. Cinco novas formas de realce foram identificadas, que são através de tamanho (aumentar o tamanho de algum elemento, a partir de algum critério), informações agrupadas (formar grupos de dados, a partir de algum critério), *slowmotion* (que é utilizado com o gênero animação [CM11b], e é utilizado para enfatizar partes da histórias que ocorreram momentos importantes, e avançar partes menos importantes ou com ênfase menor), cores e sugestão (quando o sistema sugere dados, como, por exemplo, as dez *hashtags* mais citadas [RCWR14]).

Orientação de transição diz respeito a técnicas para movimentações dentro ou entre cenas, que não causem desorientações no usuário. Estas transições podem ser, por exemplo, com o uso de animações, sob a perspectiva do leitor ou sob perspectiva do autor. Duas técnicas referentes a Orientação de transição que, além de terem sido identificadas por Stolper et al., foram encontradas nos artigos analisados, são botão próximo/anterior e *scroll*.

Stolper et al. [SLRS16] identificaram, em seu trabalho, uma sub-categoria que foi chamada de Ligação dos Diferentes Elementos da História, que possui o objetivo de traçar conexões entre diferentes elementos de história e que foi entendida como uma sub-categoria de Elementos Visuais. Inicialmente, três formas diferentes de conectar os elementos foram identificadas, que são por interação, por cor e através de animações. Todas estas técnicas foram identificadas no conjunto de trabalhos analisados, com exceção da técnica animação. Nos trabalhos analisados, foram identificadas a conexão entre elementos através de uma forma explícita, como, por exemplo, linhas ou setas, através de estrutura em árvore e por aproximação (como os elementos são organizados na interface).

A terceira categoria são os **elementos estruturais**, presentes na narrativa, que auxiliam e facilitam o seu uso. Esta também está dividida em três sub-categorias, que são (i) ordenação, (ii) interatividade e (iii) mecanismos de mensagem. A ordenação diz respeito a forma pela qual as cenas são ordenadas, podendo ser de forma randômica, direcionada de acordo com o leitor ou de forma linear. A interatividade refere-se às diferentes estratégias de manipulação do conteúdo como filtro, seleção, navegação. Através da análise dos artigos, na sub-categoria interatividade, foram identificadas mais cinco técnicas: ordenação (mudar a ordem que a informação é mostrada na interface), adicionar dados/notas (adicionar ou remover dados no *dataset* ou adicionar notas em pontos específicos da interface), organização manual (organizar como os elementos são dispostos na interface), *brushing e linking*, e comparação de dados (replicar uma determinada visualização e mudando seus critérios, a fim de comparar diferentes pontos centrais com um mesmo *template*). Por fim, os mecanismos de mensagem são as formas que as narrativas comunicam observações e comentários ao leitor, que podem ser pelo uso de títulos, *annotations*, artigos, comentários, textos introdutórios ou resumos. No que diz respeito aos mecanismos de mensagens, foi encontrada uma nova abordagem, que é através de áudio.

Para Segel e Heer [SH10] estas categorias devem equilibrar a intenção do autor da narrativa com a forma de descoberta da história por parte do leitor. Para isso, os autores dividem narrativas com visualizações sob o espectro do direcionamento da história pelo autor ou pelo leitor. Narrativas com visualizações direcionadas pelo leitor costumam ser mais interativas e não possuem ordenação, enquanto narrativas guiadas pelo autor costumam possuir pouca ou nenhuma interação e são muito dependentes de mecanismos de mensagens. No geral, as narrativas com visualizações apresentam modelos híbridos que utilizam em parte abordagens guiadas pelo leitor e em parte abordagens guiadas pelos autores.

Segel e Heer [SH10], identificaram alguns padrões de combinações no que se refere a combinação das categorias do espaço de *design* utilizadas. Os autores argumentam que, apesar das inúmeras possibilidades de combinações, algumas foram mais aparentes. Os modelos apresentados pelos autores são *Martini Glass Structure*, *Interactive Slideshow* e *Drill-Down Story* (Figura 3.9).

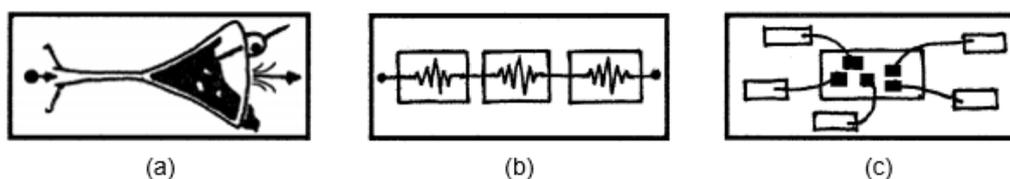


Figura 3.9 – Modelos Híbridos para narrativas. Fonte: [SH10].

Na estrutura *Martini Glass*, inicialmente a história é guiada pelo autor (a partir de questionamentos, observações ou textos). Após o direcionamento inicial do autor, o leitor é capaz de interagir livremente para explorar os dados. O modelo tem este nome pois sua estrutura lembra uma taça de Martini na qual a haste representaria a história sendo guiada

pelo autor e a parte mais expandida seria a representação da história quando ela começa a ser guiada pelo leitor (pois possui múltiplos caminhos e interatividade). A Figura 3.10 mostra um exemplo desta estrutura, na qual, no último slide, o leitor é capaz de interagir com a visualização ao clicar nos ícones presentes no mapa e, quando esta ação é feita, obter uma breve explicação sobre cada item.



Figura 3.10 – Exemplo de utilização da estrutura *Martini Glass*. Fonte: [Gua06a].

A *Interactive Slideshow* possui uma estrutura típica de *Slideshow*, entretanto, seus slides possibilitam a interação do leitor. Isto permite o que o leitor explore pontos da história antes de seguir em frente. Esta estrutura pode ser observada na Figura 3.11, que mostra padrões de aluguel de filmes (no Netflix) de doze cidades diferentes podendo o leitor ordenar os slides por ordem alfabética, por *ranking* ou pelos mais alugados. A estrutura *Drill-Down Story* apresenta um tema geral e, a partir deste, permite que o leitor escolha uma instância específica destes dados. Na Figura 3.12, é apresentado um exemplo desta representação, na qual é possível que o leitor veja detalhes das prisões da Coreia do Norte existentes no mapa.

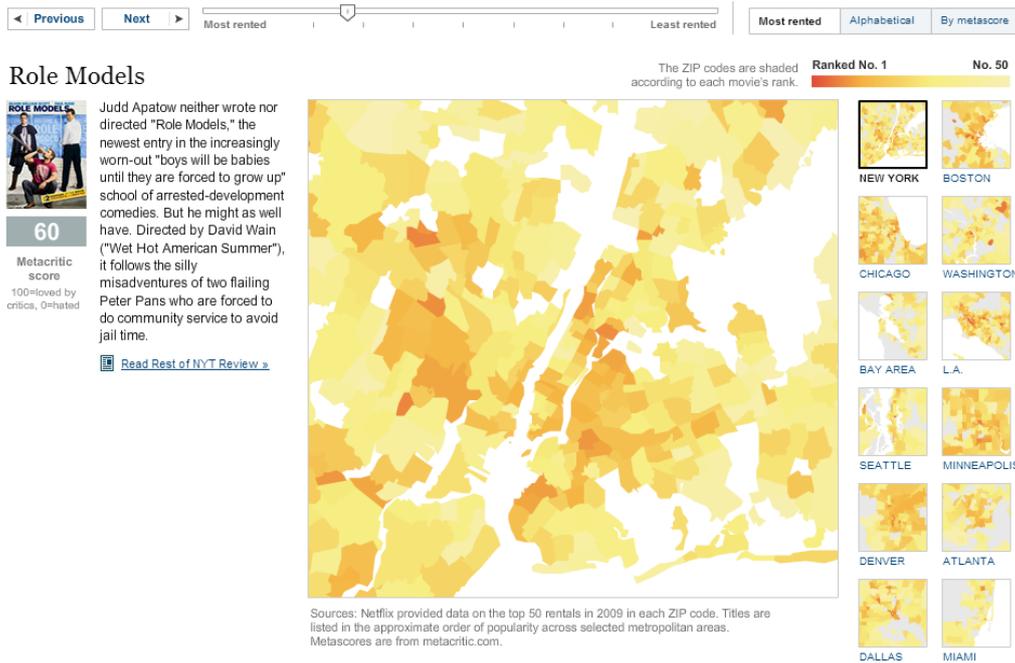


Figura 3.11 – Exemplo de utilização da estrutura *Interactive Slideshow*. Fonte: [BAMQ06].

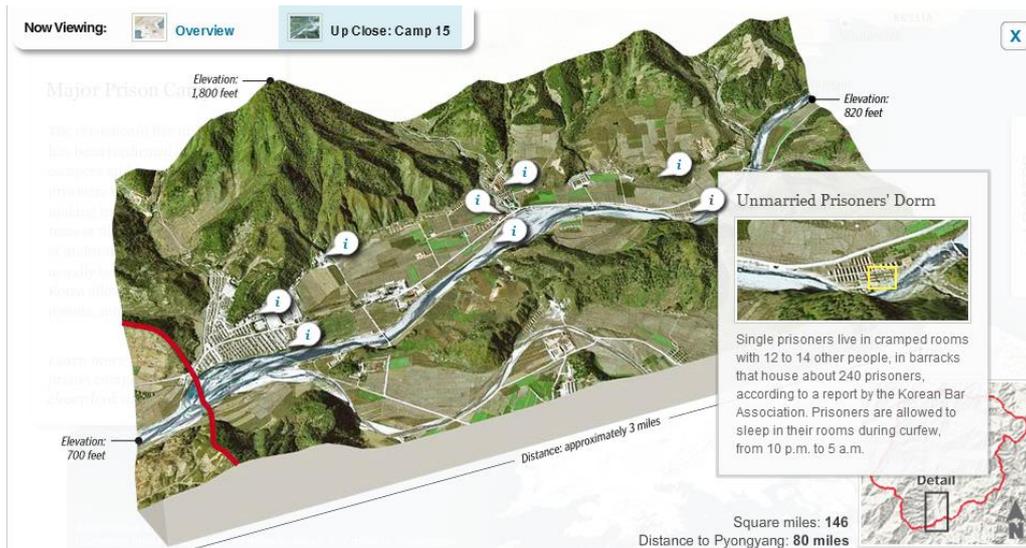


Figura 3.12 – Exemplo de utilização da estrutura *Drill Down Story*. Fonte: [Pos09].

4. TRABALHOS RELACIONADOS

Gerson e Page [GP01] argumentam que produtores e diretores de filmes, vídeos, teatro e comerciais de televisão enfrentam problemas em relação a formas de apresentar narrativas no que diz respeito a manter a atenção do público enquanto a mensagem desejada é passada. Segundo os autores, a diferença entre narrativas com visualizações e narrativas de entretenimento convencionais (como em filmes e televisão) é o uso da informação e, transmitir narrativas através de dados, torna a tarefa de entregar a mensagem muito mais complexa.

Este capítulo apresenta uma visão geral dos trabalhos que foram desenvolvidos na intersecção das áreas de visualização de informação e de narrativas. Buscou-se, na literatura, trabalhos que enfoquem o processo e as atividades para criar narrativas com visualizações, os quais são apresentados na seção 4.1, e trabalhos nos quais os autores apresentam visualizações que incorporam elementos de narrativas ou nos quais os autores pesquisaram como estas são feitas, os quais são apresentados na seção 4.2.

4.1 Processos e Atividades

Dadas as dificuldades de criar narrativas com visualizações, uma série de trabalhos focaram-se em mapear as atividades e/ou os componentes necessário para a sua criação [LRIC15, CM11a, EAB13, SH14, SH10, WR12], e, por conseguinte, tentar minimizar as dificuldades de contar estas histórias.

Enfocando seu trabalho no processo de criação de histórias para o jornalismo, Satyanarayan e Heer [SH14] propõem um modelo que contém os componentes necessários para criar narrativas com visualizações interativas. Tal modelo foi criado a partir de entrevistas qualitativas com jornalistas e estudantes de jornalismo que possuíam experiência em criação de narrativas de dados.

O modelo proposto possui os seguintes componentes: (i) parâmetros, que são os componentes presentes na interface, bem como os valores de suas propriedades; (ii) anotações textuais e gráficas, que objetivam realçar informações importantes e direcionar a atenção do leitor; (iii) gatilhos (*triggers*), que promovem mudanças na visualização, a partir de ações do usuário, tempo ou parâmetros previamente definidos; e (iv) cenas, que são blocos com a estrutura da narrativa (neste caso, cada bloco pode conter os outros componentes deste modelo).

No mesmo contexto (jornalismo), o trabalho de Weber e Rall [WR12] teve como objetivo identificar os fatores necessários para a criação de narrativas com visualizações. Para isso, utilizaram a metodologia qualitativa *Grounded Theory*. Foram conduzidas entre-

vistas com dezenove profissionais em oito empresas (cinco da Alemanha, duas da Suíça e uma dos Estados Unidos) e, com as entrevistas, identificou-se, a partir de um processo de codificação, as categorias: métodos, tempo, competências, colaboração e papéis. Estas categorias têm, como núcleo, a atitude – neste caso, atitude significa que todos os envolvidos (desenvolvedores, autores e *designers*) devam agir como jornalistas, pois todos produzem um produto jornalístico.

Segundo os autores, todos os atores devem colaborar e unir suas competências para produzir narrativas com visualizações, pois as atividades envolvidas no processo são multidisciplinares e, dificilmente, uma única pessoa terá as competências necessárias para executá-las por completo. Os entrevistados também citaram o tempo como fator importante a ser considerado pois a criação de visualizações não é feita em poucas horas. O método, neste contexto, contempla tanto questões relacionadas a colaboração, como comunicação e métodos para a criação da história propriamente dita.

Elias et al. [EAB13] também abordam o uso de narrativas, agora, no contexto de *Business Intelligence* (BI). Os autores entrevistaram cinco especialistas em BI com o objetivo de identificar, primeiramente, as principais práticas nas atividades que envolvem a transmissão dos resultados de BI, e, a partir disto, identificar os desafios que os especialistas enfrentam para criar e compartilhar suas histórias, além dos desafios que os leitores enfrentam em ler e interpretar as histórias dos autores.

No que diz respeito a análise, segundo os autores, a principal forma de reportá-las é através de relatórios. Estes contêm, geralmente, *dashboards* completos, que, quando possível, são feitos com visualizações interativas. Os leitores são ensinados através de reuniões e apresentação, além de materiais (como tutoriais e explicações) sobre o que está sendo reportado. O processo para criação envolve diversas interações com os leitores (como um gerente ou executivo), o que, segundo os autores, não ocorre em criações de histórias ou em domínios de análises visuais.

Além dos desafios, os autores questionaram os participantes a respeito da integração entre narrativa e análise. Primeiramente, questionou-se se, devido ao fato de que as análises feitas em *dashboards* eram exploratórias, as interações nas histórias deveriam ser permitidas. Após, perguntou-se se tivessem a oportunidade de criar histórias de forma centralizada (em um único lugar, com todo o suporte necessário), como a histórias deveriam, na opinião deles, ser visualizadas. No que diz respeito a interação, as respostas, em sua maioria, foram que os usuários tivessem a oportunidade de algumas interações, mas de forma limitada (como, por exemplo, possibilidade de filtrar e de selecionar resultados). No que diz respeito a como a história deveria ser mostrada, ao serem questionados em relação aos gêneros de narrativas [SH10] (já exposto no capítulo 3), todos preferiram o gênero *Annotated Chart*, embora também vissem potencial nos gêneros *Partitioned Poster*, *Slide Show* e *Flow Chart*. Por fim, os autores identificam, com base nas entrevistas, que

para o contexto de BI seria necessário um novo modelo, que consistem em uma espécie de *dashboard* anotado.

Além das entrevistas, foi conduzida uma sessão de *design* participativo com um especialista sênior em BI, e, a partir desta seção, foram identificados requisitos necessários para criação de histórias. Os requisitos identificados foram: (i) transição integrada da forma atual que criadores realizam suas análises para esta nova forma de reportar resultados; (ii) integração, que trata a necessidade de ferramentas que possibilitem criar toda a história em um único local; (iii) possibilidade de indicar focos de atenção e propiciar direcionamento aos leitores; (iv) visualizações interativas; (v) *templates* específicos para o contexto de BI; (vi) reuso de estruturas já previamente criadas, em novas histórias; e (vii) representação estática de toda história, com o objetivo de que o usuário tenha visão do contexto e que ele possa explorar a história em seu próprio ritmo.

Cruz e Machado [CM11a] também apresentam uma *engine* para construção de história, entretanto, não mencionam nenhum contexto específico. Neste trabalho, os autores utilizam a abordagem de que a história é a representação de uma fábula. Neste contexto, fabula é uma sequência de eventos, ordenados cronologicamente, causados ou vividos por autores. O mecanismo proposto tem por objetivo transformar fábulas em histórias e possui dois modelos. O primeiro, chamado de modelo de eventos, manipula o tempo dos eventos, bem como suas respectivas ações e consequências, criando, com isso, uma linha do tempo. É importante ressaltar que a simples manipulação dos eventos, guiada pelo tempo, pode não contemplar a forma pela qual o autor deseja contar a história, pois, em alguns casos, existe a manipulação do tempo na história, com recursos como analepses (mais conhecidos como *flashbacks*), prolepses (previsões ou visões do futuro) ou elipses (omissão intencional de elementos facilmente identificados, dado o contexto).

O modelo de ação lida com a representação das ações das fábulas, implementando o comportamento dos atores envolvidos na história. O resultado é um conjunto de parâmetros que podem ser utilizados para ajustar a história. Estes parâmetros podem determinar, por exemplo, quais são os atores envolvidos, ou quão representativa é a história em relação a fábula (neste contexto, a fábula pode representar o *dataset*). Diversas abordagens foram utilizadas para validar os modelos dos trabalhos apresentados, cada qual com seu respectivo objetivo. No trabalho de Cruz e Machado [CM11a], os autores demonstraram a utilização da *engine* proposta através da sua aplicação em um dataset que tratava de dados a respeito do declínio dos impérios ocidentais nos séculos 19 e 20. Neste trabalho, o foco não era validar a engine proposta, mas, sim, discutir se o ajuste dos resultados da engine (neste caso, parâmetros do modelo) pode mostrar diferentes expressões de uma mesma fábula. Os autores mostraram que, mudando os parâmetros, era possível mudar a ênfase da história. Neste caso, esta engine seria uma outra forma de manipular dados, ao invés do tratamento via *data mining*. A vantagem, na percepção dos autores, é que esta

manipulação se daria na camada de representação, e que isso promoveria a exploração e descoberta de dados mais precisos.

Weber e Rall [WR12] ilustram suas considerações através de um estudo de caso realizado na redação do New York Times. Para isso, foram realizadas entrevistas com duas pessoas que são fortemente envolvidas no processo de criação de visualização de dados, com o objetivo de obter detalhes sobre este processo, métodos de trabalho e a aplicação do conjunto de habilidades multidisciplinares para alcançar o objetivo final (construir a visualização). Após narrar e detalhar cada parte do processo da redação, os autores concluem que, além do fator central no processo de criação, que é atitude, também há a discussão da necessidade de habilidades multidisciplinares dos envolvidos e que isso poderia ter implicações em programas de universidades (no caso, seriam necessários programas acadêmicos multidisciplinares que envolvessem áreas de *design*, jornalismo, visualização de informação e computação).

Satyanarayan e Heer [SH14] instanciam o modelo proposto em um sistema, chamado de Ellipsis que é constituído por duas partes: a primeira, chamada de DSL (*domain-specific language*), é a implementação dos componentes presentes no modelo; a segunda, é uma interface gráfica de usuário, que permite a manipulação direta destes componentes, sem necessidade de experiência prévia com programação. O sistema foi avaliado por oito profissionais, do ramo de jornalismo, com experiência prévia em narrativas com visualizações. O objetivo deste estudo era obter *insights* de como o sistema proposto poderia se adaptar aos processos atuais no jornalismo e sanar possíveis necessidades que os profissionais poderiam vir a ter.

Todos os participantes foram capazes de criar histórias facilmente, através da interface gráfica e utilizar os principais recursos disponibilizados. Os *feedbacks* foram em relação a possibilidade de a ferramenta permitir colaboração entre os profissionais, pois o sistema encoraja um processo interativo entre desenvolvedores e jornalistas e, com isso, daria ao repórter um senso maior de propriedade da visualização final. Além disso, alguns outros comentários em relação a interface e como algumas *annotations* poderiam ser adicionadas também foram levantadas.

Com uma abordagem similar a utilizada no trabalho de Satyanarayan e Heer [SH14], Elias et al. [EAB13] também desenvolveram uma ferramenta para permitir a criação de histórias. Guiando-se pelos requisitos elencados (já expostos previamente nesta seção), os autores estenderam uma ferramenta já existente para incorporar os elementos de narrativas. Com isso, foram identificadas quatro entidades, que são entidades de informação (que contemplam visualizações, textos e *annotations*, por exemplo), entidades de relação (flechas, linhas e *links* de HTML), entidades organizacionais (agrupamentos visuais e sequenciamento) e, por fim, entidades de ênfase (realce e *zoom*).

Neste trabalho, além da avaliação por parte de quem cria a história, houve avaliação por parte do leitor da história. No que diz respeito a avaliação dos criadores, os

participantes acharam muito fácil e útil a forma pela qual os autores propõem a criação das histórias. Ambos os participantes acreditam que a história deva ser guiada sob a perspectiva do criador, caso contrário, o objetivo da história será perdido. Outro ponto positivo mencionado foi o *template* utilizado, que era o *Annotated Dashboard*.

Os participantes sugeriram que fosse possível utilizar dois tipos de histórias, que são histórias fixas (representam pontos específicos do *dataset*) e histórias on-line (representam a evolução dinâmica dos dados e podem ter o mesmo cenário analítico, independente dos dados). Por fim, levantou-se a necessidade de a ferramenta prover mecanismos de colaboração, tanto com diferentes criadores, quanto com os clientes. Nos testes realizados com os leitores, todos os participantes reportaram que leitura se deu de maneira fácil, não houve a necessidade de treinamento prévio e todos foram capazes de responder perguntas complexas a respeito dos dados.

Os trabalhos apresentados tinham como objetivo mostrar formas que ajudassem o processo de criação de histórias. Esta melhora veio em forma de processos, ferramentas, *engines* e *frameworks* ou mais focada em questões de *design* e formas de apresentações. No trabalho de Segel e Heer [SH10], por exemplo, embora, a partir de sua análise extensa, tenha sido possível extrair os principais gêneros de narrativas, bem como elementos de *design* utilizados, estes foram analisados no contexto de jornalismo e, quando utilizados em outro contexto, como no de BI, por exemplo [EAB13], o *template* utilizado precisou de adaptações.

Nos trabalhos de Elias et al. [EAB13] e Satyanarayan e Heer [SH14] foram apresentados componentes necessários para a criação de histórias e, em ambos os trabalhos, estes componentes foram mapeados em interfaces que permitiram que criadores de histórias criassem suas narrativas sem a necessidade de programa-las (ou seja, o criador da história não necessariamente precisaria ser um desenvolvedor). Os elementos elencados por ambos os trabalhos convergem em alguns aspectos, como, por exemplo, os componentes relacionados ao conteúdo a ser apresentado (visualizações, textos e anotações), formas de realçar conteúdo e sequenciamento e estruturação da narrativa. Outros componentes foram identificados em um modelo e não em outro, como os gatilhos (específico do trabalho de Satyanarayan e Heer [SH14]) e entidades de relação (específico do trabalho de Elias et al. [EAB13]). Outra questão interessante, que também surgiu em ambos os trabalhos, foi a necessidade de ferramentas que apoiem a colaboração. Isto pode ser pelo fato de que estes trabalhos foram aplicados em contextos em que há a necessidade de trabalho em grupo – nestes casos, visualizações no jornalismo e em relatórios de BI. Já no trabalho de Weber e Rall [WEB12], ainda que tenha sido analisado como as histórias eram criadas no contexto do jornalismo, os autores apresentam, além das categorias que influenciam o processo de criação de histórias, questões relacionadas às características pessoais dos envolvidos neste processo.

Por fim, Cruz e Machado [CM11a] atentam ao fato de que uma mesma história pode ser apresentada de várias formas, com enfoques diferentes. Para isso, sua *engine* tinha como objetivo extrair parâmetros de um determinado *dataset* e, com isso, possibilitar que esses parâmetros fossem ajustados conforme o foco que o criador da história quisesse dar. Apesar de ser uma abordagem interessante, há algumas questões ligadas à ética que não foram abordadas pelos autores. Conforme discutido por Lee et al. [LRIC15], quando visualizações são criadas para fazer com que mensagens específicas sejam melhor compreendidas e ou que apenas determinados pontos de vista sejam considerados, representações visuais podem ser mal utilizadas e conter ênfases controversas. Os autores acreditam que a mesma ética existente no jornalismo deva ser mantida na criação de narrativas com visualizações, no que diz respeito a influência que a história reportada deva ter sob o leitor.

4.2 Narrativas com Visualizações

Alguns autores enfocaram seus trabalhos em criar narrativas que utilizem visualizações [BDVA⁺15, LWW⁺13, KS12, BKH⁺16] ou apresentar como narrativas com visualizações são criadas [SH10].

Zhao et al. [BKH⁺16] focaram seu trabalho no gênero *Comic Strip*. Os autores tinham a intenção de permitir que fossem construídas narrativas utilizando painéis, que tivessem imagens ou visualizações interativas. Para isso, criou-se uma aplicação que permitiu que fossem criadas histórias, definindo seus *frames* (podendo ser adicionadas, além do conteúdo, legendas, síntese por voz ou bolhas de pensamento) e o sequenciamento da história. É possível observar, na Figura 3.7, um exemplo do resultado das histórias criadas nesta ferramenta.

Esta representação foi validada, de forma qualitativa, com doze usuários. O objetivo da avaliação era, a partir de um comparativo com slides criados em Power Point (contendo o número exato de *frames*, bem como a correspondências dos elementos contidos em ambos) coletar as seguintes métricas: engajamento, velocidade, eficiência do espaço, facilidade de uso e se a representação era agradável. A representação proposta pelos autores mostrou resultados significativamente melhores no que diz respeito a engajamento, a eficiência de espaço e a representação mais agradável. Quanto à velocidade e à facilidade de uso, não obteve-se melhoras significativas.

Balduini et al. [BDVA⁺15] focaram seu trabalho em narrativas com visualizações para grandes eventos, com dados extraídos de redes sociais (neste caso, do Twitter e do Instagram). Especificamente, os autores tinham interesse em investigar se era possível permitir que a audiência perceba padrões e observe a dinâmica dos envolvidos em um evento através de narrativas. Para isso, os autores construíram uma solução, chamada

de CitySensing, que possibilita coletar, realizar a anotação semântica, agrupar, analisar e redefinir dados capturados sobre lugares e eventos.

A interface proposta possuía duas visualizações principais, que são a geográfica e a por grafos. Na visualização geográfica era possível visualizar o volume de dados por localização e era possível verificar a evolução destes dados através de uma linha do tempo. A interface possibilitava visualizar em um nível mais amplo, como, por exemplo, por cidades, ou em um nível mais detalhado, como por bairros. Na visualização por grafos, eram mostrados os 100 usuários mais ativos, que se relacionam entre si, a partir do momento que falam do mesmo evento ou local. Na Figura 4.1(a) é apresentado um exemplo da visualização geográfica e, na Figura 4.1(b), um exemplo da visualização por grafo, existentes na solução CitySensing.

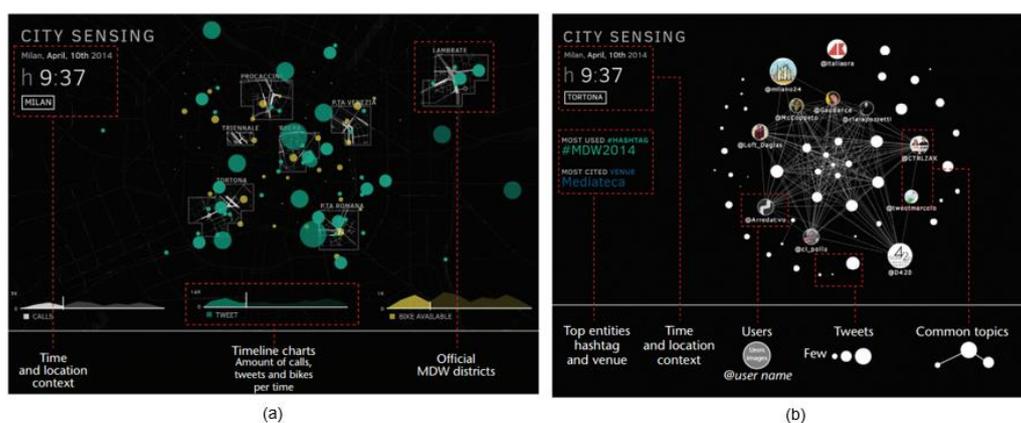


Figura 4.1 – Visualizações do CitySensing. Fonte: [BDVA⁺15].

Para avaliar a interface, os autores prepararam uma visualização e realizaram perguntas a 23 participantes, os quais nunca haviam utilizado a interface antes e não possuíam habilidades específicas com visualizações. Foram feitas oito perguntas (com alternativas de resposta que poderiam ser verdadeira, falsa ou incerta). Para duas destas perguntas não havia resposta correta (incerta) e para seis haviam respostas específicas. As respostas dos participantes estavam, em sua grande maioria, corretas. Isto, segundo os autores, é um indicativo de que a audiência pode se beneficiar, perceber padrões e observar a dinâmica dos envolvidos através dos dados.

Liu et al. [LWW⁺13] abordam o uso de narrativas de dados através de relações de entidades de uma história ao longo do tempo. Na visualização, que pode ser observada na Figura 4.2 cada entidade é representada por uma linha, que termina quando sua presença na história também termina. Ao longo da *timeline* os locais ou acontecimentos vão emergindo, conforme estes aparecem na história e a visualização mostra quais personagens estão presentes.

Foram conduzidos estudos de caso com *script* de três filmes, e os resultados foram comparados entre si. Além disso, três participantes foram entrevistados (um professor da área de cinema, um estudante de doutorado em sociologia e um professor da área de

Se percebe que é possível utilizar narrativas com visualizações em diversos contextos e com diferentes objetivos. Nesta seção, por exemplo, há trabalhos que foram utilizados no jornalismo [SH10], em grandes eventos [BDVA⁺15] e em históricos de desenvolvimento de sistemas [KS12]. Em todos os trabalhos que houve avaliação das visualizações propostas, os usuários foram capazes de entender e conseguiram vislumbrar benefícios a respeito deste tipo de visualização, mostrando que narrativas de dados podem ter um grande potencial no que diz respeito a obtenção de conhecimento com dados.

5. MODELO PARA REPRESENTAÇÃO DE NARRATIVAS INTERATIVAS POR MEIO DE VISUALIZAÇÕES COM DADOS EXTRAÍDOS DE REDES SOCIAIS

Para auxiliar na reflexão sobre a questão de pesquisa proposta neste trabalho (*Como interfaces que utilizam narrativas com diferentes técnicas de visualização podem apoiar a interação do usuário e a sua análise de dados obtidos a partir de redes sociais?*), foi construído um modelo de representação de visualizações com narrativas, com dados obtidos de redes sociais.

Para a elaboração do modelo, primeiramente, buscou-se entender como a área tem abordado a construção das narrativas no contexto de visualização de dados. Para isso, foram analisados trabalhos científicos, tanto aqueles que visavam auxiliar a construção de tais narrativas, como aqueles que apresentavam narrativas com visualizações. Além disso, buscou-se entender a teoria referente a narrativas, independente do seu uso na área de visualização.

Com isso, foram identificados sete componentes que integram a construção de narrativas com visualização de dados, que são os **dados**, os **elementos da narrativa**, as **informações**, os **gêneros**, os **elementos visuais**, os **elementos estruturais** e as **visualizações**.

Estes componentes, além de suas características próprias, possuem relação entre si. A partir dos dados, extraem-se os elementos de narrativa. A partir dos elementos da narrativa, é possível determinar quais informações serão apresentadas aos usuários. Estas são apresentadas através das visualizações, que são organizadas de uma determinada forma (gênero). Os elementos estruturais e visuais, além de estruturar a narrativa, terão influência em como será a interação com a interface. Estes componentes podem ser utilizados tanto na estrutura da interface como um todo, quanto em visualizações específicas. Na tabela 5.1, são apresentados os componentes com seus respectivos sub-componentes (foi utilizado o caractere # para representar os componentes que não possuem sub-componentes).

Na figura 5.1 é apresentada uma possível utilização deste modelo. O gênero (*partitioned poster*) é representado pela organização dos componentes, principalmente, no que se refere às múltiplas visualizações presentes na interface. No *container* Navegação, está uma possível utilização dos elementos estruturais, que apresentam o progresso da narrativa (através de uma *progress bar*) e as formas de avançar e retroceder a narrativa (botão próximo e anterior, e através de um *slider*).

Já o *container* Filtros, contém outro possível uso de elemento estrutural, apresentando, exclusivamente, formas para filtrar informações (entretanto, outras técnicas desse componente poderiam ser empregadas na interface). Por fim, no *container* Visualizações

Tabela 5.1 – Modelo proposto com os sub-componentes de cada componente

Componente	Sub-Componente
Dado	Redes Sociais
Elementos da Narrativa	Eventos
	Atores
	Localização
	Tempo
Informações	#
Gênero	<i>Partitioned Poster</i>
Visualizações	#
Elementos Visuais	Estruturas Visuais
	Realce
	Ligação dos Diferentes Elementos da História
Elementos Estruturais	Ordenação
	Interatividade
	Mecanismos de Mensagem

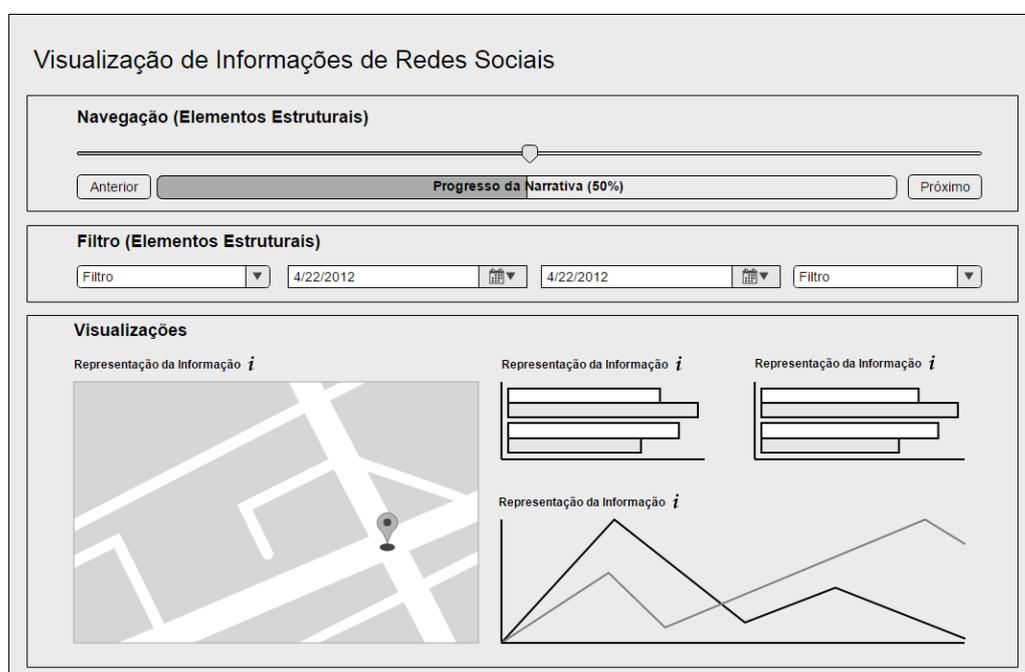


Figura 5.1 – Exemplo de utilização do modelo.

estão representadas as visualizações que apresentam informações determinadas pelo autor da narrativa (por exemplo, a quantidade de uso de algum dado, representada por um gráfico de barras). Nas seções seguintes será apresentado um maior detalhamento de cada componente deste modelo.

5.1 Dados

Considerando o contexto explorado neste trabalho, o modelo foi construído com o intuito de se utilizar dados oriundos de redes sociais. Para tal, qualquer dado de rede social que possua postagens de usuários poderia ser utilizado (como, por exemplo, dados de Facebook¹, Twitter², Instagram³). Diversas redes sociais, atualmente, fornecem API'S (*Application Programming Interface*) para extração de seus dados, pelas quais é possível obter não apenas as postagens em si, mas também informações relacionadas, como o usuário que postou, a data, a localização do usuário no momento da postagem, entre outros.

5.2 Elementos da Narrativa

Conforme comentado previamente, existem elementos que são necessários para a criação de narrativas. Estes elementos são necessários, inclusive, no contexto de visualização. Por isso, com o objetivo de auxiliar na criação de narrativas com visualizações, propõe-se uma forma para mapear uma estrutura de uma postagem para estes elementos. Para este fim, baseou-se no conjunto proposto por Cruz [CM11a], que contempla os seguintes elementos: eventos, atores, tempo e localização.

A partir da análise da estrutura dos dados a serem utilizados, propôs-se uma forma de mapear a estrutura de postagens para os elementos de narrativa:

1. Eventos: o evento pode ser identificado através de *tags*⁴. Como é comum a utilização de múltiplas *tags* em uma única postagem, a partir de um evento principal, pode-se buscar eventos relacionados ou sub-eventos.
2. Atores: os atores podem ser identificados a partir do sujeito da frase, dos usuários (que efetuaram a postagem) e/ou dos possíveis usuários mencionados na postagem.
3. Localização: a localização pode ser caracterizada pelos locais que o evento em si ocorre (como, por exemplo, os estados, as cidades ou os bairros). A localização obtida via georreferência da postagem (quando existente), também pode ser utilizada como recurso para obter informação sobre este elemento. É possível utilizar, também, locais mencionados na postagem.

¹www.facebook.com

²www.twitter.com

³www.instagram.com

⁴ *Tags* são palavras que objetivam categorizar conteúdos na internet, podendo estas ser, por exemplo, por lugar (*#nycity*), por sentimento (*#love*), por algum evento (*#olimpiada2016*), ou qualquer outra frase ou palavra que o usuário deseja.

4. Tempo: o tempo pode ser caracterizado pela data da postagem e/ou datas importantes do evento. Também é possível identificar referências de datas nas postagens.

Na Figura 5.2 é apresentado um exemplo de como tais informações poderiam ser extraídas, em uma postagem que ocorreu no Twitter. Nesta figura, cada cor de marcação representa um elemento: amarelo representa o evento, azul representa o tempo, vermelho representa o ator e verde a localização.

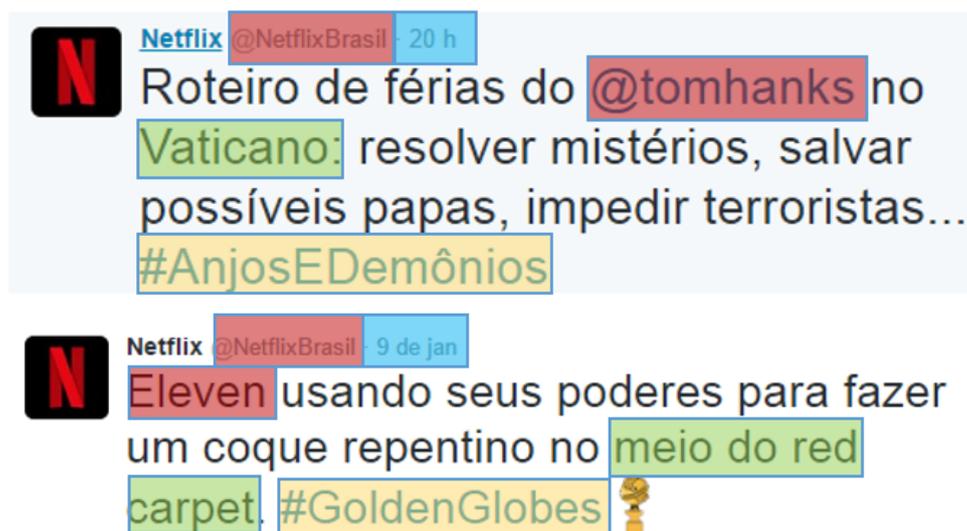


Figura 5.2 – Exemplo de elementos a serem extraídos do modelo.

Em relação aos exemplos, em ambas as postagens seria possível obter o evento através das *tags*, bem como o tempo através da data da postagem. No que se refere aos atores, estes poderiam ser obtidos pelo nome do usuário (presente em ambas as postagens), pela marcação de outro usuário (apenas na primeira postagem) ou a partir da obtenção do sujeito da frase (apenas na segunda postagem). Já a localização poderia ser obtida através da análise da postagem, presentes em ambos os exemplos.

5.3 Informações

A informação é um dos pontos centrais da área de Visualização de Informações. Há inúmeras definições para o termo informação, mas a informação pode ser entendida como a associação de dados com intuito de produzir um conteúdo que expresse um significado superior [PUC06]. Os dados, de maneira isolada, embora expressem valor, não possuem um valor informativo. Nesse sentido, a união de dados formarão informações completas [PUC06], havendo diversas combinações de dados que podem ser utilizados para formar informações.

Trazendo o conceito de dado e de informação para o presente modelo, o dado, neste contexto, é expressos pelos elementos, e a informação seria a combinação destes elementos ou dos elementos e de outras informações associadas, que são do interesse do autor da narrativa apresentar para sua audiência. Tais informações, no final, são expressas através de visualizações.

Com o intuito de verificar quais tipos de informações poderiam ser utilizadas no presente modelo, analisou-se um conjunto de visualizações disponíveis na internet [Int09b, Int09a, Int09c, Bel09b, Bel09c, Won09, Bel09a, Sen09, Nod09, Pro09, Mus09], bem como um conjunto de trabalhos científicos [SDM14, MBB⁺12, BDVA⁺15], que representavam informações especificamente para redes sociais. Procurou-se visualizações que estivessem disponíveis on-line, por meio de uma pesquisa via mecanismos de busca, e selecionou-se uma amostra dos resultados obtidos. Já no que se refere aos trabalhos analisados, estes foram obtidos através do resultado de uma revisão sistemática previamente executada.

A partir desta análise, verificou-se que é possível representar informações de apenas um elemento ou pela combinação de mais de um elemento. A seguir é apresentada uma lista dos elementos com as possíveis informações que podem ser a eles relacionadas:

- Eventos: volume de *tags*; *tags* mais frequentes; *tags* mais utilizadas por sentimento; relação entre *tags*;
- Atores: relação de/entre atores/grupo de atores; listagem de atores/grupos de atores; atores mais frequentes; volume de postagens;
- Localização: quantidade de postagens por local; local por sentimento; local da postagem;
- Tempo: quantidade de postagens por algum critério de tempo (como, por exemplo, ano, mês, dia, hora).

Conforme previamente mencionado, foram, também, encontradas informações oriundas da combinação de mais de um elemento. A lista das informações com os elementos combinados é apresentada a seguir:

- Tema + Tempo: volume do uso de *tags*; sentimento por data;
- Localização + Atores: número de seguidores por local;
- Ator + Tempo: menções ao longo do tempo; seguidores ao longo do tempo; usuários mais frequentes ao longo do tempo;
- Ator + Evento + Tempo: utilização de *tag* por usuário ao longo do tempo.

5.4 Gênero

Conforme apresentado no capítulo 3, o gênero é a forma básica de apresentar a narrativa. Ele define o *template* dos elementos que serão dispostos na interface, além de definir como serão as cenas da narrativa, e quais caminhos o leitor poderá tomar ao longo da narrativa. Baseando-se nos gêneros propostos por Segel e Heer [SH10], verificou-se quais gêneros eram viáveis ou não, para que o objetivo deste trabalho fosse atingido. Entende-se que por estar explorando a interatividade, o gênero *Film/Video/Animation* não se adequa, pois o gênero possui pouca ou nenhuma interatividade e um sequenciamento estritamente linear da narrativa.

Já os gêneros *Magazine Style* e *Annotated Chart*, também seriam uma opção de difícil viabilidade pois demandam um alto grau de informações textuais e contextualizadas, que talvez fosse difícil de se obter dos dados de redes sociais. Já os gêneros *Comic Strip* e *Slide Show*, por mostrarem uma narrativa com uma sequência estritamente linear no que se refere a interação, não permitiriam a exploração não direcionada da narrativa.

Por isso, optou-se pelo uso do gênero *Partitioned Poster*, pois este contempla a utilização de mais de uma técnica de visualização (*multi-view*) para representar a narrativa, o que está diretamente relacionado com a questão de pesquisa e com os objetivos deste trabalho (no que diz respeito ao uso de técnicas de visualização). Além de contemplar a utilização de múltiplas visualizações, o gênero não sugere uma navegação estritamente linear, permitindo uma exploração não direcionada por parte do usuário. A representação gráfica desse modelo pode ser vista na Figura 5.3.

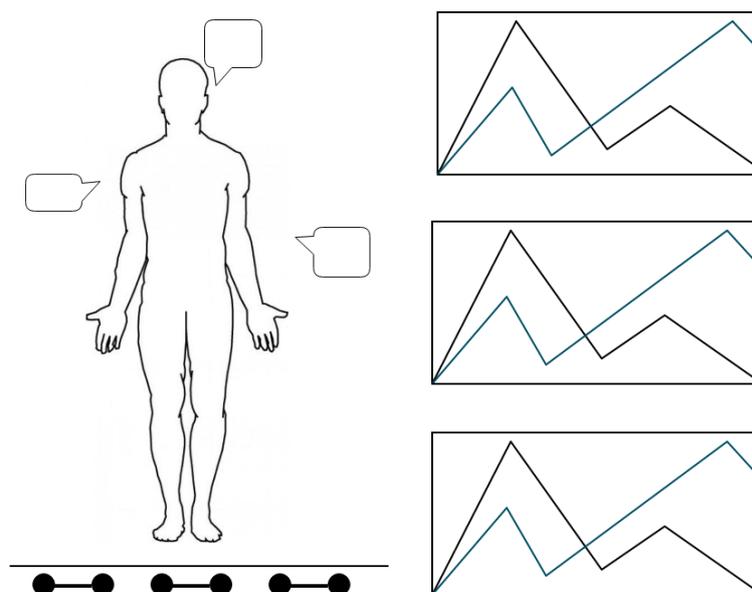


Figura 5.3 – Representação Gráfica do Gênero *Partitioned Poster*. Fonte: [GP01].

5.5 Visualizações

As visualizações são as formas pelas quais as informações são expressas na interface. É através delas que a narrativa será apresentada aos leitores. Uma mesma informação pode ser apresentada de diversas maneiras, e a escolha de qual técnica será utilizada pode variar de acordo com qual se adéque mais a narrativa a ser apresentada, além do desejo do autor da narrativa de apresentar a informação de uma determinada maneira.

Ao verificar a lista de possíveis informações que poderiam ser apresentadas (seção 5.3), verificou-se quais técnicas poderiam ser utilizadas. Esta verificação baseou-se nos estudos apresentados no capítulo 2 (seção 2.3), verificando o propósito da informação (por exemplo, mostrar uma relação ou uma proporção), e as possíveis formas de representação para tal fim. Com isso, a tabela 5.2 apresenta a lista dos elementos (e de combinações destes) com as informações que podem ser apresentadas, com as sugestões de algumas técnicas de visualizações que poderiam ser utilizadas.

5.6 Elementos Visuais

Os elementos visuais referem-se aos recursos visuais que auxiliam o leitor no uso da narrativa, de uma maneira mais ampla (ou seja, os elementos que influenciam a narrativa como um todo). As categorias que envolvem tal dimensão do espaço de *design* (conforme a seção 3.3) são as estruturas visuais, realce, orientação de transições e ligação de diferentes elementos da narrativa. No que se refere as estruturas visuais, para este modelo, entende-se que as principais técnicas a serem exploradas são a plataforma visual consistente (que diz respeito a manter uma plataforma visual consistente, a medida que o leitor interage com a interface, mudando apenas o conteúdo dos painéis da interface, mantendo o *layout* intacto), a estruturação por níveis (mostra os fatos sob demanda, a medida que o usuário interage com a interface). Formas de progresso na narrativa também são importantes, e poderiam ser incorporadas de diversas maneiras.

No que se refere à sub-categoria realce, esta pode ser explorada de diversas maneiras como, por exemplo, por cor, tamanho, foco ou sugestão. Entende-se que, neste modelo, deva existir, pelo menos, uma técnica de realce presente, pois podem haver partes da narrativa que se sobressaem em relação a outras, e, com isso, pode auxiliar o usuário a tomar um determinado caminho na narrativa. Entende-se que não seja necessária a utilização da orientação de transição, pelo fato de não haver mais de uma cena. Por fim, em relação a ligação dos diferentes elementos da história, a principal técnica que sugere-se que seja utilizada, é através de interação (no que se refere a interação nas visualizações, como forma de direcionamento da narrativa), entretanto, outras técnicas podem ser utiliza-

Tabela 5.2 – Tabela com a relação entre os elementos, as informações para apresentação e formas de representá-los.

Elemento de Narrativa	Informação	Visualizações
Eventos	Volume de <i>tags</i>	Gráfico de Pizza
		Tabela/ Listagem
		Gráfico de Barras
		Nuvem de Palavras
		Mapa de Calor
	<i>Tags</i> mais utilizadas por sentimento	Gráfico de Barras
		Gráfico de Pizza
		Mapa de Calor
		Nuvem de Palavras
Relação entre <i>Tags</i>	Grafo	
	<i>Hierarchical Edge Bundle</i>	
	<i>Treemaps</i>	
Atores	Relação de entre Atores	Grafo
		<i>Hierarchical Edge Bundle</i>
		<i>Treemaps</i>
		<i>Sunburst</i>
	Atores	Listagem/Tabela
	Atores mais frequentes Volume de Postagens	Listagem/Tabela
		Gráfico de Barra
		Nuvem de Palavras
		Mapa de Calor
		Gráfico de Pizza
Gráfico de Bolha		
Localização	Quantidade de postagens por local Local por sentimento Locais das postagens	Mapa
		Gráfico de Barra
		Listagem/Tabela
		Mapa de Calor
		Gráfico de Pizza
		Gráfico de Bolha
	Menções ao Longo do Tempo Seguidores ao longo do tempo Usuários mais frequentes ao Longo do Tempo	Gráfico de Linha
		Gráfico de áreas empilhadas
		Gráfico de Linhas
		Gráfico de Barras
		Gráfico de Pontos
		Gráfico de Escala Contínua
		<i>Streamgraph</i>
Tema + Tempo	Volume do uso de <i>Tags</i> Sentimento por data	Gráfico de Linhas (Múltiplas)
		Gráfico de Barras Empilhadas
		Gráfico de Áreas Empilhadas
		Gráfico de Barras
Localização + Atores	Número de seguidores por local	Mapa
		Mapa de Calor
		Gráfico de Pizza
		Gráfico de Barras
Ator + Evento + Tempo	Utilização de <i>Tag</i> por usuário ao longo do tempo	Gráfico de Barras Empilhadas

das em conjunto. Na tabela 5.3 são apresentados os sub-componentes deste componente, bem como a sugestão das técnicas que podem ser empregadas.

Tabela 5.3 – Sub-Componentes que compõe os elementos visuais com as respectivas técnicas.

Sub-componentes	Técnicas
Estruturas Visuais	Plataforma Visual Consistente
	Barras de Progresso
	Estruturação por Níveis
Realce	Cor
	Tamanho
	Sugestão
	Foco
Ligação dos Diferentes Elementos da História	Interação

5.7 Elementos Estruturais

Os elementos estruturais são os recursos que auxiliam a interação do leitor com a narrativa. As categorias que envolvem esta dimensão são ordenação, interatividade e mecanismos de mensagem. No que tange a ordenação, a forma que entende-se que a narrativa deva ser estruturada através da ordenação aleatória, mas, também prover mecanismos para direcionar o leitor (caso seja de interesse do leitor seguir a narrativa linearmente).

As principais técnicas de interatividade a serem utilizadas são o detalhamento de informações (*Hover Highlighting / Details*) e mecanismos para filtro e seleção. Botões de navegação também poderiam ser incorporados neste modelo. Acredita-se que as demais técnicas (como, por exemplo interatividade limitada ou instrução explícita), não devam ser exploradas, por não estarem de acordo com o propósito deste trabalho. Por fim, o mecanismo de mensagem que pode ser explorado é o cabeçalho, com o intuito de contextualizar o usuário a respeito da visualização apresentada. Na tabela 5.4 são apresentados os sub-componentes que compõe este componente, bem como a sugestão das técnicas que podem ser empregadas.

Tabela 5.4 – Sub-componentes que compõe os elementos estruturais com as respectivas técnicas.

Sub-componentes	Técnicas
Ordenação	Randômica
	Direcionada de Acordo com o Leitor
Interatividade	Detalhamento de Informações
	Botões de Navegação
	Filtro e Seleção
Mecanismos de Mensagem	Cabeçalho

6. #SEGUEATOCHA: INSTANCIACÃO DO MODELO PROPOSTO

A fim de analisar o modelo proposto, foi construída uma interface a partir de seus componentes, com foco no revezamento da tocha das olimpíadas de 2016. Embora alguns componentes sejam mais precisos no que se refere a forma a qual este deve estar contemplado na interface, outros componentes permitem implementações de forma mais livre. Nesse sentido, neste capítulo serão apresentadas as formas pelas quais cada componente foi transposto para a interface, bem como um detalhamento de questões relacionadas aos dados e os princípios que guiaram a implementação dessa interface.

6.1 Implementação da Interface

Nesta seção serão apresentadas, para cada componente do modelo proposto, as técnicas empregadas para sua implementação, bem como a justificativa para as decisões tomadas. Ao final, será apresentada a interface desenvolvida, apresentando os componentes incorporados e as opções de interatividade presentes.

6.1.1 Dados

Conforme consta no modelo, a base de dados é composta por dados advindos de redes sociais. Dentre as possibilidades existentes, escolheu-se o Twitter por já se ter acesso a um conjunto de dados coletados desta rede¹, e pelo fato de os dados serem adequados às características do modelo.

Relativo à base de dados, foram utilizados dados relacionados as Olimpíadas 2016, que estavam sendo coletados desde novembro de 2015. No período da execução deste trabalho, havia em torno de 80 mil *tweets* coletados. Dentre os *tweets* das Olimpíadas, foram selecionados aqueles que eram referentes ao revezamento da tocha (caracterizados pelo uso da *hashtag* "TochaOlimpica" ou que possuíssem a palavra Tocha). Dos 80 mil *tweets*, em torno de 26 mil *tweets* possuíam essa palavra e/ou *hashtag*.

Optou-se por utilizar esses dados pois era um evento que havia uma duração relativamente grande (o evento ocorreu entre 03/05/2016 e 05/08/2016) e por haver uma referência de localização que não era proveniente dos *tweets*. Os dados coletados através do Twitter eram o usuário, a data e a postagem. Com a data da postagem, era possível saber em qual cidade a tocha olímpica estava naquele momento, ou seja, não havia a

¹Os dados foram coletados no contexto do projeto "Visualização Dinâmica de Dados de Redes Sociais e sua inclusão em Rotinas de Produção Jornalística", Edital PRAIAS/PUCRS 2016

necessidade de utilizar a localização proveniente da postagem para fazer a associação com a postagem em si (pois apenas os *tweets* oriundos de *smartphones* possuíam tal informação). Com isso, além da base de dados dos *tweets*, também coletou-se (de forma manual) as datas e as cidades/estados pelos quais a tocha estava passando. É importante ressaltar que, oficialmente, não havia os pontos de cada cidade em que a tocha passou, portanto, a informação era a nível de cidade e não de bairro/ruas. Também não encontrou-se informação de todos os participantes que carregaram as tochas nas cidades, e, por isso, essa informação também não pode ser utilizada.

Anterior à realização do mapeamento dos elementos para a estrutura do *tweet*, foi realizado o tratamento e limpeza dos dados. Este tratamento permitiu, além de extrair as informações necessárias, extrair informações adicionais, que foram utilizadas para construir as visualizações.

Esta etapa foi implementada utilizando a linguagem de programação C#.NET (versão do *framework* 4.6). Os dados (*tweets*) estavam em um arquivo do tipo CSV; com isso, leu-se o arquivo que contém, além do *tweet* em si, informações sobre o usuário que postou, a geolocalização, a linguagem e a data. Para cada *tweet*, o processo de limpeza e identificação foi realizado. As etapas da limpeza dos dados são as seguintes:

- Identificação e remoção do caractere '#' (que caracteriza *hashtags*) no meio do *tweet*, para que esta palavra seja entendida como parte da estrutura da frase;
- Identificação e remoção do caractere '@' (que caracteriza menções a usuários) no meio do *tweet*, para que esta palavra seja entendida como parte da estrutura da frase;
- Identificação se o *tweet* é um *retweet* ou não (através da identificação do início como sendo "RT:");
- Substituição de gírias e abreviações por palavras em sua estrutura formal;
- Correção e identificação de palavras incorretas;
- Identificação das URL's (*Uniform Resource Locator*) externas do *tweet*.

No que diz respeito a informações relacionadas, foram identificados todos os usuários envolvidos no *tweet* (o usuário que postou e os usuários citados via mensagem direta), a identificação das *hashtags* e as URL's externas (como *links* para portais de notícias ou vídeos). Além disso, identificou-se, ainda, se o *tweet* é um *retweet* (se este foi publicado através do compartilhamento de outro *tweet*). A identificação destas informações foi realizada através do uso de expressões regulares.

O processo de limpeza dos dados consistiu nas seguintes etapas: identificação de *hashtags* e usuários que constam no *tweet* (para que estas estruturas sejam entendidas como parte da estrutura da frase); remoção de *tags* de HTML (*HyperText Markup Language*)

que acredita-se que foram obtidas erroneamente pela ferramenta de coleta; e a tentativa de correção de palavras escritas erroneamente, utilizando o corretor ortográfico Hunspell², que é o corretor ortográfico utilizado pelo Libre Office. Foram limitadas as palavras que eram processadas pelo corretor ortográfico, excluindo-se usuários e palavras que estavam contidas no dicionário de Palavras Vazias (*stop words*), que são palavras que são excluídas de um processamento de texto, por não exercerem influência no resultado esperado [LRU14].

Além da exclusão das palavras com a utilização do corretor ortográfico, criou-se um dicionário de gírias para substituir estas gírias por uma palavra formal (como, por exemplo, a substituição da palavra "vc" por "você") para atingir melhores resultados na identificação da estrutura da frase. Importante ressaltar que todas as mudanças realizadas no *tweet* foram salvas em uma estrutura separada, mantendo o *tweet* original para fins de futuras exibições.

6.1.2 Elementos de Narrativa

Por ter-se decidido utilizar os dados referentes a tocha olímpica, foram buscados os dados oficiais do revezamento da tocha³, no que diz respeito a datas e locais (para expressar a relação espaço-temporal, relacionada aos elementos Tempo e Localização). Entretanto, foi utilizada a data proveniente dos *tweets*, para fins de filtros e relações entre as datas de eventos e as datas dos *tweets*.

O elemento Evento foi definido como sendo o Revezamento da Tocha Olímpica de 2016. Nesse sentido não houve a necessidade de detectar eventos a partir dos *tweets*. Os *tweets* que possuem relação são aqueles que possuíam a *hashtag* TochaOlímpica ou que possuíam a palavra Tocha (neste caso, acredita-se ser pertinente o uso da palavra Tocha como critério de inclusão para os *tweets* sobre a corrida da tocha, pois são coletados apenas os *tweets* referente as olimpíadas). O elemento Atores foi caracterizado pelos usuários dos *tweets* (tantos os usuários que postaram, como os usuários mencionados). Espera-se que a audiência possa identificar sub-eventos ou eventos relacionados, selecionando ou adicionando outras possíveis *hashtags* ou atores, com o objetivo de mudar a perspectiva dos fatos da narrativa que estão sendo apresentados.

6.1.3 Informações

As tomadas de decisões referentes às informações que seriam apresentadas ao usuário seguiram o princípio de que deveriam haver informações relacionadas a cada ele-

²<https://hunspell.github.io/>

³<https://www.rio2016.com/revezamento-da-tocha>

mento de narrativa. Não foram utilizadas todas as informações previstas no modelo (seção 5.3) da lista, por se entender que, por questões de *layout*, a leitura da narrativa poderia ser prejudicada.

As informações a serem apresentadas foram definidas a partir da análise dos dados, do que poderia ser extraído, além da análise das informações que desejava-se que fossem apresentadas. Por fim, chegou-se ao seguinte conjunto de informações:

- Eventos: volume de *tags* e *tags* mais frequentes;
- Atores: listagem de atores e atores mais frequentes;
- Localização: quantidade de postagens por local e locais do evento;
- Tempo: quantidade de postagens por dia e quantidade de postagens por hora.

6.1.4 Gênero

O modelo pressupõe o uso do gênero *Partitioned Poster*. O gênero, que não possui troca de cenas (ou seja, toda a narrativa é apresentada em um *frame* único), tem como base a utilização de múltiplas visões. As visões podem ou não ser visualizações (poderiam ser utilizadas, por exemplo, informações textuais). Neste caso, foram utilizadas apenas visualizações para a construção da narrativa.

6.1.5 Visualizações

Para a escolha das visualizações, verificou-se as informações que seriam apresentadas. Procurou-se utilizar a mesma técnica de visualização para um mesmo tipo de informação (por exemplo, utilizou-se gráfico de barras sempre que se queria apresentar o volume de uso de um elemento, quando não se havia a presença do tempo). Verificando-se as possíveis técnicas de visualização para as informações a serem apresentadas, escolheu-se aquelas julgadas mais pertinentes para o propósito do trabalho:

- Eventos:
 - Volume de *hashtags*: gráfico de barras;
 - *Hashtags* mais frequentes: *word cloud*.
- Atores:
 - Listagem de atores: gráfico de barras;

- Atores mais frequentes: *word cloud*.
- Localização:
 - Quantidade de postagens por local: gráfico de barras;
 - Locais do evento: mapa;
 - Rota entre os locais do evento: mapa.
- Tempo
 - Quantidade de postagem por dia: gráfico de linha;
 - Quantidade de postagem por hora: gráfico de linha.

6.1.6 Elementos Visuais

No que se refere aos elementos visuais, para a sub-categoria estruturas visuais foi utilizada a estruturação por níveis (na qual o nível de detalhe da narrativa vai aumentando conforme o usuário vai interagindo com a interface). No caso desta interface, o usuário pode analisar a narrativa sob a perspectiva do país ou sob a perspectiva de algum estado específico. Além disso, outros elementos podem ser escolhidos para a análise, podendo o usuário analisar a narrativa sob o ponto de vista de um ator, ou de um evento específico. Dado que as visualizações mudam conforme os filtros escolhidos pelos usuários, pode ocorrer que um novo ator ou evento fique aparente devido aos filtros que o usuário escolheu, caracterizando, com isso, a estruturação por níveis. Além disso, a plataforma se mantém consistente conforme o usuário interage com a interface, caracterizando, com isso, a adoção da técnica Plataforma Visual Consistente.

Já na sub-categoria realce, foi utilizado realce por sugestão (pelos gráficos de barras, que mostram os dados mais relevantes) e por tamanho (nas *word clouds*). O objetivo era utilizar este tipo de técnica para apresentar as informações mais evidentes do *dataset* a fim de mostrar o que estava mais aparente, de acordo com os filtros do usuário. Por fim, a ligação dos elementos se dá pela interação, a qual, em uma determinada visualização, influencia nas informações que serão mostradas nas demais visualizações da interface. As técnicas utilizadas em cada sub-categoria deste componente podem ser visualizadas na Tabela 6.1.

6.1.7 Elementos Estruturais

No que se refere aos elementos estruturais, para a sub-categoria ordenação, o sistema desenvolvido é uma mescla de direcionada de acordo com o leitor e acesso randô-

Tabela 6.1 – Técnicas de Elementos Visuais utilizadas.

Sub-categoria	Técnicas Utilizadas
Estruturas Visuais	Plataforma Visual Consistente
	Estruturação por Níveis
Realce	Sugestão
	Tamanho
Ligação dos Diferentes Elementos da História	Interação

mico. Ao ser apresentado à interface, o leitor tem a possibilidade de seguir qualquer caminho, de forma não linear e, a qualquer momento, adicionar os filtros e focos que deseja. Entretanto, caso seja de interesse do leitor, ele é capaz de seguir uma ordem linear, seguindo a ordem que a tocha percorreu, o que é apresentado ao usuário, ou através dos elementos de navegação, presentes no sistema.

As técnicas de interatividade incorporadas foram detalhamento de informações e filtro e seleção. No que se refere ao detalhamento de informações, quando os leitores clicam em qualquer ponto das visualizações, são apresentadas informações relacionadas, como data e quantidade de *tweets*. Os filtros foram incorporados de duas formas, que são por um conjunto de campos e pelas próprias visualizações. No filtro pelos campos, era apresentado um conjunto de opções (período, usuário, *hashtag* e estado) para que o usuário filtrasse livremente as informações que desejasse. No caso dos filtros pelas visualizações, poderia filtrar a narrativa de acordo com as informações apresentadas na visualização (por exemplo, na *tag cloud* de usuários, era possível selecionar qualquer usuário que fosse apresentado, apenas clicando no usuário desejado, ou, também, selecionar algum estado no próprio mapa que mostrava o percurso da tocha). Por fim, os mecanismos de mensagens que foram incorporados nesta interface foram cabeçalhos, que tinham o objetivo de auxiliar o leitor, indicando qual informação estava sendo mostrada. As técnicas utilizadas em cada sub-categoria deste componente podem ser visualizadas na Tabela 6.2.

Tabela 6.2 – Técnicas de Elementos Estruturais utilizadas.

Sub-categoria	Técnicas Utilizadas
Ordenação	Randômica
	Direcionada de Acordo com o Leitor
Interatividade	Detalhamento de Informações
	Filtro e Seleção
Mecanismos de Mensagem	Cabeçalho

6.2 #SegueATocha: Interface para a Narrativa Proposta

Considerando as características de cada componente, foi desenvolvida uma interface cujo o objetivo era analisar as potencialidades do modelo proposto. Procurou-se prover

pelo menos uma visualização para cada elemento de narrativa. Além disso, em todas as visualizações existentes na interface é possível selecionar o dado que está sendo mostrado, para mudar a perspectiva da narrativa, adicionando filtros pelo elemento selecionado. Além da interação através das visualizações, foi provido um conjunto de filtros para o usuário, para que ele pudesse filtrar os elementos livremente.

Inicialmente é apresentada uma série de visualizações, com dados de todo o *dataset*. Conforme desejo do leitor, é possível, através da interação com as visualizações ou através do uso do filtro, refinar os dados e mudar a perspectiva da narrativa, podendo este refinamento ser, por exemplo, por algum usuário específico, por alguma *hashtag* ou por algum estado. A figura 6.1 mostra a visão geral da interface, apresentando, principalmente, as informações expressas através das visualizações. No mapa (Figura 6.1 A), estão sendo apresentados os locais do evento, bem como a rota entre eles. Já no gráfico de linha (Figura 6.1 B) é apresentada a quantidade de postagens por dia e a quantidade de postagens por hora (quando se está exibindo informações de apenas um dia). Os *tweets* são apresentados em forma de tabela (Figura 6.1 C). Os gráficos de barra apresentam a quantidade de postagens por local, listagem dos atores e o volume de *hashtags* (Figura 6.1 D, E e F, respectivamente). Já nas *word clouds* são apresentados os atores mais frequentes e as *hashtags* mais frequentes (Figura 6.1 G e H, respectivamente).

No mapa, foi utilizada uma linha que representa o caminho percorrido pela tocha, em duas visões diferentes: a primeira, que mostra a rota entre os estados; e a segunda, que mostra a rota entre as cidades de um determinado estado. Na visão dos estados, o ponto geográfico que é marcado no mapa é a capital. É possível selecionar o estado através do mapa e, com isso, mudar a perspectiva para o estado selecionado. Ao lado do nome do estado ou cidade, é apresentado um número, que tem por objetivo orientar o usuário, mostrando a ordem do revezamento. A figura 6.2 (a) apresenta a visão do mapa com as rotas entre os estados e a figura 6.2 (b) apresenta um exemplo de visão detalhada da rota entre as cidades do estado do Rio Grande do Sul.

No gráfico de linhas é apresentada a relação entre o volume de *tweets* e a data da postagem (Figura 6.3 (a)). Quando é clicado em algum determinado ponto, é apresentada a data referente ao ponto selecionado, bem como a quantidade de *tweets* naquele período. Além disso, é mostrada a opção para filtrar pela data referente ao ponto selecionado (Figura 6.3 (b)). Quando o filtro utilizado pelo usuário possui as mesmas datas de início e de fim, ao invés de apresentar um gráfico de data x volume de *tweets*, é apresentado um gráfico que mostra o volume de *tweets* no decorrer das horas deste dia.

Além do gráfico de linha, foram implementados mais dois tipos de visualização, que são o gráfico de barras e a *word cloud*. Ao total, três gráficos de barras foram disponibilizados (Figura 6.4). O primeiro mostra as cidades com maior número de *tweets*, o segundo as *hashtags* mais utilizadas, e, o terceiro, os usuários mais aparentes. Na figura 6.4 (a) é apresentado um exemplo de detalhamento, bem como a opção de filtro, que está disponí-

vel quando o usuário seleciona o dado em determinado ponto do gráfico. Esta opção está disponível em todos os gráficos de barras presentes na interface.

Referente às *word clouds*, em ambas é possível clicar em um determinado usuário ou *hashtag* e, com isso, utilizá-lo como filtro. Na figura 6.5 (a) é apresentada a *word cloud* referente aos usuários, e na figura 6.5 (b), a *word cloud* referente às *hashtags*. Por fim, os *tweets* são apresentados em uma tabela, como é apresentado na figura 6.6. Esta tabela apenas apresenta os *tweets* e não possui nenhuma interatividade.

Além de possuir a opção de filtrar e selecionar elementos através das visualizações, existe, também, a possibilidade de filtro através de um conjunto de campos disponibilizados na interface. É possível filtrar por data de início e de fim, por usuário, por *hashtags* ou por estado. Os filtros (6.7) podem ser mostrados ou escondidos, para que estejam visíveis apenas quando os usuários os quiser.

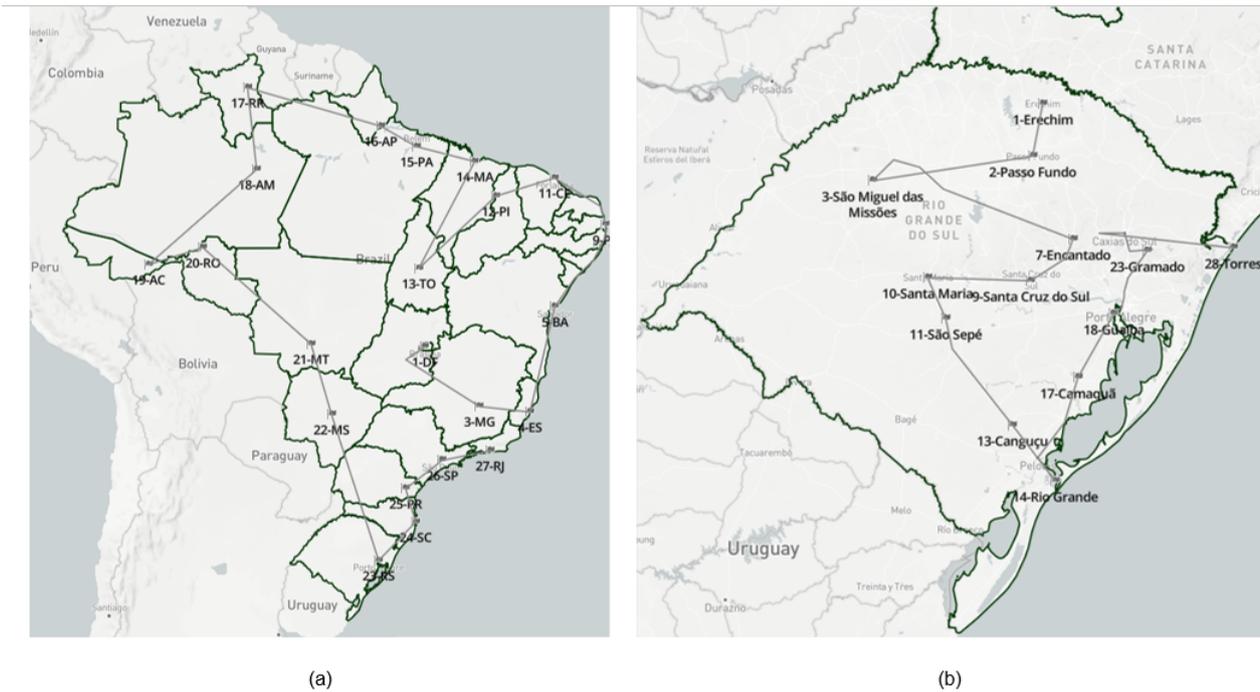


Figura 6.2 – (a) Mapa com as rotas entre os estados e (b) Exemplo de visão do mapa detalhando o estado do Rio Grande do Sul.

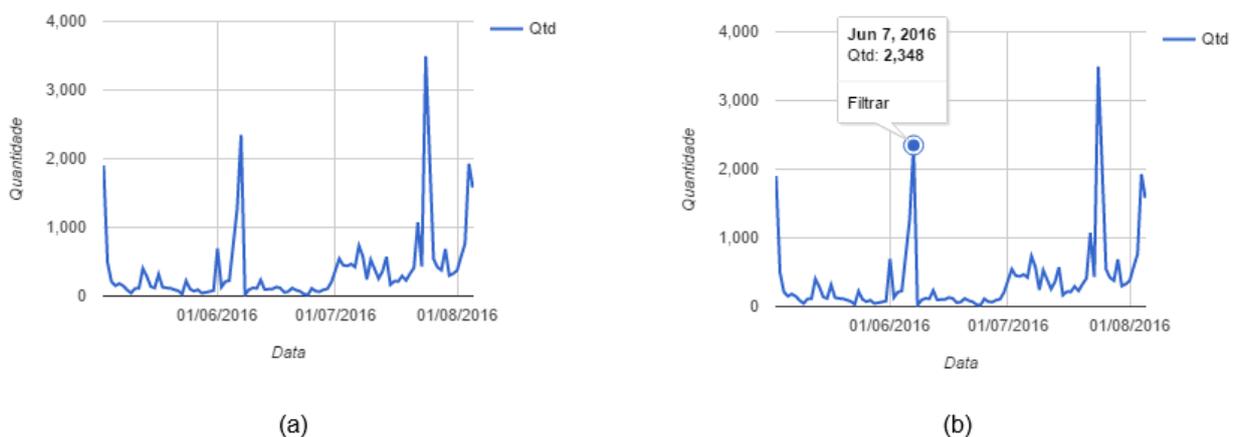


Figura 6.3 – (a) Gráfico de Linha com a relação dos *tweets* x Data e (b) Filtro e detalhamento do Gráfico de Linha.

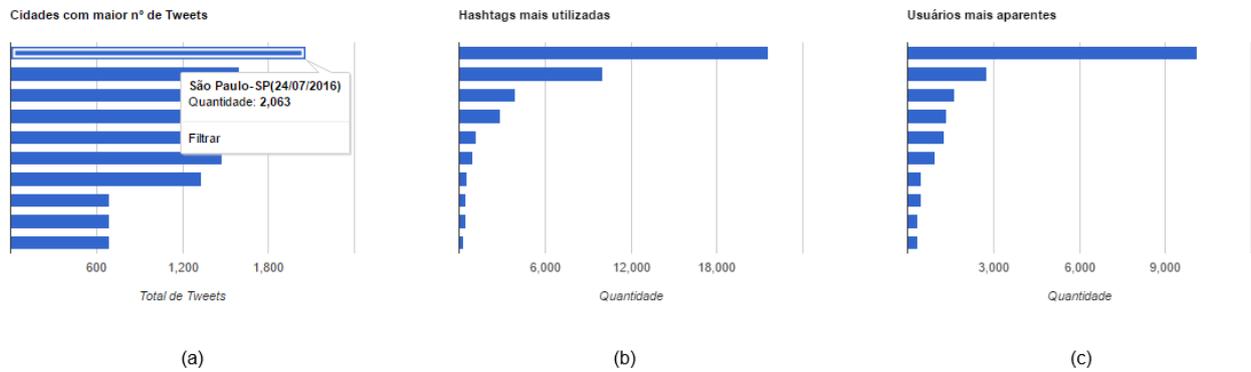


Figura 6.4 – (a) Gráfico de barras com as cidades com maior número de *tweets* e detalhamento e opção de filtro; (b) Gráfico de barras com as *hashtags* mais utilizadas e (c) Gráfico de barras com os usuários mais aparentes.

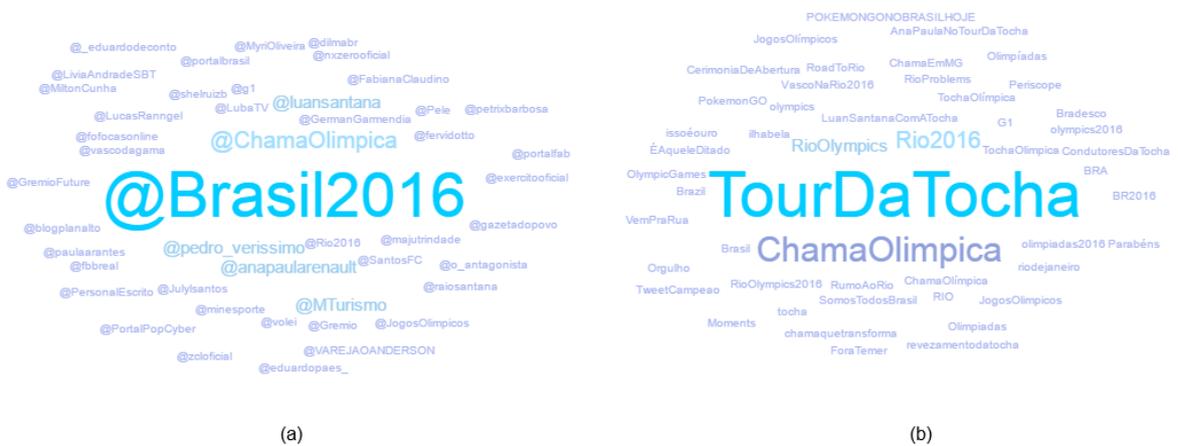


Figura 6.5 – (a) *Word cloud* com os usuários e (b) *Word cloud* com hashtags.

Tweet
A #ChamaOlimpica entra em espaço aéreo nacional nesta madrugada! Acompanhe aqui a cobertura dessa chegada! #Rio2016 https://t.co/LYTqR2VGbD
RT @portalfab: A #ChamaOlimpica entra em espaço aéreo nacional nesta madrugada! Acompanhe aqui a cobertura dessa chegada! #Rio2016 https://...
RT @portalfab: A #ChamaOlimpica entra em espaço aéreo nacional nesta madrugada! Acompanhe aqui a cobertura dessa chegada! #Rio2016 https://...
RT @portalfab: A #ChamaOlimpica entra em espaço aéreo nacional nesta madrugada! Acompanhe aqui a cobertura dessa chegada! #Rio2016 https://...
RT @portalfab: A #ChamaOlimpica entra em espaço aéreo nacional nesta madrugada! Acompanhe aqui a cobertura dessa chegada! #Rio2016 https://...
RT @portalfab: A #ChamaOlimpica entra em espaço aéreo nacional nesta madrugada! Acompanhe aqui a cobertura dessa chegada! #Rio2016 https://...
RT @portalfab: A #ChamaOlimpica entra em espaço aéreo nacional nesta madrugada! Acompanhe aqui a cobertura dessa chegada! #Rio2016 https://...
RT @AviacaoGovBr: A chama olímpica já está a bordo do 767-300 da LATAM rumo a Brasília! #TourDaTocha #Rio2016 https://t.co/TY2dst5KeA
RT @portalfab: A #ChamaOlimpica entra em espaço aéreo nacional nesta madrugada! Acompanhe aqui a cobertura dessa chegada! #Rio2016 https://...
Só bobagem... #rio2016 #chamaolimpica #olimpiadas https://t.co/nFYsiwmClx https://t.co/voF7kFpLNQ
1 2 3 ... 10916 10917 10918 > >

Figura 6.6 – Tabela com os *tweets*.

FILTROS
Fechar

▼ PERÍODO

Dt. Início

Dt. Fim

▼ USUÁRIO

Usuário

▼ HASHTAG

Hashtag

▼ LOCAL

Escolha um Estado

Filtrar

Figura 6.7 – Opções de Filtro.

7. #SEGUEATOCHA: ANÁLISE DO MODELO PROPOSTO

A fim de analisar a narrativa criada a partir do modelo proposto, foi elaborada uma etapa de estudos com usuários. O principal objetivo desses estudos foi analisar as potencialidades do modelo proposto, verificando seus pontos positivos e negativos, bem como as oportunidades de melhoria. Os estudos foram realizados em duas etapas, com um total de 11 participantes (6 na primeira etapa e 5 na segunda). Neste capítulo serão apresentadas as etapas desses estudos, apresentando o procedimento seguido, a metodologia de análise, a execução de cada etapa, uma visão geral de como se deu a análise da narrativa por parte dos participantes e, ao final, uma discussão geral dos resultados obtidos.

7.1 Procedimento

A fim de incentivar a exploração da interface e a análise dos dados apresentados, os participantes foram convidados a explorar os dados de duas maneiras distintas: na primeira, a exploração era livre; na segunda, era fornecido um conjunto de informações e o usuário deveria encontrá-las, utilizando os recursos providos na interface. Com isso, objetivava-se analisar o uso da ferramenta tanto de uma maneira mais exploratória, a fim de analisar informações sem expectativas prévias, quanto para encontrar informações específicas.

Para atingir tais objetivos, um protocolo com uma série de procedimentos foi elaborado (Apêndice B) e suas etapas podem ser vistas na Figura 7.1. Primeiramente, foram apresentados aos participantes os objetivos da pesquisa e uma breve explicação de como seria realizado o estudo. Após foi solicitado que os participantes assinassem o termo de consentimento livre e esclarecido e, então, foi aplicado um questionário com algumas questões iniciais, relacionadas ao uso de redes sociais e experiência com uso de computadores.

Antes da etapa da exploração livre da interface, era explicado aos participantes o funcionamento da interface, bem como as possibilidades de interação existentes. Após, os participantes deveriam interagir com a interface, durante o tempo que achassem pertinente. Era mencionado, ainda, que os participantes poderiam tirar dúvidas com o pesquisador presente em qualquer momento durante a interação. Após a exploração livre, o participante era questionado a respeito dos dados que foram possíveis de serem extraídos, a partir da interface.

A etapa seguinte consistia em realizar uma série de tarefas, pré-definidas, possuindo questões específicas, para as quais o usuário deveria buscar respostas através da interface (no total, era solicitado que o usuário buscasse 5 informações, da forma que ele

achasse mais pertinente). Por fim, na última etapa do estudo, o usuário deveria responder questões focadas no uso da interface e na sua experiência de uso.

Considerando-se que narrativas com visualizações são amplamente usadas no contexto de jornalismo, as quais são criadas para o grande público, sem a necessidade de conhecimentos específicos prévios, optou-se por participantes de perfil amplo, apenas com pré-requisito de possuir familiaridade com computação, não demandando de conhecimento em análise de dados ou serem usuários de alguma rede social específica.

Com o objetivo de analisar se o protocolo criado atendia as necessidades da pesquisa, realizou-se um estudo piloto, com um usuário, que se enquadrava no perfil esperado. Esse estudo permitiu a análise dos procedimentos previstos, a fim de se refinar os instrumentos de estudo e verificar o tempo de execução do procedimento [MG05].

Com o estudo piloto realizado, verificou-se que algumas questões não eram muito relevantes para o foco de análise (especificamente, as questões G, H, I e J da etapa 2 (vide Anexo B)), dado que não objetivava-se contrapor dados referentes à experiência do usuário com visualizações. Além disso, como objetivava-se verificar qual seria o foco dos participantes, na narrativa, havia uma pergunta, ao final da primeira etapa, que solicitava que o participante descrevesse os fatos que ele havia conseguido extrair a partir da narrativa apresentada pela interface. O participante que realizou o teste não soube responder a esta questão. Por isso, removeu-se esta parte etapa e o protocolo foi alterado para que o participante, no momento da interação, narrasse os fatos que estavam sendo apresentados a partir da interface. O protocolo alterado após o teste piloto, pode ser verificado no Apêndice C. O tempo total gasto nessa etapa durou em torno de 45 minutos.

7.2 Metodologia de Análise

Para análise dos dados coletados nesses estudos, decidiu-se utilizar a metodologia Análise de Conteúdo. Segundo Holsti [Hol69], a análise de conteúdo é uma técnica para fazer inferências, identificando, objetiva e sistematicamente, características da mensagem. Nesse sentido, quaisquer fonte de dados, como gravações de vídeo, uso de *software* ou música, por exemplo, podem ser utilizadas, não atendo-se somente a fonte de dados textuais.

O objetivo deste tipo de análise é a identificação de categorias, através de uma ou mais passagens de texto ou outra fonte de dados, que expressem a mesma ideia teórica e descritiva [Gib09]. O processo de análise, que é chamado de codificação, pode ser realizado a partir de duas abordagens distintas [LFH10], que são a codificação *a priori* e codificação emergente. Na codificação *a priori*, é utilizado um conjunto de categorias já conhecido (obtido, por exemplo, da literatura) e, então, é identificado, na fonte de dados da pesquisa, as passagens que se enquadrem na categoria do conjunto utilizado. Já na codifi-

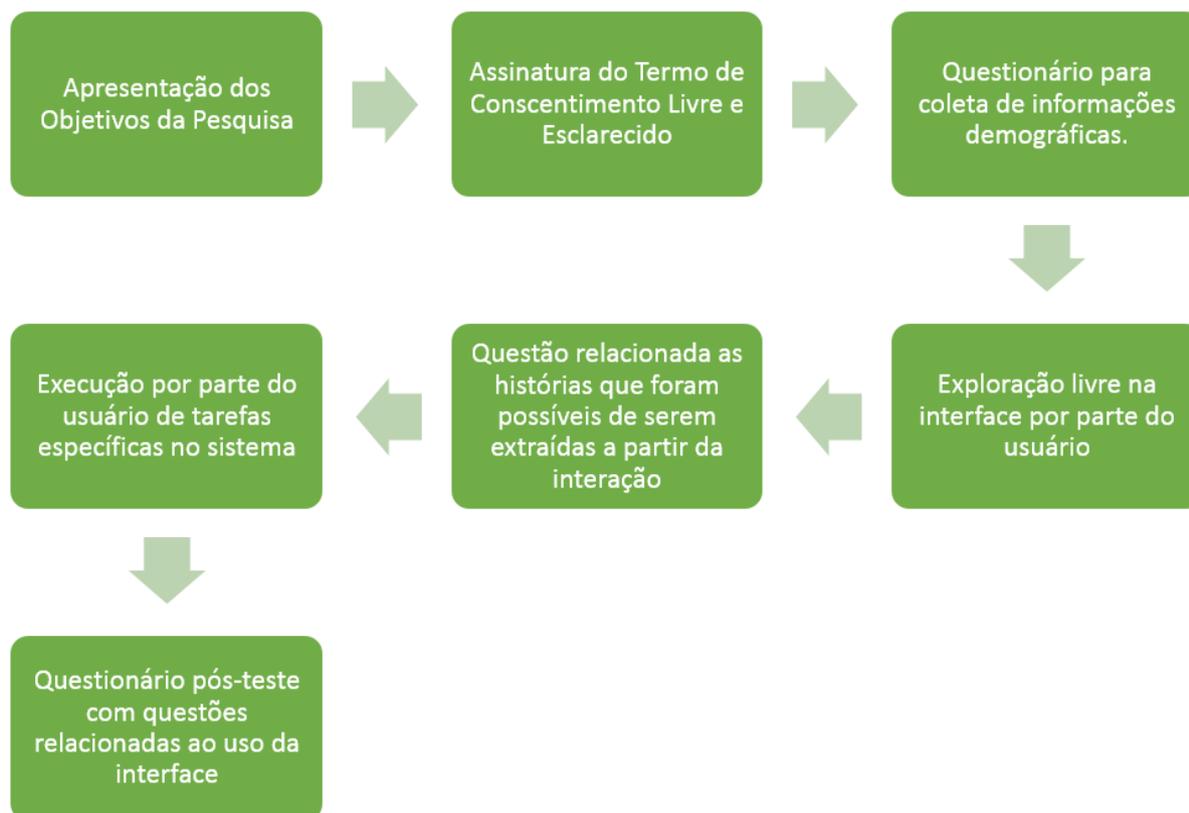


Figura 7.1 – Representação gráfica da etapa de estudo com usuários.

cação emergente, as categorias não são conhecidas, e são identificadas pelo pesquisador, através de sua análise. A utilização de uma ou de outra varia de acordo com quão o tópico de estudo já foi estudado previamente e se há um conjunto de categorias já existente. No caso da presente pesquisa, não foi encontrado nenhum conjunto de categorias, por isso, utilizou-se a abordagem codificação emergente.

7.3 Execução da Primeira Etapa dos Estudos

Foram realizados estudos individuais com seis participantes. A amostra selecionada se deu a partir da escolha por conveniência. O grupo era composto, em sua maioria, por estudantes da área da computação, com exceção de dois participantes. A idade dos participantes variava entre 20 e 28 anos, dos quais, 4 eram do gênero feminino e 2 eram do gênero masculino. Quanto a experiência com uso geral de computadores, todos consideravam-se avançados.

No que se refere ao uso de redes sociais, todos eram usuários e utilizavam, principalmente, o Facebook, o Instagram e o Twitter. Alguns participantes também utilizavam o Snapchat, o LinkedIn, o Youtube e o Reddit. As principais finalidades pelas quais as re-

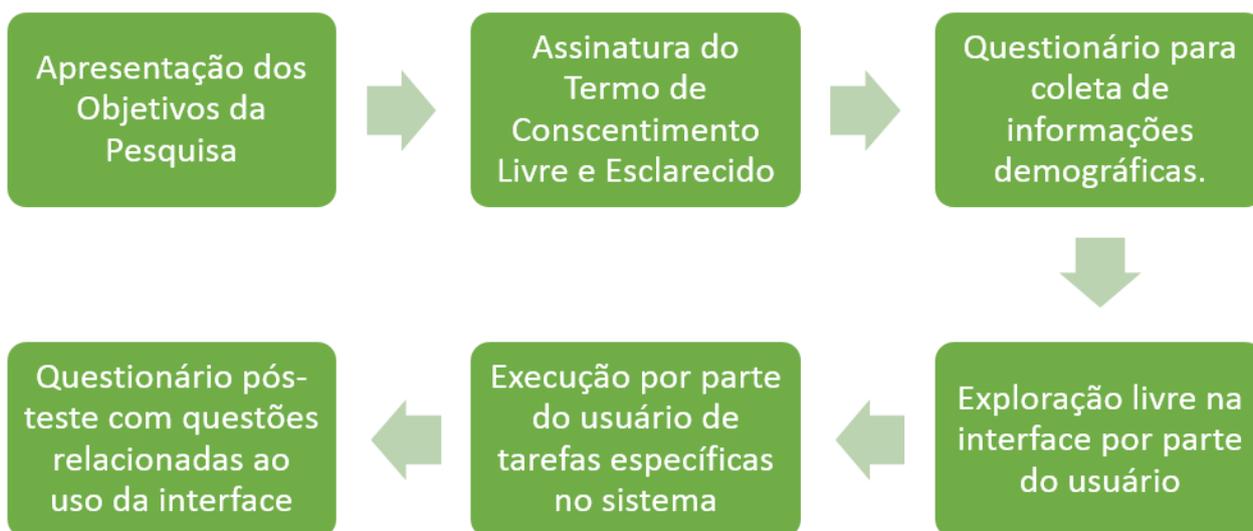


Figura 7.2 – Representação gráfica da etapa de estudo com usuários após a execução do teste piloto.

des sociais eram utilizadas, que foram citadas, foram comunicação com amigos e família, ver conteúdos diversos, ler notícias, entretenimento e também para questões de trabalho e universidade. A tabela 7.1 apresenta um resumo do perfil dos participantes desta etapa.

Tabela 7.1 – Perfil dos participantes da primeira etapa do estudo

Participante	Gênero	Idade	Escolaridade	Experiência no Uso de Computadores	Utiliza Redes Sociais
P1	Feminino	28	Doutorado (Em andamento)	Avançado	Sim
P2	Feminino	21	Superior (Em andamento)	Avançado	Sim
P3	Masculino	21	Superior (Em andamento)	Avançado	Sim
P4	Feminino	29	Mestrado (Em andamento)	Avançado	Sim
P5	Feminino	24	Superior Completo	Avançado	Sim
P6	Masculino	27	Mestrado (Em andamento)	Avançado	Sim

Todos os participantes tinham conhecimento a respeito do evento (revezamento da tocha). Quando perguntado sobre os fatos relacionados ao evento que os participantes tinham conhecimento, os principais fatos que os participantes se recordavam foram o de pessoas que caíram (alguns participantes lembraram de algumas pessoas específicas) e de pessoas que tentavam apagar a tocha.

Após a coleta das informações iniciais de perfil, e após o uso do sistema pelo participante, realizou-se uma entrevista para coletar a impressão dos participantes a respeito da interface utilizada. Durante o uso da interface, o pesquisador anotava algumas impressões a respeito da utilização. A partir das anotações, da entrevista com o participante e, também, da análise do uso do sistema (feita pela gravação da tela do computador), realizou-se o processo de codificação dos dados coletados. Com isso, chegou-se a 3 macro-categorias, que são Interface, Interatividade e Dados. As subseções seguintes apresentarão os resultados da execução das tarefas específicas (solicitas na Etapa 5 do Apêndice C e os resultados para cada categoria da codificação.

7.3.1 Execução das Tarefas Específicas

Na etapa 5 do protocolo (Apêndice C), era solicitado que os participantes executassem 5 tarefas específicas através da interface. As tarefas eram buscar informações que foram solicitadas a eles, com o objetivo de analisar como se daria a interação do usuário, quando questões precisas eram solicitadas. No que se refere a execução das tarefas, todos os participantes conseguiram encontrar as informações, com exceção da P4 na T5. A sumarização da execução das tarefas, pelos participantes, pode ser verificada na Tabela 7.2.

Tabela 7.2 – Resultado da Execução das Tarefas Específicas da Etapa 1.

Participante	T1	T2	T3	T4	T5
P1	Sim	Sim	Sim	Sim	Sim
P2	Sim	Sim	Sim	Sim	Sim
P3	Sim	Sim	Sim	Sim	Sim
P4	Sim	Sim	Sim	Sim	Não
P5	Sim	Sim	Sim	Sim	Sim
P6	Sim	Sim	Sim	Sim	Sim

Os usuários eram capazes de buscar as informações através dos filtros pelos campos (laterais) ou através das visualizações. Na T1 (busca do usuário com maior número de *tweets* em um estado), todos os participantes filtraram através do mapa, embora a opção de buscar pelos campos estivesse também disponível. Para encontrar as informações, todos os participantes recorreram ao gráfico de barras.

Já no que se refere à T2 (cidade na qual houve mais *tweets* entre um período), todos os participantes buscaram através dos filtros laterais (dado que era a única forma de selecionar duas datas) e, para verificar o usuário, P5 e P6 recorreram à *word cloud* e, os demais, ao gráfico de barras.

Na T3 (cidade que o evento estava ocorrendo em uma determinada data) os participantes também utilizaram os filtros laterais, com exceção de P3 que buscou a data pelo gráfico de linha. A análise da informação ocorreu através do próprio mapa (dado que o usuário era direcionado para o ponto a qual a tocha estava passando, quando o filtro pela data era utilizado) por todos os participantes, com exceção de P3 e P5, que verificaram através do gráfico de barras.

Quanto à T4 (volume de *tweets* em uma data), alguns participantes (P1 e P5) filtraram a informação através dos filtros laterais, e verificaram através do rodapé da tabela, que mostrava o total de registros. Já os demais, verificaram no gráfico de linhas a data, e verificaram a quantidade, através do próprio gráfico.

Por fim, a T5 (número de *tweets* de um estado) foi a única que houve um participante que não conseguiu concluir. Para filtrar pela data, todos os participantes optaram

por utilizar os filtros laterais. Entretanto a maioria dos participantes tiveram certa dificuldade de encontrar a informação (que era apresentada no rodapé da tabela de *tweets*), e, P4 não conseguiu encontrá-la.

7.3.2 Interface

No que se refere a interface, foram extraídas 3 sub-categorias, que são Características, Pontos Positivos e Sugestões. Dentro da categoria característica, houve uma divisão entre as características positivas e negativas. As categorias identificadas podem ser vistas na tabela 7.3. Na tabela, há apenas a estrutura hierárquica dos categorias, entretanto a estrutura com as falas originárias, bem como os respectivos participantes, podem ser vistas no Apêndice E.

No que se refere as características positivas, a interface foi apontada como amigável, completa, limpa, fácil, organizada, curiosa, interessante, provendo muitas possibilidades e orientando usuários para buscar dados desconhecidos. No que se refere às categorias curiosa e interessante, a participante se referia, de forma positiva, a maneira a qual as informações estavam dispostas na interface *"Achei bem curiosa e interessante a forma que ele coloca as informações"* (P1) - ele, nesse caso, é a interface que apresentava a narrativa. Já na categoria fácil, além da participante P4 que falou explicitamente que era fácil no momento em que estava interagindo com a interface, o participante P1 comentou que, dada a quantidade de informação, a interface, além de fácil de utilizar, estava organizada. Além disso, segundo P2, a interface provê muitas possibilidades para análise, dado que era disponibilizada qualquer informação que o usuário quisesse, e de diversas formas diferentes. Segundo P5, a interface orienta a exploração para dados desconhecidos: *"eu não sabia quais usuários mais tuitados, nem sabia quais hashtags mais utilizadas, então isso aqui (nuvem) me auxiliaram a saber por qual buscar."*

Já as características negativas apontadas foram limitada, confusa e que necessita tempo para aprendizado. No que se refere a categoria limitada, o participante referia-se ao fato de que só era possível selecionar, como filtro, uma porção dos dados, os quais eram apresentados através das visualizações (como, por exemplo, usuários mais aparentes, cidades que mais tuitaram, *hashtags* mais frequentes), não sendo possível navegar entre os dados que eram menos recorrentes. P2, além de vislumbrar que a interface provia muitas possibilidades para análise, considerou isso, também um ponto negativo, por considerar que a interface podia se tornar confusa em alguns momentos, dado que haviam muitas possibilidades. Para ela, a interface causou certa desorientação em alguns momentos. Segundo a participante: *"Mas ao mesmo tempo, isso acabou se tornando um pouco confuso, pois, tinha horas que eu não sabia exatamente para onde ir pois haviam muitos itens na tela"*. Por fim, P2, quando comentou que usuários poderiam se beneficiar da interface, caso

Tabela 7.3 – Estrutura de Codificação Para o Código Interface.

Estrutura da Codificação				
Interface	Características	Positivas	Amigável	
			Completa	
			Limpa	
			Fácil	
			Organizada	
			Curiosa	
			Interessante	
			Provê muitas Possibilidades	
			Orienta para Dados Desconhecidos	
		Negativas	Limitada	
			Confusa	
			Tempo para aprendizado	
			Pontos Positivos	
			Mapa Interativo	
	Forma de Expressar as Relações			
	Foco			
	Apresenta as Rotas			
	Filtro conforme Interação			
	Rotas provê ordenação			
	Sugestões			
	Cor para realce			
	Cor para foco			
	Tipos de Gráficos por Usuário			
	Acesso aos Filtros Sempre Disponíveis			
	Tamanho pra Realce			
	Aumentar tamanho da Fonte			
	Mostrar Informação (<i>label</i>) nos Gráficos			
	Filtrar Cidade pelo Mapa			
	Filtro de Cidade por Campo			
	Formas para Limpar os Filtros Específicos			
Voltar ao Estado Anterior				
<i>Word cloud</i> de Cidades				
Limitar o Usuário				
Sugestão nos Filtros por Campo				
Filtro para os Menos Tuitados				
Mostrar Data na Rota				

investissem tempo para entender sobre suas possibilidades, entendeu que, para ela, a interface demandava de tempo para aprendizado, o que caracteriza-se em uma interface não muito intuitiva.

Alguns pontos positivos apontados foram mais pontuais da interface, como, por exemplo, o fato de o mapa ser interativo (poder utilizá-lo como filtro), apresentar as rotas percorridas pela tocha e o fato de poder ver a ordenação através das tochas (segundo P5, "aqui eu já consigo ver a ordem daí eu já sei então "depois de Santa Catarina, ela foi para

o RS”)). Outras questões de mais alto nível também foram apontadas pelos participantes. Primeiramente, a forma de expressar as relações, através das múltiplas visualizações, e que, a partir da seleção de um filtro em uma visualização, afetava a informação que era apresentada nas demais visualizações, foi apontado como um ponto positivo. Além disso, o fato de as informações atualizarem-se conforme o usuário vai interagindo, também foi apontado como um ponto positivo; segundo o P5: *“os tweets vão atualizando conforme tu vais interagindo é muito bacana.”*. Por fim, quando o usuário utilizava algum filtro, como por data, ou por cidade, o mapa era direcionado para o ponto referente ao filtro selecionado (por exemplo, se um período de 3 dias era selecionado no filtro, o usuário era direcionado para a parte do mapa a qual a tocha percorreu durante esses dias). Este comportamento, que recebeu o nome de foco, também foi apontado como um ponto positivo, por dois participantes (P5 e P6).

Por fim, os participantes fizeram diversas sugestões para melhorar a interface. Primeiramente foi sugerida uma melhor exploração das cores, para dois fins distintos (para realce e para foco). O uso de cor para realce foi sugerido tanto para as informações das *popups* dos gráficos, que apareciam quando era clicado em algum ponto do gráfico (*“explorar mais as cores para essas informações que ficam mais escondidinhas, que tu tens que clicar em cima da cidade, talvez se ela fosse mais coloridinha eu visse ‘ah aqui está a quantidade”* - P1), ou para o uso nos gráficos para estes chamarem mais atenção (P3 e P6). Já no que se refere ao foco, P1 vislumbrou que, caso uma melhor exploração das cores fosse adotada, talvez ele conseguisse focar melhor nas informações desejadas.

Para P4, o tamanho da fonte utilizada na tabela dos *tweets* deveria ser maior, sendo que o tamanho atual dificultou a análise dessas informações. P3 vislumbrou que o realce por tamanho, agora nas visualizações, também poderia ser melhor explorado na interface. Outra sugestão que surgiu foi a utilização de gráficos por interesse do usuário. Para P2, talvez alguns usuários pudessem se interessar por informações mais quantitativas ou qualitativas, dependendo do seu perfil. Segundo a participante, talvez fosse interessante prover as visualizações que o usuário fosse, de fato, utilizar.

Algumas sugestões de novas formas de filtro também surgiram. Primeiramente, dado que a interface possuía *scroll*, foi sugerido que os filtros laterais ficassem sempre disponíveis no topo da interface, para facilitar o seu acesso. Além disso, foi sugerido filtro por cidades, de duas maneiras distintas: por um campo nos filtros laterais e pelo próprio mapa (ambas sugestões de P5). Foi sugerido, ainda, formas de limpar filtros específicos e formas para voltar ao estado anterior do sistema (também ambas por P5), conforme se dá a exploração da interface, para que a navegabilidade e o sequenciamento da interação fossem melhorados. Na filtragem de campos, foi mencionado que seria interessante utilizar sugestão dos valores, conforme o usuário vai digitando nos campos, dado que dificilmente o usuário irá saber os valores para filtrar. Haver a possibilidade de filtrar pelos *tweets* menos *twitados* também foi mencionado.

Além disso foi sugerido que houvesse uma *Word cloud* que apresentasse as cidades que houvessem mais *tweets* e que os gráficos de barras tivessem uma *label* que indicasse o valor (sem precisar interagir com o gráfico para descobrir o valor que estava sendo apresentado). No mapa, sugeriu-se que o usuário fosse limitado no *zoom*, forçando-o a ver apenas as partes do mapa nas quais o evento aconteceu e, na rota, apresentar a data ao lado do nome da cidade, ao invés de mostrar a ordem que a tocha percorreu.

7.3.3 Interatividade

As categorias extraídas referente a interatividade foram expectativas, sugestões, sentimentos e filtros. A categoria filtro foi dividida entre filtro lateral e visualizações. As categorias identificadas podem ser vistas na tabela 7.4. Na tabela, há apenas a estrutura hierárquica das categorias, entretanto a estrutura com as falas originárias, bem como os respectivos participantes, podem ser vistas no Apêndice F.

Tabela 7.4 – Estrutura de codificação para o código Interatividade - Etapa 1.

Estrutura da Codificação				
Interatividade	Expectativas		Navegação pelos gráficos	
			Acessar <i>Links</i> Existentes	
			Externo ao Software	
	Sugestões		Acesso aos <i>Links</i> Externos	
			Filtro visual para seleção de período	
			Ver a visualização maior em outra página	
	Sentimentos		Curiosidade	
			Desorientação	
			Motivação	
	Filtros	Laterais		Tarefa
				Questões Específicas
				Datas
				Completo/Possibilidades
Visualizações			Livre	
			Curiosidade	

A categoria expectativas diz respeito a ações que o usuário gostaria de ter executado, mas que o sistema não suportava. Estas, que foram indicadas por P1, foram a navegação no gráfico (ou seja, poder avançar a história mostrada no gráfico, pelo gráfico) e a possibilidade de acessar *links* externos existentes no *tweet* (alguns *tweets* apresentavam *links*, que, via interface, não eram possíveis acessar).

Algumas sugestões de melhoria também foram apontadas, no que se refere a interatividade, algumas delas, oriundas de expectativas não sanadas pelo software. A primeira é Acesso aos *Links* Externos presentes nos *tweets* (P1). Além disso, a única forma de filtrar por um determinado período era através dos filtros laterais. Nesse caso, P2 sugeriu que um

filtro visual para seleção de período fosse adicionado na interface. A participante, ainda, sugeriu que isto poderia ser feito através de um *slider*: *"Uma coisa que talvez pudesse ter era as barrinhas de tempo para arrastar o período de início e período de fim que daí é um filtro mais visual"*. Um filtro visual para data também foi sugerido por P3. Foi sugerido, ainda, que fosse possível detalhar uma determinada visualização em outra página (P4), para que esta fosse aberta em um tamanho maior, permitindo, com isso, uma melhor análise dos dados apresentados nessa visualização.

Outra categoria que emergiu a partir da análise desses dados foi sentimentos. Esta categoria diz respeito aos sentimentos instigados e apontados pelos próprios participantes, no decorrer do uso da interface. A partir de uma informação extraída pelos dados que estavam sendo apresentados, P1 ficou curiosa e quis saber mais detalhes sobre aquilo que lhe era apresentado na interface. Segundo a participante: *"aquela hora eu fiquei curiosa de saber porque tão falando de "Pato Branco" (cidade)"*. Outro sentimento apontado pela participante P2 foi motivação. Para ela, a forma que foram apresentadas as informações e os dados em si motivam a interação com a interface para extrair informação: *"Para um evento de olimpíada, muita gente se valeu disso para promover manifestação ou para promover o pokemon go, para várias coisas diferentes. E enxergar isso de uma forma que tu enxergues o todo te motiva sim a interagir e procurar"*. Por fim, outro sentimento que foi apontado por dois participantes foi desorientação. P2 acredita que o fato de haver muitos componentes e muitas possibilidades de interação possa desorientar o usuário na hora da análise dos dados. Já P4 foi mais precisa e citou em que momento se sentiu desorientada.

Por fim, no que se refere aos filtros presentes, haviam dois tipos, que eram por meio dos filtros laterais, e pelas visualizações presentes na interface. Para os participantes P2, P3 e P5, os filtros laterais eram úteis apenas para filtrar por datas. Já P1 apontou que os filtros laterais são úteis quando se tem tarefas a serem executadas, e, P2, para questões específicas. P6 considerou o filtro lateral como completo, possuindo muitas possibilidades de combinações de uso: *"Se eu quero um filtro mais completo, aí eu vou ali que tu consegue do jeito que tu quiser"*. Já os filtros pelas visualizações, segundo P1, são utilizados quando se há uma exploração livre (por curiosidade, de acordo com P2).

7.3.4 Dados

No que se refere a categoria Dados, foram extraídas 2 sub-categorias, que são Problemas e Sugestões. As categorias identificadas podem ser vistas na tabela 7.5. Na tabela, há apenas a estrutura hierárquica dos categorias, entretanto a estrutura com as falas originárias, bem como os respectivos participantes, podem ser vistas no Apêndice G.

O principal problema identificado (P3 e P6), foi referente aos fato de haver muitas informações repetidas. P3 mencionou que, por causa desses problema, não sentiu *"vontade*

Tabela 7.5 – Estrutura de codificação para o código Dados - Etapa 1.

Estrutura da Codificação		
Dados	Problemas	Informação Repetida
	Sugestões	Agrupar Informações de <i>Retweets</i>
		Mapa de Calor pelo Número de <i>Tweets</i>
		Mostrar Data e Hora dos <i>Tweets</i>

de ver". P6 também mencionou que, dado que os *tweets* possuíam a fonte muito pequena e que haviam muitos repetidos, não "*chamou a atenção em lê-los*".

No que se refere as sugestões, foi sugerido que os *tweets* fossem mostrados de forma agrupada, possivelmente pelo problema identificado nessa categoria (P3). Foi sugerido, ainda, que a data e a hora dos *tweets* fossem apresentadas na tabela. Por fim, P3 e P6 sugeriram utilizar mapas de calor para mostrar a quantidade de *tweets* no mapa.

7.4 Execução da Segunda Etapa dos Estudos

Após a coleta das primeiras impressões, foram feitos alguns ajustes na interface, com o objetivo de melhorar a sua usabilidade, a qual vinha afetando os participantes em focar-se na narrativa. Além disso, houve uma pequena alteração no procedimento, na etapa que era solicitado que o usuário realizasse as tarefas específicas: ao invés de solicitar que o participante buscasse no sistema uma série de informações, foi solicitado que o participante escolhesse um determinado estado e analisasse os dados que eram mostrado para aquele estado. Nessa parte da execução dos estudos, os participantes não deveriam interagir com os filtros. O motivo pelo qual esta parte da execução foi alterada foi para aprimorar os possíveis fatos que poderiam ser extraídos da interface (diferentes de informações isoladas). Por se entender que, estruturalmente, não houve mudanças, após esta alteração, não foi realizado um novo teste piloto.

O protocolo que foi utilizado para a realização desta segunda etapa de estudos (que, a partir desse momento, será referenciada como E2, e, a execução da primeira etapa, E1) pode ser verificado no Apêndice D.

Conforme mencionado previamente, após a coleta das impressões dos participantes, na primeira etapa dos estudos, alguns ajustes e correções de erros foram realizados na interface. Os ajustes foram feitos com base nos *feedback* dos participantes, sendo feitas alterações nos principais pontos levados, que puderam ser feitos.

Primeiramente, foi mudada a disposição dos componentes da interface. Alguns participantes apontaram que, no gráfico de linha, era difícil de selecionar o ponto desejado. Isto ocorria porque a largura do gráfico era muito pequena para a quantidade de pontos clicáveis. Para isso, foi aumentada a largura desse gráfico para que esta questão fosse

melhorada (Figura 7.3). A tabela que continha os *tweets* também teve aumento de tamanho, para que fosse possível adicionar informações de usuário, adicionar a possibilidade de acessar o *link* externo, bem como o tamanho da fonte aumentado, para facilitar a leitura desses *tweets* (Figura 7.4).

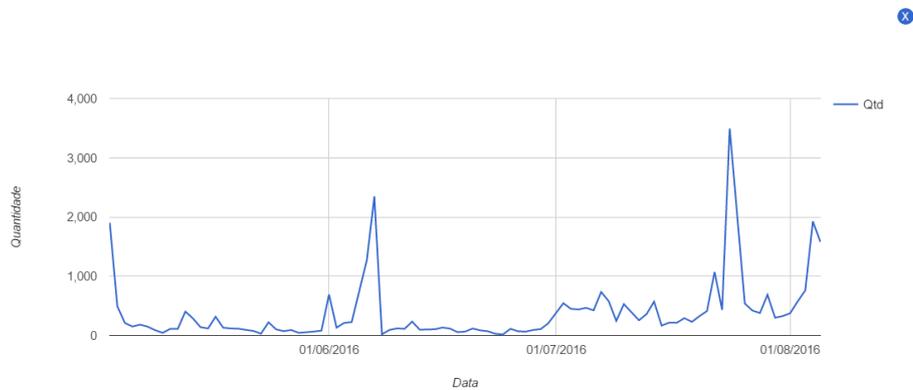


Figura 7.3 – Gráfico de linhas após as alterações.

Tweet	Usuário	Acessar Link
RT @Brasil2016: ??? HO-HO-HO...o penúltimo do #TourDaTocha é Elio Maria Lazzarotto, que trabalha há 34 anos como Papai Noel. https://t.co/QV...	paperbackscribe	Acessar Link Externo
O condutor final do #TourDaTocha em Pato Branco é o empresário Cláudio Petrycoski. https://t.co/tmb7jkC7wE	Brasil2016	Acessar Link Externo
Claudio já patrocinou campeonatos nacionais do "esporte da paz" e investiu na prática do Tchoukball #TourDaTocha https://t.co/GLk5PEwnIQ	Brasil2016	Acessar Link Externo
No futebol, apoiou o Pato Branco Futebol Clube por muitos anos, e investiu em outras modalidades #TourDaTocha https://t.co/QioB0vk3E9	Brasil2016	Acessar Link Externo
RT @Brasil2016: O condutor final do #TourDaTocha em Pato Branco é o empresário Cláudio Petrycoski. https://t.co/tmb7jkC7wE	allinesci	Acessar Link Externo
RT @Brasil2016: Claudio já patrocinou campeonatos nacionais do "esporte da paz" e investiu na prática do Tchoukball #TourDaTocha https://t...	allinesci	Acessar Link Externo
Petrycoski sempre ajudou diversas iniciativas comunitárias relacionadas à prática esportiva. #TourDaTocha https://t.co/587X1LxqZ	Brasil2016	Acessar Link Externo
RT @Brasil2016: No futebol, apoiou o Pato Branco Futebol Clube por muitos anos, e investiu em outras modalidades #TourDaTocha https://t.co/...	allinesci	Acessar Link Externo
Cláudio acendeu a pira com a @ChamaOlimpical #TourDaTocha encerrado em Pato Branco! https://t.co/biDJeTN0Ph	Brasil2016	Acessar Link Externo
RT @Brasil2016: Cláudio acendeu a pira com a @ChamaOlimpical #TourDaTocha encerrado em Pato Branco! https://t.co/biDJeTN0Ph	allinesci	Acessar Link Externo

Showing 1 to 10 of 2263 records

Figura 7.4 – Tabela com os *tweets* após as alterações.

Foi, também, adicionada uma *label* nos gráficos de barra, para especificar o valor que estava sendo apresentado. Além disso, no gráfico de barras das cidades, era apresentada a relação entre data e cidade. Em alguns casos, ocorria de a tocha percorrer, em um mesmo dia, cidades diferentes. Nesse caso, estas cidades eram apresentadas separadamente. Os participantes não percebiam que as datas eram as mesmas, mas com cidades distintas. A fim de explicitar que a tocha percorreu mais de uma cidade, no mesmo dia, agrupou-se as cidades, para serem apresentadas juntas. Ambas as alterações podem ser verificadas na figura 7.5.

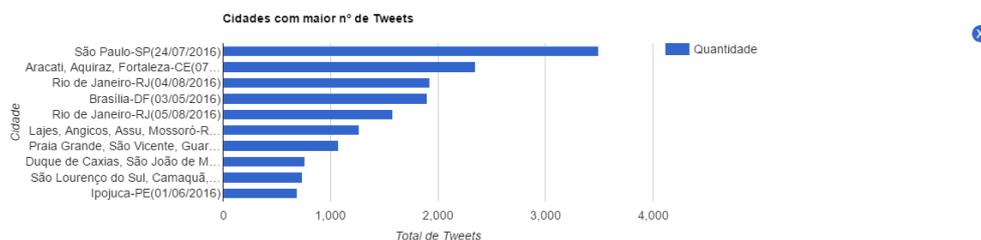


Figura 7.5 – Tabela com os *tweets* após as alterações.

Foi adicionada a possibilidade de limpar os filtros em todas as visualizações, para que o usuário pudesse voltar ao estado anterior, quando interagia com uma determinada visualização. Além disso, no mapa, o usuário tem a possibilidade de clicar em uma determinada cidade, e, quando isto é feito, informações de data e quantidade de *tweets* são apresentadas ao usuário. Na figura 7.6 é possível verificar ambas as mudanças (detalhe das cidades adicionado no mapa e a alteração referente à limpeza dos filtros, ilustrada no canto superior direito). Por fim, por conta de tais alterações, a disposição dos elementos presentes na interface teve que ser alterada. Na figura 7.7 é apresentada uma visão geral da interface, após as alterações descritas nesta seção.

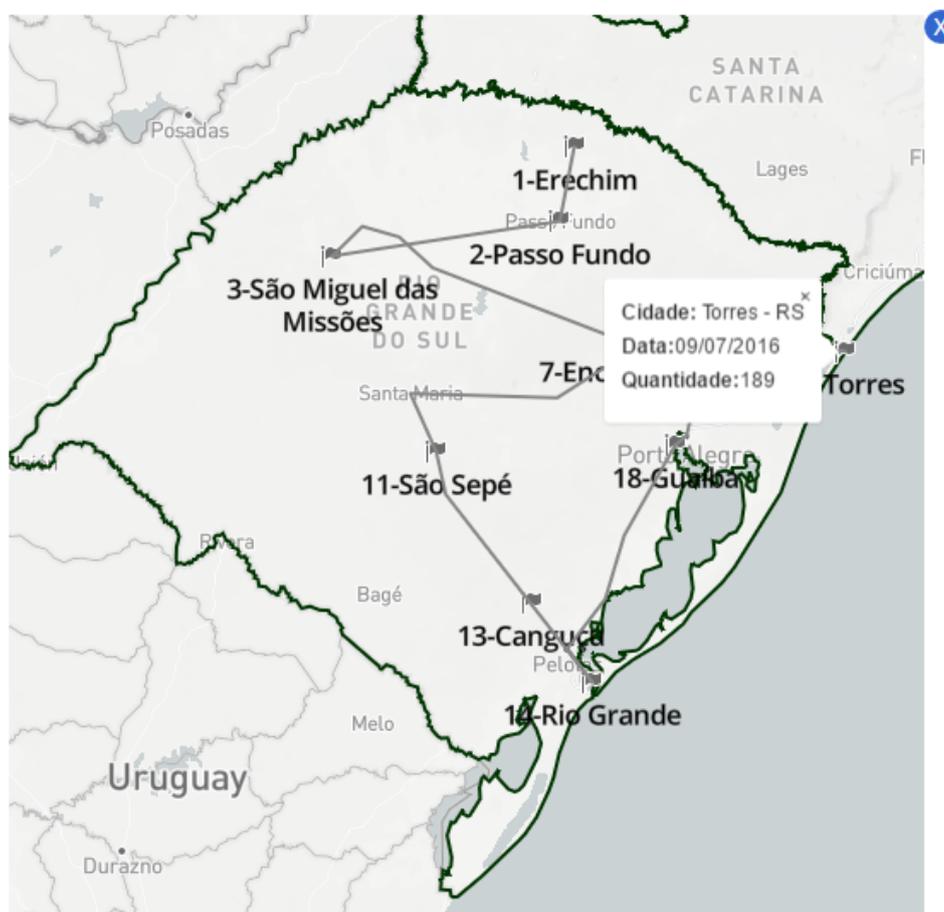


Figura 7.6 – Mapa com os detalhes e opção de limpeza das visualizações, após as alterações.

Participaram dessa etapa da pesquisa, ao total, 5 indivíduos, dos quais 4 eram do gênero masculino e 1 do gênero feminino. As idades variavam entre 20 e 35 anos. Todos os participantes consideravam-se usuários avançados de computadores, com exceção de uma, que considerava-se uma usuária intermediária. As redes sociais mais citadas foram Facebook, Twitter e Instagram, mas outras também apareceram, como Youtube, Snapchat e LinkedIn. Os principais motivos de uso de redes sociais apontados foram comunicação,

Tabela 7.6 – Perfil dos participantes da segunda etapa do estudo

Participante	Gênero	Idade	Escolaridade	Experiência no Uso de Computadores	Utiliza Redes Sociais
P7	Masculino	35	Mestrado (em andamento)	Avançado	Sim
P8	Feminino	20	Superior (em andamento)	Intermediário	Sim
P9	Masculino	20	Superior (em andamento)	Avançado	Sim
P10	Masculino	26	Superior Completo	Avançado	Sim
P10	Masculino	27	Mestrado (em andamento)	Avançado	Sim

Todos os participantes tinham conhecimento a respeito do evento. Dos fatos relacionados ao evento, que os participantes se lembravam, e que foram citados, além de pessoas caindo e das tentativas de apagar a tocha, foram o incidente que ocorreu na amazônia relacionado a onça abatida¹, sobre o fato do Pelé não ter aceitado participar do evento por questões de saúde² e sobre um incidente relacionado a assalto no RS.

Seguindo os mesmos passos da primeira etapa, após a coleta das informações iniciais de perfil, e após o uso do sistema por parte do participante, realizaram-se entrevistas, para coletar as impressões dos participantes a respeito da interface. Nessa etapa, também anotou-se impressões durante o uso do sistema, bem como analisou-se a gravação do uso da interface (gravação da tela do computador). A partir dessas fontes de dados, realizou-se o mesmo processo de codificação dos dados coletados. Chegou-se às mesmas 3 macrocategorias da etapa 1, que são Interface, Interatividade e Dados. As subseções seguintes apresentarão, primeiramente as mudanças de interface realizadas no sistema, e, após, os resultados da codificação para cada categoria.

7.4.1 Execução da Análise de um Estado

Após do uso livre da interface pelos participantes, era solicitado que eles selecionassem um determinado estado e analisassem as informações apresentadas. O objetivo dessa etapa era analisar as informações que os participantes poderiam extrair. Alguns participantes focaram sua análise em fatos específicos, e outros fizeram uma leitura geral das visualizações.

P7 selecionou o estado de Minas Gerais. Através da leitura dos *tweets*, ele foi capaz de verificar que houve pessoas que tentaram apagar a tocha, durante a passagem por Ouro Preto. O participante também tentou ver se haviam *hashtags* que indicassem algo relacionado e percebeu que não havia. Após isto, o participante selecionou a cidade, através do gráfico de barras e analisou as informações de uma maneira geral.

Já P8, que selecionou o estado de Santa Catarina, começou a analisar os usuários que postaram. Após verificar que não havia nenhum usuário desconhecido, percebeu um

¹<http://esportes.estadao.com.br/noticias/jogos-olimpicos,onca-e-abatida-durante-revezamento-da-tocha-e-gera-revolta-nas-redes,10000058417>

²<http://extra.globo.com/esporte/rio-2016/pele-nao-vai-acender-pira-olimpica-por-falta-de-condicoes-fisicas-19859328.html>

usuário que lhe chamou a atenção (usuário @NBA). O participante decidiu focar sua análise nesse usuário. Após selecionar o usuário, o participante leu os *tweets* e, então, analisou o resto das visualizações.

P9 selecionou o estado da Bahia. O participante focou sua análise nos usuários. Verificou que haviam alguns famosos e selecionou um dele (@ivetesangalo). Ao começar a ler os *tweets*, percebeu que haviam apenas menções a este usuário e não *tweets* do usuário em si. Por este motivo, decidiu parar de analisar.

P10 escolheu o estado de Tocantins. O participante analisou as visualizações de maneira geral, verificando as quantidades, as cidades e os horários. No gráfico de linha, tentou deduzir o motivo do pico (acreditava que era porque, naquele momento, o revezamento da tocha estava ocorrendo). Por fim, P11 selecionou Acre, e viu que haviam menções ao caso de uma onça que foi executada durante o percurso (lembrou que o caso ocorreu no estado do Amazonas, e deduziu que o assunto se propagou até os próximos estados). Após a verificação do fato, o participante concluiu sua análise.

7.4.2 Interface

No que se refere a interface, foram extraídas 6 sub-categorias, que são Características, Pontos Positivos, Pontos Negativos, Possibilidades de Uso, Sugestões e Sentimentos. As categorias identificadas podem ser vistas na tabela 7.7. Na tabela, há apenas a estrutura hierárquica dos categorias, entretanto a estrutura com as falas originárias, bem como os respectivos participantes, podem ser vistos no Apêndice H.

No que se refere as características, foram identificadas as mesmas 3 da primeira etapa do estudo, que são simples, fácil e amigável. Um participante apontou que a interface como um todo era fácil de ser utilizada (P11). Já outro participante (P8) apontou que, além da interface como um todo ser fácil, era fácil de limpar os filtros e mudar o filtro desejado. Segundo o usuário: *"É bem fácil de usar e de limpar os filtros e poder pesquisar outra coisa que tu queiras"*. A outra característica, nova até então, que foi identificada, foi que a interface possibilita a criação de hipóteses. Segundo o participante P11: *"...me instigou, tipo, que que aconteceu dia tal, porque tal hashtag está aparecendo, qual é a hashtag que foi usada... as perguntas emergem facilmente."*

Já no que se refere aos pontos positivos, o único ponto positivo que converge com os pontos positivos da etapa anterior foi a forma de apresentar as relações (que se davam a partir da interação com o sistema). Foi, também, elencado como ponto positivo o fato de se mostrar destaque a partir do tamanho (segundo P7, referindo-se às *word clouds*), de apresentar a relação entre os locais e o evento, que, segundo P10 é uma informação que, até então, ele nunca havia visto em redes sociais. Além disso, a quantidade de filtros e a possibilidades de combinação entre eles também foram apontados como pontos positivos.

Tabela 7.7 – Estrutura de codificação para o código Interface - Etapa 2.

Estrutura da Codificação		
Interface	Características Positivas	Simple
		Fácil
		Amigável
		Possibilita Criação de Hipóteses
	Pontos Positivos	Forma de mostrar Relação
		Destaque por Tamanho
		Relação entre Local e Evento
		Combinações de Filtros
		Formas de Interação
		Quantidades de Filtro
		Extrair Informações Úteis
	Pontos Negativos	Pouca Ajuda do Sistema
		Disposição dos Conteúdos
		<i>Scroll</i>
		Escala dos Gráficos Adaptáveis
	Possibilidade de Uso	Comportamento das Pessoas
		Saber o que as pessoas estão falando
		Informações Qualitativas
		Informações Quantitativas
	Sugestões	Informações mais Visíveis
		Cor para destaque
		<i>Feedback</i> Estado Atual do Sistema
		Indicar critérios de ordenação dos dados
Separar por tipo de informação		
Escalas únicas		
Agrupamento por Cor		
Sentimentos	Frustração	

Por fim, a forma de interação com a interface e a possibilidade de extrair informações úteis, foram os outros pontos positivos elencados (ambos identificados por P11).

Por outro lado, alguns pontos negativos também foram identificados pelos participantes. Primeiramente, segundo P7 e P8, o sistema não provê ajuda adequada. Segundo P7, "*faltou um pouco explicar o que cada gráfico realmente mostra*". Já P8 aponta que o sistema não informa as possibilidades presentes na interface, e que foi mais da intuição dele a forma de interagir: "*Acho que o ponto negativo né, foi muito da minha intuição e não o sistema me informando*". Outros dois pontos apontados foram a disposição dos conteúdos, que, embora apontado pelo P8, este não sabia alguma sugestão para melhorar. Além disso, o fato de ter *scroll* na interface também foi apontado como algo negativo: "*ter que ficar indo e vindo é ruim*". Por fim, o fato de as escalas serem adaptáveis, de acordo com o

conjunto de dados que estava sendo mostrado, dificultou a comparação desses dados com outros filtros.

Alguns participantes também vislumbraram possibilidades de uso para a visualização apresentada. P8 sugeriu que era possível utilizar o sistema para saber o que as pessoas estavam falando, além de poder ver informações qualitativas ou quantitativas, dado que existiam gráficos que apresentavam cada tipo de informação. Já P7 vislumbrou que a interface permite trabalhar com o comportamento das pessoas. Segundo o participante: "*Se eu quisesse trabalhar com o comportamento das pessoas acho que seria uma ferramenta válida para ser utilizada*".

No que se refere às sugestões dadas, duas delas foram identificadas, também, na primeira etapa, que são o uso de cores para destaque e a separação das visualizações por tipo de informação (na etapa 1, foi sugerido que fossem mostradas as visualizações de acordo com o tipo de informação que o usuário gostaria de analisar - no que se refere a informações qualitativas ou informações quantitativas). Já no que se refere as novas sugestões, a primeira foi que as informações ficassem sempre visíveis (P8), pois, no mapa, algumas informações eram suprimidas conforme o *zoom* era diminuído. Além disso, foi sugerido que fossem melhoradas as questões relacionadas ao *feedback* atual do sistema. P7 sugeriu que faltou apresentar informações referente ao que cada gráfico representa, e também que faltou indicar que quando é selecionado um filtro por alguma visualização, as demais são alteradas, adicionado o filtro selecionado. Já P9 sugeriu que os filtros laterais fossem melhor indicados.

Foi sugerido, ainda, que os critérios de ordenação dos *tweets* nas tabelas fossem melhores explicitados ao usuário (P10 não sabia que os *tweets* estavam ordenados por data). Além disso, P11 sugeriu utilizar escalas únicas, para facilitar a comparação com o *dataset* como um todo, não importando o filtro utilizado, e que as informações de um mesmo elemento (como, por exemplo, atores) fossem todas de uma mesma cor - sugerindo, com isso, um agrupamento por cor. Por fim, na categoria sentimento, P8 indicou que frustrou-se, pois, ao selecionar um usuário, acreditava que iria ler os *tweets* que este usuário postou, não mostrando aqueles aos quais o usuário foi mencionado.

7.4.3 Interatividade

As categorias extraídas, referentes a interatividade foram, sugestões e filtros. A categoria filtro foi dividida entre filtro lateral e visualizações. As categorias identificadas podem ser vistas na tabela 7.8. Na tabela, há apenas a estrutura hierárquica dos categorias, entretanto a estrutura com as falas originárias, bem como os respectivos participantes, podem ser vistos no Apêndice I.

Tabela 7.8 – Estrutura de codificação para o código Interatividade - Etapa 2.

Estrutura da Codificação			
Interatividade	Filtros	Laterais	Questões Específicas
		Visualizações	Exploração
	Sugestões		Filtrar por usuário que postou
			Filtrar cidade pelo mapa
		Locais das Postagens	

No que se refere aos filtros, as mesmas situações de uso que foram identificadas na etapa 1, foram identificadas nessa etapa. Os filtros laterais podem ser utilizados para questões específicas, e, os filtros inerentes às visualizações, para explorar de forma mais livre. Já as sugestões, houve uma que já havia sido citada previamente, que foi a possibilidade de filtrar as cidades pelo mapa. Outras duas sugestões foram apontadas pelos participantes. A primeira foi a possibilidade de filtrar pelo usuário que postou (P8), pois, para a usuária, seria interessante filtrar por usuário que originou os *tweets*. Além disso, P10 sugeriu apresentar, além do mapa que já era apresentado, um mapa com informações referentes ao local onde a postagem foi realizada e, que isto, seria uma informação bem relevante para ele.

7.4.4 Dados

No que se refere a categoria Dados, foram extraídas 2 sub-categorias, que são Problemas e Sugestões. As categorias identificadas podem ser vistas na tabela 7.9. Na tabela, há apenas a estrutura hierárquica dos categorias, entretanto a estrutura com as falas originárias, bem como os respectivos participantes, podem ser vistos no Apêndice J.

Tabela 7.9 – Estrutura de codificação para o código Dados - Etapa 2.

Estrutura da Codificação		
Dados	Sugestões	Agrupar Informações de <i>Retweets</i>
		<i>Tweets</i> mais Interagidos
	Problemas	<i>Spam</i>
		Informações não sucintas

No que se refere às sugestões, uma já havia sido mencionada na primeira etapa, que era agrupar os *tweets retuitados*. Entretanto P10 sugeriu que fosse mostrada uma lista de *tweets* mais importantes - ou seja, os *tweets* mais curtidos, *retwitados*, ou que fossem mais relevantes (não especificando qual critério de relevância). Já referente aos problemas identificados, P7 mencionou que havia muito *spam* (referindo-se a informações repetidas) e P10 mencionou que sentiu falta de informações mais sucintas. Segundo ele: "A questão dos *tweets* relacionados aqui ou dos *tweets* em geral que estão relacionados com os filtros... é

aquilo que eu falei. É legal de ver, mas eu não navegaria por 840 registros, para ver coisas, para procurar coisas que interessam".

7.5 Explorando a Narrativa

Através da análise da interação com a interface, além dos próprios relatos dos participantes, foi possível verificar que era possível dar diferentes focos na análise, bem como extrair informações distintas, a partir da implementação do modelo proposto. Na primeira etapa dos testes, além da exploração livre, os participantes também deveriam executar tarefas pontuais solicitadas pelo pesquisador.

Em todos os casos as informações foram obtidas de forma correta. Já na exploração livre, apesar dos diferentes níveis de detalhes informados por cada participante, no que se refere à informação extraída, todos os participantes mostraram-se dispostos a explorar a interface e tentar extrair informações úteis. Isto dá indícios que o modelo pode ser utilizado tanto para a análise de informações previamente desejadas (como, por exemplo, uma *hash-tag* específica ou algum usuário específico), quanto para descobrir, através da interação, o foco desejado pelo usuário.

Além disso, houve participantes que focaram sua análise em gráficos que apresentavam informações quantitativas (gráfico de barras e gráfico de linhas), e participantes que focaram sua análises em informações mais qualitativas, providas pelas *word clouds*. Já na etapa dois, quando era solicitado que o participante escolhesse um estado e analisasse as informações daquele estado no geral, era feita uma leitura dos gráficos, e das informações que eram apresentadas, não focando-se em algum fato específico.

Em alguns casos, os participantes pareciam apenas estarem interagindo e analisando as informações de forma isolada, não estabelecendo relações entre as visualizações apresentadas na tela. Além disso, alguns participantes procuravam contextualizar sua análise, a partir da leitura dos *tweets*. Entretanto, devido aos problemas apontados pelos participantes, como, por exemplo, informações repetidas e tamanho de fonte, eles não se motivavam em analisar os dados a partir dessa leitura.

No que tange aos focos dados pelos participantes, houve aqueles que buscaram usuários famosos (como uma participante de um *reality show* e de cantores que ficaram evidentes a partir dos dados). Outros participantes analisavam estados com os quais possuíam alguma relação, sendo o estado de nascimento ou estados com amigos próximos. Houve um participante que analisou os dados a partir de picos nos gráficos da interface. Houve também aqueles que descobriram as cidades pelas quais a tocha passou, bem como conseguiram verificar a ordem de passagem (e que, de fato, houve um certo sequenciamento lógico que a tocha percorreu entre os estados, que era no sentido anti-horário).

Outra usuária conseguiu vislumbrar que haviam diversas informações relacionadas à política, e, com isso, chegou a seguinte análise: *"Me chamou a atenção a questão política, como isso refletiu. Teve várias tags de reclamação eram o olimpíadas, rioproblems, teve várias coisas que tu vê que o pessoal, por essa visualização, que foi um evento diferente porque geralmente as pessoas tem um orgulho quando isso acontece... tu tens um orgulho do seu país sediar a copa, por exemplo. Mas foi um caso totalmente diferente, tu vê que está refletido aqui. Ela pode não ser a tag mais aparente, mas tem muitas tags do mesmo nicho. Se eu não tivesse no Brasil na época e visse o sistema, e não soubesse do cenário político, eu talvez conseguisse entender que tinha alguma coisa conflituosa. Isso me chamou bastante atenção. Eu sabia que era uma época conflituosa, mas não tinha essa noção que nas redes sociais, as pessoas tinham aproveitado o evento. Inclusive eu vi um post com as tags misturadas "olimpíadas" "foratemer" então vou aproveitar para pôr minha manifestação".* Outro participante conseguiu lembrar, a partir do refinamento dos seus filtros, um fato referente a onça que foi executada durante a passagem por Manaus ³.

A interface proposta, além de possuir diferentes possibilidades de análise, no que se refere ao tipo de informação a ser extraída, também possui variadas possibilidades de caminhos que o usuário pode percorrer. Acredita-se que a partir desse modelo, é possível atender demandas de diversos usuários, desde aqueles que já possuem uma ideia clara de análise, pois é possível que o usuário refine os resultados desejados de diversas maneiras e com diversas opções, mas também possibilita que usuários que não possuam conhecimento profundo a respeito dos dados, e nem interesse prévio em alguma informação específica, também consigam analisar os dados e extrair fatos a partir da interface. Acredita-se, ainda, que com refinamentos apresentados nas seções prévias, o modelo tem potencialidades para atingir resultados mais satisfatórios para a análise de dados.

7.6 Discussão

A análise dos dados obtidos na etapa dos estudos com usuários, possibilitou identificar as potencialidades do modelo proposto e alguns fatores que, entende-se, seriam mais significativos para a melhoria da interação por parte do usuário, com possíveis interfaces construídas a partir desse modelo.

Houveram sugestões que foram mais específicas da forma como determinados componentes foram incorporados na interface. Como exemplos, podem ser citadas as rotas disponíveis no mapa, a sugestão de filtrar por cidade ou rota, e de mostrar a data que a tocha estava naquela cidade, ao invés de mostrar por ordem. Além dessas, sugestões de

³<http://g1.globo.com/am/amazonas/noticia/2016/06/erramos-diz-rio-2016-apos-morte-de-onca-presente-em-tour-da-tocha.html>

implementação de determinadas visualizações (como mapa de calor, relacionando os locais com o número de *tweets*, ou uma *word cloud* com as cidades que mais houveram *tweets*).

A despeito das questões mais específicas, apontadas pelos participantes, houveram outras que foram mais significativas para a melhora do modelo. Primeiramente, a utilização das visualizações como forma de filtragem para direcionamento da narrativa, foi apontada, por diversos participantes, como um ponto positivo. Além disso, a quantidade de filtros e a possibilidade de combinação que a interface provia, também foi apontado como um fator importante para que os usuários extraíssem informações da interface. Nesse mesmo sentido, por possuir diversas opções de filtragem, e, também, diversas visualizações, identificou-se que a interface provia variadas possibilidades de caminhos a serem seguidos.

Outro fator interessante, foi a questão dos filtros. Na interface haviam duas possibilidades de utilização, que era através de campos e através das visualizações. Os participantes, sempre que iniciavam a exploração, de forma livre, utilizavam os filtros pelas visualizações. Quando era solicitado que fosse buscadas informações específicas, os filtros por campos eram utilizados. Entretanto, era possível selecionar o estado facilmente pelo mapa. Na maioria dos casos, os participantes optavam por utilizar o filtro pelo mapa, mesmo sendo apontado, por diversos participantes, que os filtros pelos campos eram úteis quando tarefas ou informações específicas eram solicitadas. Em um único caso, o participante selecionou o estado pelo campo, e justificou que era de sua preferência. Um participante sugeriu que fosse adicionada uma opção de "filtro visual" para seleção de período, pois era o único filtro que não havia duas possibilidades de uso. A questão do uso dos filtros visuais se mostrou importante, não somente para a utilização como um todo, mas também para facilitar a escolha do caminho que o usuário pudesse seguir, e pela facilidade de selecionar o dado que desejava utilizar. Esta facilidade também foi levantada como ponto positivo por parte dos participantes.

Alguns pontos, levantados pelos participantes, devem ser melhorados e explorados na construção da narrativa. Primeiramente, deve-se aprimorar a forma de orientação do usuário, relativo a questões de ajuda do sistema e na questão de sequenciamento da narrativa. O sequenciamento da narrativa refere-se a prover melhores mecanismos para a orientação do usuário na história, tanto nas formas de seguimento da narrativa, quanto a mecanismos para mostrar ao usuário em qual parte da narrativa ele se encontra no momento. Já na ajuda da interface, é importante que a interface forneça informações a respeito das possibilidades que ela possui, além de mostrar ajuda a respeito das informações que estão sendo apresentadas.

Muitos participantes sugeriram um melhor uso das cores na interface. O uso das cores, segundo os participantes, deveria ser usado tanto para destaque de determinadas informações, como para agrupar informações de um mesmo elemento (atores receberiam uma determinada cor e todas as visualizações referente a este elemento, utilizaria esta cor).

Além disso, sugeriu-se que fossem apresentados gráficos por tipos de usuários, ou seja, para aqueles que gostariam de ver informações quantitativas, seriam apresentadas apenas as visualizações que possuíam tais informações e o mesmo caso seria para aqueles que quisessem ver as informações qualitativas.

A forma que foi utilizada para expressar as relações entre as informações, na qual, a utilização de um filtro influenciava em todas as visualizações presentes na interface também foi apontado como um ponto positivo. Este recurso foi um dos mais importantes e foi um dos principais fatores que auxiliaram a manter a narrativa, com tantos componentes, de uma forma coerente, pois, mesmo com tantas possibilidades e opções, qualquer interação influenciaria no sistema como um todo, forçando que o usuário analisasse as visualizações de forma relacionada, e não de forma isolada.

Os participantes vislumbraram diversas possibilidades de uso para esta interface. Através dela é possível analisar informações quantitativas ou qualitativas, buscar tarefas específicas ou explorar de forma livre, indicando que o modelo pode ter variadas aplicações para diferentes públicos.

8. CONCLUSÃO

A produção de dados vem aumentando massivamente com o passar dos anos. Já é entendido que a informação é um dos recursos mais valiosos da sociedade [CC09]. Embora a quantidade de informações, atualmente, seja grande, fazer com que as informações relevantes estejam disponíveis para as pessoas exige grandes esforços, de diversas áreas de estudo. Dentre as áreas que auxiliam a atingir tal objetivo, está a de Visualização de Informações, área na qual o presente trabalho se encontra.

Recentemente, foi identificado que apresentar grandes volumes de dados, por meio de narrativas, é uma forma muito efetiva, no que se refere ao engajamento e à retenção das informações extraídas, a partir deste tipo de visualização [Fig14]. Com isso, cada vez mais se realizam estudos que fazem a intersecção entre ambas as áreas - Visualização de Informação e Narrativas - focando-os nos mais variados objetivos. Ao mesmo tempo, redes sociais são uma fonte de dados de muita valia, pois é possível extrair informações relacionadas a um determinado assunto, e com uma variação grande de usuários da rede. Dado este contexto, viu-se que uma oportunidade de pesquisa era o estudo da representação de dados extraídos de redes sociais que utilizasse elementos de narrativas para a sua concepção. Com isso, emergiu a seguinte questão, base para guiar o andamento da presente pesquisa: *"Como interfaces que utilizam narrativas com diferentes técnicas de visualização podem apoiar a interação do usuário e a sua análise de dados obtidos a partir de redes sociais?"*.

Com a questão de pesquisa formulada, identificou-se os objetivos principais da pesquisa. Resumidamente, objetivava-se, ao final deste trabalho, um melhor entendimento de como se dá a utilização de técnicas de visualização para construir uma narrativa e uma proposta de representação de dados de redes sociais neste contexto.

Primeiramente, identificou-se que já existia um *framework* concebido a partir da análise de diversas narrativas, para portais de notícias [SH10]. Este *framework* - que os autores chamaram de espaço de *design* - tem sido amplamente usado na literatura relacionada e foi fundamental para o entendimento dos componentes de interface necessários para a construção de narrativas com visualizações. Tendo ele como base, analisou-se outros trabalhos que implementaram narrativas com visualizações, verificando-se as técnicas que foram utilizadas por seus autores. Com isso, foi possível, além do entendimento das técnicas para construção de narrativas, uma extensão do atual espaço de *design*, com novas técnicas identificadas a partir desta análise.

A partir do entendimento do conceito de narrativa, fora do âmbito de visualização de dados, verificou-se que narrativas possuem um conjunto de elementos que as compõem. A composição e a relação entre esses elementos formam a narrativa em si. Existem diversos conjuntos de elementos existentes e, para esta pesquisa, escolheu-se aquele que era

composto pelos elementos eventos, tempo, atores e localização, dado que este conjunto já havia sido utilizado no campo de visualização e, também, por serem facilmente obtidos com os dados que iriam ser utilizados. Escolhido o conjunto de elementos que iria ser utilizado, foi proposta uma forma de extrair cada elemento de uma postagem de rede social.

O objetivo de desenvolvimento de uma forma de representação dos dados com redes sociais, utilizando elementos de narrativas, resultou em um modelo, composto por sete componentes, que são dados, elementos de narrativa, informações, visualizações, gênero, elementos visuais e elementos estruturais. A concepção do modelo se deu a partir do estudo da área de narrativas, bem como de trabalhos relacionados à visualização de dados e a narrativas. Após a identificação dos componentes do modelo, analisou-se as possibilidades de cada componente, e definiu-se os limites, no que se refere as técnicas que deveriam ser focadas. Definido o modelo, criou-se uma instância desse modelo, uma interface web que implementa partes de cada componente do modelo, com dados extraídos da rede social Twitter. Tal interface foi utilizada para análise do modelo com usuários.

Um das questões que foram identificadas a partir dos estudos realizados com os usuários foi a identificação das diversas possibilidades de uso do modelo. A instância do modelo conta com visualização de diversos tipos de informação, propiciando que os usuários analisassem tipos de informações diferentes. Além disso, pode ser utilizada tanto para a exploração de informações específicas, a partir de seus filtros, quanto para uma exploração mais livre, com a descoberta de informações sem um conhecimento ou interesse prévio nos dados associados. De fato, este era um dos objetivos do modelo: ao mesmo tempo que não exigiu um usuário especialista e/ou com interesse prévio para extrair informações por meio da interface, supriu a necessidade de usuários com questões mais pontuais ou algum interesse prévio, no que se refere a extração de informações.

Este estudo revelou, também, que a utilização dos filtros, através das visualizações, foi um fator determinante para que os usuários interagissem com a interface. Se, por um lado, haviam muitas possibilidades para o usuário, no que se refere aos caminhos e interação com o sistema, por outro, o fato da interação em um componente específico (seleção de algum critério de filtro, a partir de qualquer forma existente) influenciar as informações apresentadas em todas visualizações forçava que o usuário analisasse os dados providos nas visualizações como um todo, e não como componentes isolados.

Assim, diferente das visualizações com narrativas existentes, que possuem um conjunto de dados muito mais focado, o presente trabalho utilizou um *dataset* amplo e que possibilitava que os usuários focassem em quaisquer informações que desejassem, podendo, este foco, ser totalmente diferente de um usuário para outro.

Por fim, no decorrer da pesquisa, algumas limitações foram identificadas. Primeiramente, houve limitações técnicas, principalmente na extração dos elementos de narrativa e no tratamento dos dados. Há diversos recursos de processamento de linguagem natural (PLN) que poderiam ser utilizados para extração de informações mais precisas destes ele-

mentos (como recursos semânticos e de identificação de eventos). Além disso, por se tratar de dados extraídos de redes sociais, muitas vezes era utilizada uma linguagem não-formal, contendo gírias ou outras estruturas de linguagem, que dificultavam o uso de recursos de PLN, por trabalharem apenas com a linguagem formal. Mas como o foco do trabalho não era PLN, estes pontos ficaram para próximos passos da pesquisa.

Já no que se refere ao método de análise adotado, este sugere que, para minimizar o viés, o processo de codificação seja feito por mais de uma pessoa, para determinar o quão válido e confiável são as categorias extraídas. Dado que a presente pesquisa foi realizada apenas por um pesquisador, este processo não pode ser realizado. Além disso, houve limitação no número de usuários que participaram dos estudos, além de não haver variação com o tipo de amostra.

A partir destes pontos e de reflexões sobre a pesquisa como um todo, os seguintes trabalhos futuros foram elencados:

- Implementação das melhorias de interface elencadas pelos usuários, com foco em questões relacionadas a ajuda do sistema, a formas de sequenciamento da narrativa e a melhor uso das cores, principalmente para realce de informações;
- Utilização de recursos de PLN a fim de melhorar a forma de extração das informações, relacionadas aos elementos da narrativa, bem como melhorar o processo de limpeza dos dados;
- Realização de estudos com outros perfis de usuários, como, por exemplo, pessoas que tem conhecimento profundo de construção de narrativas ou profissionais especializados em análise de dados;
- Estudo de outros possíveis contextos aos quais este modelo pode ser aplicado e construção de novas instâncias para posterior análise.

Por fim, acredita-se que os objetivos propostos neste trabalho foram atingidos, sendo sua principal contribuição o apoio à construção de narrativas interativas, por meio de visualizações com dados extraídos de redes sociais, contemplando, não somente o apoio à construção da interface em si, mas, também, à estruturação dos dados e a definição de sua representação.

REFERÊNCIAS BIBLIOGRÁFICAS

- [And06] Anderson, E. “Iris data set parallel coordinates”. Capturado em: <https://goo.gl/nVwH3D>, Acessado em: Fevereiro de 2017.
- [Bal09] Bal, M. “Narratology: Introduction to the theory of narrative”. University of Toronto Press, 2009, 264p.
- [BAMQ06] Bloch, M.; Amanda, C.; McGinty, J. C.; Quea, K. “A peek into netflix queues”. Capturado em: <https://goo.gl/sxRkrL>, Acessado em: Fevereiro de 2017.
- [BDVA+15] Balduini, M.; Della Valle, E.; Azzi, M.; Larcher, R.; Antonelli, F.; Ciuccarelli, P. “Citysensing: Fusing city data for visual storytelling”, *IEEE MultiMedia*, vol. 22–3, 2015, pp. 44–53.
- [Bel09a] Belmonte, N. “#eleições2014 - análise do crescimento de seguidores dos candidatos à presidência”. Capturado em: <https://interactive.twitter.com/brazil2014/>, Acessado em: Fevereiro de 2016.
- [Bel09b] Belmonte, N. “#euro2012: A summary for the action on twitter during the european football tournament”. Capturado em: <https://euro2012.twitter.com/>, Acessado em: Fevereiro de 2016.
- [Bel09c] Belmonte, N. “#sotu2015: See the state of the union address minute by minute on twitter”. Capturado em: <http://twitter.github.io/interactive/sotu2015/#p1>, Acessado em: Fevereiro de 2016.
- [Ber06] Berg, S. “O que é gestalt?” Capturado em: <https://goo.gl/C7RPLA>, Acessado em: Fevereiro de 2017.
- [Ber11] Bertin, J. “Semiology of Graphics: Diagrams, Networks, Maps”. ESRI Press, 2011, 438p.
- [Ber12] Berg, M. “Human abilities to perceive, understand, and manage multi-dimensional information with visualizations”, Tese de doutorado, Aalto University - Department of Media Technology, Greater Helsinki, Finlandia, 2012, 101p.
- [BKH+16] Bach, B.; Kerracher, N.; Hall, K. W.; Carpendale, S.; Kennedy, J.; Riche, N. H. “Telling stories about dynamic networks with graph comics”. In: Proceedings of the Conference on Human Factors in Information Systems (CHI), 2016, pp. 3670–3682.
- [Bos06a] Bostock, M. “Flare imports hierarchical edge bundling”. Capturado em: <https://goo.gl/JUcDFb>, Acessado em: Fevereiro de 2017.

- [Bos06b] Bostock, M. “Radial tidy tree”. Capturado em: <https://goo.gl/iPV5i6>, Acessado em: Fevereiro de 2017.
- [BS06] Blake, R.; Sekuler, R. “Perception”. McGraw-Hill Companies, Incorporated, 2006, 652p.
- [Car06] Carto. “Madrid subway complaints by station”. Capturado em: <https://goo.gl/823DOZ>, Acessado em: Fevereiro de 2017.
- [CC09] Costa, P. L. V. A.; Carvalho, A. G. “Sociedade da informação: avanços e retrocessos no acesso e no uso da informação”, *Informação & Sociedade*, 2009.
- [CM11a] Cruz, P.; Machado, P. “Generative storytelling for information visualization”, *IEEE Computer Graphics and Applications*, vol. 31–2, 2011, pp. 80–85.
- [CM11b] Cruz, P.; Machado, P. “Visualizing empires decline: Applying generative storytelling for information visualization”. Capturado em: <https://src.acm.org/binaries/content/assets/src/2010/pedrocruz.pdf>, Acessado em: Janeiro de 2017.
- [COT97] Campo, M.; Orosco, R.; Teyseyre, A. “Automatic abstraction management in information visualization systems”. In: IEEE Conference on Information Visualization, 1997, pp. 50–56.
- [Cra00] Craine, K. “Designing a document strategy”. Mc2 Books, 2000, 166p.
- [EAB13] Elias, M.; Aufaure, M.-A.; Bezerianos, A. “Storytelling in visual analytics tools for business intelligence”. In: Proceedings of the IFIP TC.13 International Conference on Human-Computer Interaction (INTERACT), 2013, pp. 280–297.
- [Ell99] Ellis, W. D. “A Source Book of Gestalt Psychology”. Routledge, 1999, 403p.
- [Est06] Estadão. “Criminalidade bairro a bairro”. Capturado em: <https://goo.gl/5U4uOX>, Acessado em: Fevereiro de 2017.
- [FAC16] “Facebook reports first quarter 2016 results”. Capturado em: <https://investor.fb.com/investor-news/default.aspx>, Acessado em: Julho de 2016.
- [FCLC01] Freitas, C. M. D. S.; Chubachi, O. M.; Luzzardi, P. R. G.; Cava, R. A. “Introdução à visualização de informações”, *Revista de Informática Teórica e Aplicada*, 2001, pp. 143–158.
- [Fig14] Figueiras, A. “How to tell stories using visualization”. In: Proceedings of the International Conference on Information Visualisation (IV), 2014, pp. 18–18.
- [Gan04] Gancho, C. V. “Como analisar narrativas”. Editora Ática, 2004, 79p.

- [Gib09] Gibbs, G. “Análise de Dados Qualitativos”. Artmed Editora S.A., 2009, 198p.
- [Gol09] Goldstein, E. B. “Sensation and Perception”. Cengage Learning, 2009, 496p.
- [GP01] Gershon, N.; Page, W. “What storytelling can do for information visualization”, *Communications of the ACM*, vol. 44–8, 2001, pp. 31–37.
- [Gua06a] Guardian, T. “Iran’s nuclear programme”. Capturado em: <https://goo.gl/nEIXPc>, Acessado em: Fevereiro de 2017.
- [Gua06b] Guardian, T. “Oscars 2010 best picture nominees”. Capturado em: <https://goo.gl/KFiYSI>, Acessado em: Fevereiro de 2017.
- [HB05] Heer, J.; Boyd, D. “Vizster: Visualizing online social networks”. In: Proceedings of the Proceedings of the Symposium on Information Visualization (INFOVIS), 2005, pp. 32–39.
- [HDR+13] Hullman, J.; Drucker, S.; Riche, N. H.; Lee, B.; Fisher, D.; Adar, E. “A deeper understanding of sequence in narrative visualization”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 19–12, 2013, pp. 2406–2415.
- [HFM07] Henry, N.; Fekete, J.-D.; McGuffin, M. J. “Nodetrix: a hybrid visualization of social networks”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 13–6, 2007, pp. 1302–1309.
- [HMM00] Herman, I.; Melançon, G.; Marshall, M. S. “Graph visualization and navigation in information visualization: A survey”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 6–1, 2000, pp. 24–43.
- [Hol69] Holsti, O. “Content analysis for the social sciences and humanities”. Addison-Wesley Pub. Co., 1969, 235p.
- [Hol06] Holten, D. “Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 12–5, 2006, pp. 741–748.
- [IGS06] IGSP. “Como funciona o alerta de tsunami?” Capturado em: <https://goo.gl/55KJf4>, Acessado em: Fevereiro de 2017.
- [Ins16] Instagram. “Instagram press”. Capturado em: <https://www.instagram.com/press/>, Acessado em: Julho de 2016.
- [Int09a] InteractiveTwitter. “Election2016: Us presidential candidate twitter buzz”. Capturado em: <https://interactive.twitter.com/candidateRace16/>, Acessado em: Fevereiro de 2016.

- [Int09b] InteractiveTwitter. “How every #gameofthrones episode has been discussed on twitter”. Capturado em: <https://interactive.twitter.com/game-of-thrones/#?episode=60>, Acessado em: Fevereiro de 2016.
- [Int09c] InteractiveTwitter. “#worldcup: Beyond the home team”. Capturado em: <https://interactive.twitter.com/wcmentions/#BRA>, Acessado em: Fevereiro de 2016.
- [JS91] Johnson, B.; Shneiderman, B. “Tree-maps: A space-filling approach to the visualization of hierarchical information structures”. In: Proceedings of the Conference on Visualization, 1991, pp. 284–291.
- [Kan03] Kanashiro, M. “A cidade e os sentidos: sentir a cidade”, *Desenvolvimento e meio ambiente*, vol. 7, 2003, pp. 155–160.
- [Kir12] Kirk, A. “Data Visualization: a successful design process”. Packt Publishing, 2012, 206p.
- [KK11] Khan, M.; Khan, S. S. “Data and information visualization methods, and interactive mechanisms: A survey”, *International Journal of Computer Applications*, vol. 34–1, 2011, pp. 1–14.
- [KM13] Kosara, R.; Mackinlay, J. “Storytelling: The next step for visualization”, *Computer*, vol. 46–5, 2013, pp. 44–50.
- [Kri06] Kriebel, A. “How common is your birthday? find out exactly with an interactive heat map”. Capturado em: <https://goo.gl/tqwCJS>, Acessado em: Fevereiro de 2017.
- [KS12] Kuhn, A.; Stocker, M. “Codetimeline: Storytelling with versioning data”. In: Proceedings of the International Conference on Software Engineering (ICSE), 2012, pp. 1333–1336.
- [LFH10] Lazar, J.; Feng, J. H.; Hochheiser, H. “Research methods in human-computer interaction”. John Wiley & Sons, 2010, 560p.
- [LHB10] Lidwell, W.; Holden, K.; Butler, J. “Universal Principles of Design, Revised and Updated: 125 Ways to Enhance Usability, Influence Perception, Increase Appeal, Make Better Design Decisions, and Teach Through Design”. Rockport Publishers, 2010, 272p.
- [LR96] Lamping, J.; Rao, R. “The hyperbolic browser: A focus+ context technique for visualizing large hierarchies”, *Journal of Visual Languages & Computing*, vol. 7–1, 1996, pp. 33–55.

- [LRIC15] Lee, B.; Riche, N. H.; Isenberg, P.; Carpendale, S. “More than telling a story: A closer look at the process of transforming data into visually shared stories”, *IEEE Computer Graphics and Applications*, vol. 35–5, 2015, pp. 84–90.
- [LRU14] Leskovec, J.; Rajaraman, A.; Ullman, J. D. “Mining of massive datasets”. Cambridge University Press, 2014, 326p.
- [LWW⁺13] Liu, S.; Wu, Y.; Wei, E.; Liu, M.; Liu, Y. “Storyflow: Tracking the evolution of stories”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 19–12, 2013, pp. 2436–2445.
- [Mas13] Massarolo, J. C. “Storytelling transmídia: narrativa para múltiplas plataformas”, *Tríade: Revista de Comunicação, Cultura e Mídia*, vol. 1–2, 2013, pp. 335–347.
- [MBB⁺12] Marcus, A.; Bernstein, M. S.; Badar, O.; Karger, D. R.; Madden, S.; Miller, R. C. “Processing and visualizing the data in tweets”, *ACM SIGMOD Record*, vol. 40–4, 2012, pp. 21–27.
- [MG05] Mackey, A.; Gass, S. M. “Second Language Research: Methodology and Design”. Routledge, 2005, 452p.
- [Mus09] Mustilica. “Trending topics within last 24 hours in worldwide”. Capturado em: <http://tt-history.appspot.com/>, Acessado em: Fevereiro de 2016.
- [NCo06] NComVA. “Ncomva dashboard demo”. Capturado em: <https://goo.gl/KQCPAa>, Acessado em: Fevereiro de 2017.
- [Nod09] Nod, B. “Find influencers with twitter visualization”. Capturado em: <http://bluenod.com/>, Acessado em: Fevereiro de 2016.
- [Nöl07] Nöllenburg, M. “Geographic Visualization”. Springer Berlin Heidelberg, 2007, pp. 257–294.
- [Pen06] PennState. “Choropleth maps”. Capturado em: <https://goo.gl/VjiCil>, Acessado em: Fevereiro de 2017.
- [Pin06] Pinna, D. M. d. S. “Animadas personagens brasileiras: a linguagem visual das personagens do cinema de animação contemporâneo brasileiro”, Dissertação de mestrado, Departamento Artes e Design, PUC-Rio, Rio de Janeiro, Brasil, 2006, 230p.
- [Pos09] Post, W. “On the map: Five major north korean prison camps”. Capturado em: <http://www.washingtonpost.com/wp-srv/special/world/north-korean-prison-camps-2009/>, Acessado em: Fevereiro de 2017.

- [Pro09] Prod, J. "Twitter network". Capturado em: <https://twitter.jeffprod.com/demo.php>, Acessado em: Fevereiro de 2016.
- [PUC06] PUC-Rio. "Dados, informação e conhecimento". Capturado em: <https://goo.gl/nQguf2>, Acessado em: Fevereiro de 2017.
- [RCWR14] Regattieri, L.; Chartier, R.; Windsor, J.; Rockwell, G. "Tweeviz: Following twitter hashtags to support storytelling". In: Hypertext and Social Media Conference (HYPERTEXT), 2014, pp. 52–56.
- [RND15] Rodríguez, M. T.; Nunes, S.; Devezas, T. "Telling stories with data visualization". In: Proceedings of the Workshop on Narrative and Hypertext (NHT), 2015, pp. 7–11.
- [Ros06] Rosenthal, S. "Football drawings". Capturado em: <https://goo.gl/nncD8P>, Acessado em: Fevereiro de 2017.
- [Ros10] Rosenbaum, S. "Seth godin: Mark cuban is completely wrong about aggregators". Capturado em: <https://goo.gl/TUyTss>, Acessado em: Julho de 2016.
- [SDM14] Stojanovski, D.; Dimitrovski, I.; Madjarov, G. "Tweeviz: Twitter data visualization". In: Proceedings of the Conference on Data Mining and Data Warehouses (SiKDD), 2014.
- [Sen09] SentimentViz. "Sentiment viz - tweet sentiment visualization". Capturado em: https://www.csc2.ncsu.edu/faculty/healey/tweet_viz/tweet_app/, Acessado em: Fevereiro de 2016.
- [SH10] Segel, E.; Heer, J. "Narrative visualization: Telling stories with data", *IEEE Transactions on Visualization and Computer Graphics*, vol. 16–6, 2010, pp. 1139–1148.
- [SH14] Satyanarayan, A.; Heer, J. "Authoring narrative visualizations with ellipsis", *Computer Graphics Forum*, vol. 33–3, 2014, pp. 361–370.
- [SI10] Steele, J.; Iliinsky, N. "Beautiful visualization: looking at data through the eyes of experts". O'Reilly Media, 2010, 418p.
- [SLRS16] Stolper, C. D.; Lee, B.; Riche, N. H.; Stasko, J. "Emerging and recurring data-driven storytelling techniques: Analysis of a curated collection of recent stories", Relatório Técnico, Microsoft Research Department, 2016, 14p.
- [Sto06] Stone, M. "Choosing colors for data visualization". Capturado em: <https://goo.gl/L6igZ6>, Janeiro de 2016.

- [Tim06a] Times, F. “What happens when a country defaults”. Capturado em: <https://goo.gl/9FVnVL>, Acessado em: Fevereiro de 2017.
- [Tim06b] Times, N. Y. “Paths to the top of the home run charts”. Capturado em: <https://goo.gl/7Wa8uW>, Acessado em: Fevereiro de 2017.
- [Twi16] Twitter. “Twitter usage statistics”. Capturado em: <https://goo.gl/e5pgJh>, Acessado em: Julho de 2016.
- [War12] Ware, C. “Information visualization: perception for design”. Elsevier, 2012, 513p.
- [Wei09] Weigend, A. “The social data revolution(s)”. Capturado em: <https://www.instagram.com/press/>, Acessado em: Julho de 2016.
- [WLY+14] Wu, Y.; Liu, S.; Yan, K.; Liu, M.; Wu, F. “Opinionflow: Visual analysis of opinion diffusion on social media”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 20–12, 2014, pp. 1763–1772.
- [Won09] Wongsuphasawat, K. “Premier league: Where are your club’s followers?” Capturado em: <https://interactive.twitter.com/premierleague/#?mode=team&teamId=all>, Acessado em: Fevereiro de 2016.
- [Wor06] “The new worldmapper”. Capturado em: <https://goo.gl/jKnHb3>, Acessado em: Fevereiro de 2017.
- [WR12] Weber, W.; Rall, H. “Data visualization in online journalism and its implications for the production process”. In: Proceedings of the International Conference on Information Visualisation (IV), 2012, pp. 349–356.
- [Yau11] Yau, N. “Visualize This: The FlowingData Guide to Design, Visualization, and Statistics”. John Wiley & Sons, 2011, 384p.
- [YHYO01] Yoshioka, T.; Herman, G.; Yates, J.; Orlikowski, W. “Genre taxonomy: A knowledge repository of communicative actions”, *ACM Transactions on Information Systems*, vol. 19–4, 2001, pp. 431–456.
- [ZCW+14] Zhao, J.; Cao, N.; Wen, Z.; Song, Y.; Lin, Y.-R.; Collins, C. “#fluxflow: Visual analysis of anomalous information spreading on social media”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 20–12, 2014, pp. 1773–1782.

APÊNDICE A – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Modelo para Construção de Narrativas com Dados Extraídos de Redes Sociais

Faculdade de Informática/PUCRS
Avenida Ipiranga, 6681 – Prédio 32 - 90619-900 – Porto Alegre – RS
Tel: (51) 3320-3558

Termo de Consentimento Livre e Esclarecido

Agradecemos a todos os participantes desta pesquisa, a inestimável contribuição que prestam para o avanço da pesquisa sobre Interação Humano-Computador.

O objetivo desta pesquisa analisar as possibilidades de construção de narrativas a partir das visualizações apresentadas. Para isto, os participantes dos testes são convidados a utilizar o sistema interativo, enquanto são observados por um ou mais pesquisadores. Esta observação será registrada em papel e, também, através de um software de captura de telas, que armazena tudo o que acontece na tela do computador, e, eventualmente, através de vídeo. Estas informações nos trarão dados importantíssimos para verificar a qualidade dos sistemas em questão.

Lembramos que o objetivo deste estudo não é avaliar o participante, mas sim avaliar o aplicativo computacional que o participante estará usando durante os testes. O uso que se faz dos registros efetuados durante o teste é estritamente limitado a atividades de pesquisa e desenvolvimento, garantindo-se para tanto que:

1. O anonimato dos participantes será preservado em todo e qualquer documento divulgado em foros científicos (tais como conferências, periódicos, livros e assemelhados) ou pedagógicos (tais como apostilas de cursos, *slides* de apresentações, e assemelhados).
2. Todo participante terá acesso a cópias destes documentos após a publicação dos mesmos.
3. Todo participante que se sentir constrangido ou incomodado durante uma situação de teste pode interromper o teste e estará fazendo um favor à equipe se registrar por escrito as razões ou sensações que o levaram a esta atitude. A equipe fica obrigada a descartar o teste para fins da avaliação a que se destinaria.
4. Os participantes que forem menores de idade terão, obrigatoriamente, que apresentar o consentimento de seu responsável, para participação no estudo, o qual será declarado ciente do estudo a ser realizado através de sua assinatura no presente Termo de Consentimento.
5. Todo participante tem direito de expressar por escrito, na data do teste, qualquer restrição ou condição adicional que lhe pareça aplicar-se aos itens acima enumerados (1, 2, 3 e 4). A equipe se compromete a observá-las com rigor e entende que, na ausência de tal manifestação, o participante concorda que rejam o comportamento ético da equipe somente as condições impressas no presente documento.
6. A equipe tem direito de utilizar os dados dos testes, mantidas as condições acima mencionadas, para quaisquer fins acadêmicos, pedagógicos e/ou de desenvolvimento contemplados por seus membros.

[a ser preenchido pelo observador]
Sistema: _____ Data: __/__/____
Condições especiais (caso não haja condições especiais, escreva "nenhuma"):

<input type="checkbox"/> continua no verso

Por favor, indique sua posição em relação aos termos acima:

- Estou de pleno acordo com os termos acima.
 Em anexo registro condições adicionais para este teste.

Assinatura do participante

Assinatura do responsável
(caso o participante seja menor de idade)

Assinatura do observador

Nome do Participante: _____

Nome do Responsável (se o participante for menor de idade): _____

Pesquisador Responsável: Prof. _____ – Faculdade de Informática - PUCRS

APÊNDICE B – PROTOCOLO DE ANÁLISE ORIGINAL

1. Introdução dos objetivos da pesquisa aos participantes, bem como a apresentação e a assinatura do Termo de Consentimento Livre e Esclarecido (Apêndice A).
2. Questões referente ao perfil do participante, sobre o uso de sistemas para visualização de informação, uso de redes sociais e conhecimento sobre o evento do Revezamento da Tocha Olímpica:
 - (a) Gênero
 - (b) Idade
 - (c) Escolaridade
 - (d) Experiência em computação?
 - (e) Utiliza redes sociais?
 - (f) Caso positivo para pergunta anterior, quais e com qual finalidade?
 - (g) Já utilizou algum sistema que oferecesse funcionalidades para visualizar dados grandes volumes de dados?
 - (h) Caso positivo para pergunta anterior, qual sistema e para qual finalidade?
 - (i) Já sentiu necessidade de utilizar algum sistema que oferecesse funcionalidades para visualizar grandes volumes de dados?
 - (j) Caso positivo para pergunta anterior, para qual finalidade?
 - (k) Você tem conhecimento à respeito do revezamento da tocha olímpica ?
 - (l) Caso positivo, existem algum fato que você se lembra que ocorreu, relacionado ao revezamento da tocha?
3. Demonstração do sistema proposto, das formas de interação com ele, do que estava sendo mostrado em cada visualização. Era apresentado ao usuário, de maneira geral, as opções disponíveis ao usuário, como interagir com cada visualização, quais eram as formas de filtro e o que estava sendo representado na interface.
4. Execução, por parte do usuário, da primeira etapa do uso da interface proposta. Nesta etapa, os usuários eram livres para interagir na interface, até que achassem pertinente. Os usuários poderiam solicitar ajuda ou tirar dúvidas a qualquer momento, durante o uso.
5. Pergunta aberta ao usuário, referente ao uso da interface na etapa anterior: "A partir das interações feitas com as visualizações, descreva a história contada por estes dados? "

6. Execução, por parte do usuário, da segunda parte do teste. Nesta etapa, era solicitado ao usuário que ele realizasse, no sistema, cinco tarefas, da forma que achasse pertinente, que são:
- T1: Usuário com maior número de *tweets* no estado do RJ.
 - T2: Cidade que mais houve mais *tweets* entre o período de 06/06/2016 e 08/06/2016.
 - T3: Cidade que o evento estava ocorrendo na data de 08/08/2016.
 - T4: Volume de *tweets* da data 07/08/2016.
 - T5: Número de *tweets* do estado de RS.
7. Questionário pós-teste com perguntas referente à experiência de uso com o sistema, que são:
- (a) Como você considera a experiência com a interface? Cite os pontos positivos e os pontos negativos que achares pertinentes.
 - (b) Você se sentiu motivado em interagir nas visualizações?
 - (c) Quais foram as formas de interação que mais auxiliaram a encontrar as informações desejadas?
 - (d) Quais as formas de interagir (filtros) que você mais preferiu?
 - (e) Você acredita que as visualizações apresentadas auxiliaram a encontrar as informações desejadas?
 - (f) Porque você preferiu essa ao invés da outra forma pra encontrar as informações (dependendo do que for observado).
 - (g) Houve algum fato que você observou, através da interface, que lhe chamou mais a atenção?
 - (h) Você tem sugestões de melhoria?
 - (i) Espaço para comentários.

APÊNDICE C – PROTOCOLO DE ANÁLISE FINAL - ETAPA 1

1. Introdução dos objetivos da pesquisa aos participantes, bem como a apresentação e a assinatura do Termo de Consentimento Livre e Esclarecido (Apêndice A).
2. Questões referente ao perfil do participante, sobre o uso de sistemas para visualização de informação, uso de redes sociais e conhecimento sobre o evento do Revezamento da Tocha Olímpica:
 - (a) Gênero
 - (b) Idade
 - (c) Escolaridade
 - (d) Experiência em computação?
 - (e) Utiliza redes sociais?
 - (f) Caso positivo para pergunta anterior, quais e com qual finalidade?
 - (g) Você tem conhecimento à respeito do revezamento da tocha olímpica ?
 - (h) Caso positivo, existem algum fato que você se lembra que ocorreu, relacionado ao revezamento da tocha?
3. Demonstração do sistema proposto, as formas de interação com ele, o que estava sendo mostrado em cada visualização. Era apresentado ao usuário, de maneira geral, as opções disponíveis ao usuário, como interagir com cada visualização, quais eram as formas de filtro e o que estava sendo representado na interface.
4. Execução, por parte do usuário, da primeira etapa do uso da interface proposta. Nesta etapa, os usuários eram livres para interagir na interface, até que achassem pertinente. Os usuários poderiam solicitar ajuda ou tirar dúvidas, a qualquer momento, durante o uso.
5. Execução, por parte do usuário, da segunda parte do teste. Nesta etapa, era solicitado ao usuário que ele buscasse, no sistema, cinco informações, da forma que achasse pertinente, que são:
 - T1: Usuário com maior número de *tweets* no estado do RJ.
 - T2: Cidade que mais houve mais *tweets* entre o período de 06/06/2016 e 08/06/2016.
 - T3: Cidade que o evento estava ocorrendo na data de 08/08/2016.
 - T4: Volume de *tweets* da data 07/08/2016.
 - T5: Número de *tweets* do estado de RS.

6. Questionário pós-teste com perguntas referente à experiência de uso com o sistema, que são:
- (a) Como você considera a experiência com a interface? Cite os pontos positivos e os pontos negativos que achares pertinentes.
 - (b) Você se sentiu motivado em interagir nas visualizações?
 - (c) Quais foram as formas de interação que mais auxiliaram a encontrar as informações desejadas?
 - (d) Quais as formas de interagir (filtros) que você mais preferiu?
 - (e) Você acredita que as visualizações apresentadas auxiliaram a encontrar as informações desejadas?
 - (f) Porque você preferiu essa ao invés da outra forma pra encontrar as informações (dependendo do que eu observar).
 - (g) Houve algum fato que você observou, através da interface, que lhe chamou mais a atenção?
 - (h) Você tem sugestões de melhoria?
 - (i) Espaço para comentários.

APÊNDICE D – PROTOCOLO DE ANÁLISE FINAL - ETAPA 2

1. Introdução dos objetivos da pesquisa aos participantes, bem como a apresentação e a assinatura do Termo de Consentimento Livre e Esclarecido (Apêndice A).
2. Questões referente ao perfil do participante, sobre o uso de sistemas para visualização de informação, uso de redes sociais e conhecimento sobre o evento do Revezamento da Tocha Olímpica:
 - (a) Gênero
 - (b) Idade
 - (c) Escolaridade
 - (d) Experiência em computação?
 - (e) Utiliza redes sociais?
 - (f) Caso positivo para pergunta anterior, quais e com qual finalidade?
 - (g) Você tem conhecimento à respeito do revezamento da tocha olímpica ?
 - (h) Caso positivo, existem algum fato que você se lembra que ocorreu, relacionado ao revezamento da tocha?
3. Demonstração do sistema proposto, as formas de interação com ele, o que estava sendo mostrado em cada visualização. Era apresentado ao usuário, de maneira geral, as opções disponíveis ao usuário, como interagir com cada visualização, quais eram as formas de filtro e o que estava sendo representado na interface.
4. Execução, por parte do usuário, da primeira etapa do uso da interface proposta. Nesta etapa, os usuários eram livres para interagir na interface, até que achassem pertinentes. Os usuários poderiam solicitar ajuda ou tirar dúvidas a qualquer momento, durante o uso.
5. Execução, por parte do usuário, da segunda parte do teste. Nesta etapa, era solicitado ao usuário que ele selecionasse um estado e analisasse as informações apresentadas.
6. Questionário pós-teste com perguntas referente à experiência de uso com o sistema, que são:
 - (a) Como você considera a experiência com a interface? Cite os pontos positivos e os pontos negativos que achares pertinentes.
 - (b) Você se sentiu motivado em interagir nas visualizações?

- (c) Quais foram as formas de interação que mais auxiliaram a encontrar as informações desejadas?
- (d) Quais as formas de interagir (filtros) que você mais preferiu?
- (e) Você acredita que as visualizações apresentadas auxiliaram a encontrar as informações desejadas?
- (f) Porque você preferiu essa ao invés da outra forma pra encontrar as informações (dependendo do que eu observar).
- (g) Houve algum fato que você observou, através da interface, que lhe chamou mais a atenção?
- (h) Você tem sugestões de melhoria?
- (i) Espaço para comentários.

APÊNDICE E – ESTRUTURA DE CODIFICAÇÃO PARA O CÓDIGO INTERFACE - ETAPA 1

Estrutura da Codificação		Fala Originária	Participante	
Interface	Positivas	Amigável	P1	
		Completa	P1	
		Limpa	P1	
	Características	Fácil	Acho que para a quantidade de informação está bem organizado e fácil de mexer Foi bem fácil na hora que eu estava explorando.	P1 P4
		Organizado	Acho que para a quantidade de informação está bem organizado e fácil de mexer	P1
		Curiosa	Achei bem curiosa e interessante a forma que ele coloca as informações	P1
		Interessante	Achei bem curiosa e interessante a forma que ele coloca as informações	P1
		Provê Muitas Possibilidades	Qualquer informação que eu queira, da forma que eu queira ela está na tela	P2
		Orienta para Dados Desconhecidos	Eu não sabia quais usuários mais twitaram, nem sabia quais hashtags mais utilizadas, então isso aqui (nuvem) me auxiliaram a saber por qual buscar.	P5
		Limitada	Aqui (nas visualizações) eu só vou conseguir selecionar o que está aparecendo aqui.	P6
Negativas	Confusa	Mas ao mesmo tempo, isso acabou se tornando um pouco confuso, pois, tinha horas que eu não sabia exatamente para onde ir pois haviam muitos itens na tela	P2	
	Tempo para aprendido	...poder investir um tempo e entender todos os componentes da interface e entender aonde está cada coisa e vai conseguir usar isso muito bem	P2	

Estrutura da Codificação		Fala Originária	Participante
Interface	Pontos Positivos	Mapa Interativo	P3
		Forma de Expressar as Relações	P3
	Foco	Quando tu filtras ali pelo lado esquerdo daí tu filtras e ele já te leva para as cidades certas.	P6
		Isso eu achei legal	P5
		Eu achei bom tu poder clicar no mapa e ele já te redirecionar para o estado para tu analisar	P5
		Também já poder ver toda a rota aqui	P5
		Apresenta as Rotas	P5
		Filtro conforme Interação	P5
		Rotas provê ordenação	P5
		Aqui eu já consigo ver a ordem daí eu já sei então “depois de santa Catarina, ela foi para o RS”, e ali no lateral eu não consigo ter essa visualização	P5

Estrutura da Codificação		Fala Originária	Participante
Interface	Sugestões	Explorar mais as cores para essas informações que ficam mais escondidinhas que tu tens que clicar em cima da cidade, talvez se ela fosse mais coloridinha eu visse ‘ah aqui está a quantidade	P1
		Se eu tivesse um pouco maiores ou por exemplo um gráfico assim (?) ou um gráfico com pizza se tivessem gráficos mais coloridos e com outras cores, que fossem mais chamativos talvez fosse mais legal	P3
		Se bem que a cor chamaria mais atenção, talvez a divisão entre cores assim. Para chamar um pouco mais a atenção	P6
		Talvez se usar mais cores eu consiga focar melhor o que eu quero.	P1
		Ele não fosse se interessar por essa parte das cidades, ou das hashtags, ou dos usuários mais aparentes, nessa forma de gráfico....	P2
		Talvez para um usuário final só as tag clouds bastariam	P3
		Cabeçalho dava para deixar ele fixo, aí tu não precisas subir para filtrar	P3
		Se eu tivesse um pouco maiores ou por exemplo um gráfico assim (?) ou um gráfico com pizza se tivessem gráficos mais coloridos e com outras cores, que fossem mais chamativos talvez fosse mais legal	P3
		Pequeno a parte que aparecem os tw’s, a letra poderia ser maior	P4
		Eu acho que poderia mostrar a cidade antes mesmo de eu clicar (nos gráficos). Eu gostei do gráfico, só acho que poderia já mostrar a informação, antes mesmo de clicar	P5
Interface	Sugestões	O mapa seria legal poder filtrar por cidade	P5
		Mas também não depender só do mapa. Talvez ter um filtro por campo para ir para a cidade.	P5
		Acho que não precisar ficar limpando o filtro sempre. Tipo limpar um filtro específico – eu quero continuar no RS e quero limpar o filtro da data e continuar utilizando ele	P5
		Cliquei aqui, daí quando apareceu o resultado, eu vou ali remover os filtros e vou lá e clico e volta como era antes.	P6
		Ah podia ter uma Word Cloud por cidade também, acho que ficaria legal	P6
		Talvez limitar o zoom do que aparece no mapa	P6
		O que acontece, sobre esses filtros, eu raramente vou saber o nome do usuário ou nome da hashtag... o que dá para fazer é colocar um auto complete nos campos	P6
		Se tivesse alguma forma de eu ver a menos tweetados, ou sei lá, ver todas, daí eu chego aqui e escolher	P6
		A questão da ordem que a tocha passou, foi possível extrair, embora foi possível saber a ordem, não deu para ter noção da data que passou por lá (talvez fosse interesse mostrar a data ao invés da ordem)	P6
		Mostrar Data na Rota	P6

APÊNDICE F – ESTRUTURA DE CODIFICAÇÃO PARA O CÓDIGO INTERATIVIDADE - ETAPA 1

Estrutura da Codificação		Fala Originária	Participante
Expectativas	Navegação pelos gráficos	Àquela hora eu queria ver o gráfico mais adiante	P1
	Acesso a Links Existentes	Eu gostaria de poder acessar o link do tweet	P1
Sugestões	Acesso aos Links Externos	Talvez possibilitar que eu clique no link, daí eu conseguiria saber a história toda	P1
	Filtro visual para seleção de período	Uma coisa que talvez pudesse ter era as barrinhas de tempo para arrastar o período de início e período de fim que daí é um filtro mais visual Se tivesse uma questão mais visual nas datas talvez eu usaria ao invés de usar os campos	P2
	Visualização maior em outra página	Outra coisa que seria legal era tu clicar e ver sabe tipo maior ou em outra página	P3
	Curiosidade	Fiquei curiosa de saber porque tão falando de "Pato Branco" (cidade)	P4
Sentimentos	Desorientação	Mas ao mesmo tempo, isso acabou se tornando um pouco confuso, pois, tinha horas que eu não sabia exatamente para onde ir pois haviam muitos itens na tela	P1
	Motivação	Aquela parte de quantidade de tw, não lembro qual que eu fiquei na dúvida que se era um lugar ou outro Para um evento de olimpíada muita gente se valeu disso para promover manifestação ou para promover o pokeemgon , para várias coisas diferentes. E enxergar isso de uma forma que tu enxergues o todo te motiva sim a interagir e procurar	P2
Interatividade	Tarefa	Mas nas partes das tarefas, porque eram datas específicas o filtro lateral é mais fácil	P1
	Laterais	Se é uma coisa muito específica, muito pontual, talvez é mais fácil ir pelo filtro... já é uma coisa que está na minha cabeça que eu quero ver e é mais fácil e rápido do que vir na cloud procurar e clicar. Então é mais fácil abrir o filtro e escrever "fora teme!"	P2
		Questões Específicas	Na questão da data, é uma coisa bem precisa, bem específica eu nem sei se conseguia achar essa data tão precisa indo por esse mapa aqui
	Filtros	O filtro lá do menu foi mais usado por causa das datas	P3
		Os filtros laterais eu utilizei eles para filtrar entre datas	P5
		Se eu quero um filtro mais completo, aí eu vou ali que tu consegue do jeito que tu quiser	P6
	Completo/Possibilidades	Quando eu estava explorando sozinha e via o que ia me chamando. Para essa situação assim mais livre eu gostei da parte mais visual	P1
	Livre		
	Visualizações		
	Curiosidade	Agora se é uma coisa que eu estou mexendo, estou de curiosa, tanto que eu nem usei esse filtro antes quando eu mexi livremente, eu nem abri. Eu fiquei navegando, na curiosidade. Clica em um, vê o que aparece aqui	P2

APÊNDICE G – ESTRUTURA DE CODIFICAÇÃO PARA O CÓDIGO DADOS - ETAPA 1

Estrutura da Codificação		Fala Originária	Participante
Problemas	Informação Repetida	Porque era muito retweet então não tive vontade de ver Os tweets em si, como ficaram pequenos e por ter muito tweet repetido (retweet) não me chamou a atenção em lê-los	P3
	Agrupar Informações de Retweets	Tu poderias contar a quantidade. Talvez mostrar um só e não todos	P6
Dados	Sugestões	Além disso, talvez colorir o mapa, por número de tweets talvez.	P3
	Mapa de Calor pelo Nº de Tweets	De repente usar aqueles mapas de calor com relação ao número de tweets	P6
	Mostrar Data e Hora dos Tweets	Ter a data e a hora do tweet, porque facilitaria né, saber qual data e hora foi postado.	P6

APÊNDICE H – ESTRUTURA DE CODIFICAÇÃO PARA O CÓDIGO INTERFACE - ETAPA 2

Estrutura da Codificação		Fala Originária	Participante
Interface	Característica da Interface	Simplex	P8
		Fácil	P11
	Amigável	P8	
	Criação de Hipóteses	P8	
		P11	

Estrutura da Codificação		Fala Originária	Participante
Interface	Pontos Positivos	eu fui descobrindo que quando eu clicava num dado num determinado gráfico eu conseguia filtrar os tw em relação a eles e as áreas que teve né, por regiões por estado, onde teve mais incidências né, daqueles tweets. Achei legal.	P7
		Eu gostei desse gráfico que mostra por tamanho aqui o destaque	P9
	Relação entre Local e Evento	Coisas que eu gostei: o mapa relacionando o trajeto da tocha, porque tipo é um mapa relacionado diretamente ao evento e não aos tweets, que é uma coisa que eu não estou acostumado a ver em redes sociais.	P10
	Combinações de Filtros	Tipo você conseguir ver o tweet daquela cidade, com aquela hashtag... é um negócio que eu gostaria de ter em rede social. É uma coisa que para mim é novidade	P10
	Formas de Interação	Positivos foi a quantidade de filtros e o jeito de interagir com a interface	P11
	Quantidades de Filtro	Positivos foi a quantidade de filtros e o jeito de interagir com a interface	P11
	Extrair Informações Úteis	acho que deu uma facilidade de saber o que estava ocorrendo em conseguir informações úteis.	P11

Estrutura da Codificação		Frase Originária	Participante
Interface	Pontos Negativos	Até porque faltou um pouco explicar o que cada gráfico realmente mostra	P7
		Pouca Ajuda do Sistema	P8
	Escala dos Gráficos Adaptáveis	Um ponto negativo foi a disposição das coisas. Eu achei meio... ok eu sei que isso é aqui, mas porque tu me disseste eu não achei muito intuitivo, não sei como fazer pra ficar melhor.	P8
		Disposição dos Conteúdos <i>Scroll</i>	P8

Estrutura da Codificação		Frase Originária	Participante	
Interface	Possibilidade de Uso	Comportamento das Pessoas	P7	
		Saber o que as pessoas estão falando	P8	
	Estrutura da Codificação	Informações Qualitativas	Também o dos usuários. Isso porque, pelo que eu estava procurando sobre o que as pessoas estavam falando seria aqui (em relação a word cloud de hashtags),	P8
		Informações Quantitativas	Mas se fosse o caso de eu saber o quanto elas falaram daí com certeza eu viria aqui embaixo (nos gráficos de barra)	P8

Estrutura da Codificação		Frase Originária	Participante
Interface	Informações mais Visíveis	Eu não sabia que ela tinha passado por outra cidade eu tive que aumentar muito ela para eu ver (zoom, porque o mapa esconde os pontos quando estão muito juntos). Talvez se tivesse isso aqui de outra cor e tipo um pontinho só marcando onde tem uma cidade, aí já dá para ver que eu vou encontrar uma cidade ali que a tocha passou...mesmo eu tendo que ampliar a cidade eu gostaria de saber que tem alguma coisa naquele ponto do mapa	P8
		Eu só acho que podia ser mais colorido ali naquela parte que eu falava dos lugares onde a tocha passou podia ser de outra cor para ter mais destaque	P8
		Eu gostaria talvez de coisas que destacassem mais e talvez com o uso de cores	P9
	Feedback Estado Atual do Sistema	Acho que talvez um pouco mais de informação do que cada gráfico representa de início	P7
		Mostrar uma maneira melhor de como tu alterando uma informação em um determinado gráfico tu consegue visualizar nos outros também.	P7
		Faltou um feedback melhor para dizer que quando eu cliquei ali em cima mudou as informações também mudaram para se relacionar com as informações que eu busquei né.	P7
	Sugestões	Acho que poderia estar mais bem indicado o acesso a este filtro (laterais)	P9
		Faltou uma informação dizendo que isso aqui é relacionado a hora, porque para mim não ficou claro no início, embora faça todo o sentido, mas tipo, quando você está olhando uma interface que é bem descritiva, como é o caso da tua interface, porque ela é bem fácil de visualizar, isso aqui ficou relativamente difícil numa interface que é predominantemente simples	P10
		Tem visualizações que eu extraio números, e tem visualizações que eu extraio informações (textuais) ...não sei como poderia ser, talvez duas páginas	P11
	Escalas únicas	Eu achei embaralhado umas coisas, por exemplo a escala dos gráficos deveriam ser sempre únicas ou alguma opção pra setar ela única...isso me confundiu na hora de filtrar aqui, porque eu queria saber o quão grande esse pico foi em relação aos outros picos. Como o gráfico de adaptou, eu não consegui ver.	P11
		As cores eu gostei, achei razoável, mas talvez poderia ser uma determinada cor para a seção, por exemplo se fosse uma cor para a seção de usuário eu saberia que aquela cor estaria sempre falando de usuário	P11

Estrutura da Codificação		Frase Originária	Participante
Interface	Sentimentos	Frustração	P8
		Aí quando eu fui lá, só tinham as coisas que estavam marcando ela, não tinha o que ela em si tinha postado	

APÊNDICE I – ESTRUTURA DE CODIFICAÇÃO PARA O CÓDIGO INTERATIVIDADE - ETAPA 2

Estrutura da Codificação		Fala originária	Participante	
				Laterais
Interatividade	Filtros	Exploração	Acho que se eu estivesse procurando algo mais específico eu viria para cá (filtros)	P9
		Visualizações	Mas se eu quisesse mais explorando como eu estava fazendo eu viria para cá (para os gráficos) – se eu não sei por onde começar.	P9
			poder ver qual foi a pessoa que escreveu	P8
			Talvez se eu pudesse filtrar pelo mapa e selecionar a cidade seria bom.	P9
	Sugestões	Locais das Postagens	E a questão do mapa de tweets, que tipo eu acharia legal de ver, não somente o mapa da tocha. Eu entendi o propósito do mapa, que é você filtrar os tweets quando a tocha está naquele lugar, mas daí tipo você poder ver um mapa e por exemplo eu clicar aqui em uma hashtag e ele mostrar, sem nenhum outro filtro com aquelas hashtags, tipo, iria imprimir o mapa do globo inteiro e não só do Brasil.tipo olha só tem muita gente em porto alegre e muita gente em Londres, por algum motivo. Sabe, ok, é difícil você tirar conclusão do motivo, mas saber que tem uma correlação é curioso	P10

APÊNDICE J – ESTRUTURA DE CODIFICAÇÃO PARA O CÓDIGO DADOS - ETAPA 2

Estrutura da Codificação		Fala Originatória	Participante
Sugestões	Agrupar Informações de RT's	Disso de poder selecionar por usuário porque tem muito retweet, tipo retweet não importa muito. Ok, foi muito retweetado, mas só vendo assim, eu não consigo ver o quanto isso foi retweetado ou quanto uma coisa foi original	P8
	Dados	<p>Talvez pudesse ter algo que formatasse os rt's para que ficassem todos juntos se não fica essa lista interminável de coisas iguais</p> <p>E eu senti falta de uma lista de tweets mais interagidos, além dessas né.</p>	<p>P9</p> <p>P10</p>
Problemas	Spam	Apesar daqui dos tweets aqui de baixo serem meio perdidos, ter muito spam e informação repetida aqui, alguma coisa dá para pescar	P7
	Informações não sucintas	A questão dos tweets relacionados aqui ou dos tweets em geral que estão relacionados com os filtros. É aquilo que eu falei. É legal de ver, mas eu não navegaria por 840 records, para ver coisas, para procurar coisas que interessam,	p10