

ESCOLA POLITÉCNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO
DOUTORADO EM CIÊNCIA DA COMPUTAÇÃO

CAROLINE QUEIROZ SANTOS

**CIViS: MODELO DE *DESIGN* DE INTERFACE CUSTOMIZÁVEL PARA
APOIAR A CONSTRUÇÃO DE VISUALIZAÇÕES NARRATIVAS
INTERATIVAS DE DADOS EXTRAÍDOS DE MÍDIAS SOCIAIS**

Porto Alegre
2018

PÓS-GRADUAÇÃO - *STRICTO SENSU*



Pontifícia Universidade Católica
do Rio Grande do Sul

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
ESCOLA POLITÉCNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**CIViS: MODELO DE *DESIGN* DE
INTERFACE CUSTOMIZÁVEL PARA
APOIAR A CONSTRUÇÃO DE
VISUALIZAÇÕES NARRATIVAS
INTERATIVAS DE DADOS
EXTRAÍDOS DE MÍDIAS SOCIAIS**

CAROLINE QUEIROZ SANTOS

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

Orientador: Prof. Milene Selbach Silveira

**Porto Alegre
2018**

Ficha Catalográfica

S237c Santos, Caroline Queiroz

CIViS - Modelo de Design de Interface Customizável para Apoiar a Construção de Visualizações Narrativas Interativas de Dados Extraídos de Mídias Sociais / Caroline Queiroz Santos . – 2018.

147 p.

Tese (Doutorado) – Programa de Pós-Graduação em Ciência da Computação, PUCRS.

Orientadora: Profa. Dra. Milene Selbach Silveira.

1. Análise de dados. 2. Visualização narrativa. 3. Engenharia Semiótica. 4. End-User Development. 5. Interação Humano-Computador. 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

Caroline Queiroz Santos

CI VIS: modelo de *design* de interface customizável para apoiar a construção de visualizações narrativas interativas de dados extraídos de mídias sociais

Tese apresentada como requisito parcial para obtenção do grau de Doutor 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 12 de abril de 2018.

BANCA EXAMI NADORA:

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

Profa. Dra. Simone Diniz Junqueira Barbosa (PUC-Rio)

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

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

DEDICATÓRIA

Dedico este trabalho aos meus pais.

AGRADECIMENTOS

Cursar e concluir o doutorado só foi possível devido ao apoio e colaboração de muitas pessoas, em diferentes aspectos. Assim, eu não poderia deixar de agradecê-las e dizer o quanto foram importantes nessa desafiadora trajetória da minha vida. Embora não seja fácil ser justa nos agradecimentos, pois cada ajuda e colaboração direta ou indireta não tem como ser medida, eu espero conseguir expressar aqui a minha gratidão a cada um.

Começo agradecendo à minha orientadora Prof.^a Dra. Milene Selbach Silveira, por ter aceitado me orientar, pelas valiosas discussões, pelo incentivo às produções ao longo desses quatro anos, enfim, por todo o trabalho realizado na orientação. Agradeço aos professores da PUCRS com os quais cursei disciplinas e aos que colaboraram com trabalhos que realizamos durante o doutorado, em especial à Prof.^a Dra. Isabel Harb Manssour.

Agradeço também aos meus colegas do Departamento de Computação (DECOM) da UFVJM, por terem apoiado meu afastamento para qualificação de doutorado, o que me permitiu dedicação exclusiva ao curso durante três anos, em especial à querida Maria Lúcia, e à Capes, pela concessão de bolsa-auxílio, o que me permitiu isenção de mensalidade durante 48 meses.

Minha gratidão eterna e imensa aos meus pais e irmãos, que sempre apoiam todas as minhas empreitadas e sempre torcem pelo meu sucesso e felicidade. O amor, a paciência e o entusiasmo de vocês foram fundamentais para eu chegar até aqui. Eu não seria nada sem vocês!

Não posso deixar de agradecer aos colegas que acompanharam e colaboraram, seja em discussões, escritas, desabafos, terapias coletivas, cafés, comemorações de trabalhos concluídos/aceitos e também nos *happy hours* da vida. Cito aqui as *Golden Girls*: Angelina Ziesemer, Luana Müller e Luciana Espíndola ♡; a amiga querida Letícia Machado (dona do grupo da Letícia); a querida colega de apê e grande amiga/filha Aline Zanin (que aprendeu a abraçar!); a minha *Best Teacher Forever* Carol Paz (pelas aulas-terapias e pela grande amizade); os amigos que ganhei em Porto Alegre, além desses já citados, e que ajudaram a deixar minha vida "mais leve" nessa *city*: Diogo, Glaucio, Eduardo Ghidini e Grazi; e, por último, mas não menos importante, a minha irmã Juliana Queiroz, que foi e é meu porto seguro, minha amiga, minha amada cúmplice de vida!

Agradeço, finalmente, aos meus familiares (incluindo meu cunhado Fernando) e aos amigos que torceram e compartilharam das minhas alegrias e tristezas durante o doutorado. Listar todos aqui seria correr o risco de cometer a injustiça de esquecer um ou outro, mas, não posso deixar de citar e ressaltar a importância da presença e da força constante, quase diária, via *Facetime*, da querida amiga Clarissa Ramalho, que aqui representa todos os demais amigos.

A todos vocês, o meu muito obrigada!

CIViS: MODELO DE *DESIGN* DE INTERFACE CUSTOMIZÁVEL PARA APOIAR A CONSTRUÇÃO DE VISUALIZAÇÕES NARRATIVAS INTERATIVAS DE DADOS EXTRAÍDOS DE MÍDIAS SOCIAIS

RESUMO

A análise de dados de mídias sociais tem sido uma prática crescente, tornando a visualização de dados um aspecto importante no contexto de apresentação dos grandes volumes de dados tipicamente extraídos dessas mídias. Um dos desafios nesta área está na apresentação eficiente desses grandes volumes de dados, provendo técnicas que apoiem os processos de *design*, análise e compreensão destes dados. Avançando neste sentido, esta tese apresenta o CIViS, um modelo de *design* de interface customizável para apoiar a construção de visualizações narrativas interativas de dados extraídos de mídias sociais. Trata-se de um modelo descritivo, fundamentado na teoria da Engenharia Semiótica e sua visão sobre *End-user Development*, assim como em conceitos de visualização narrativas. O modelo considera que o analista de dados de mídias sociais tem interesse em organizar a apresentação de seus dados de forma a contar uma história com eles. Assim, o CIViS propõe um conjunto de componentes que estruturam o espaço de *design*, tendo o propósito de auxiliar *designers* e analistas de dados *co-designers* a refletirem sobre diferentes aspectos que influenciam a construção de visualizações narrativas interativas customizáveis. Realizamos quatro estudos e, a partir da triangulação dos resultados, obtivemos subsídios para a consolidação da ideia do modelo, dando prosseguimento à sua criação. Após a concepção do modelo CIViS, realizamos um estudo com foco no uso do modelo por analistas de dados de mídias sociais. Com a conclusão do estudo de uso do CIViS, obtivemos resultados que sinalizam formas pertinentes de representar aspectos relevantes de narrativas. Também verificamos que o modelo é satisfatoriamente descritivo para auxiliar no *design* de visualizações, sejam elas narrativas ou não. Por fim, apresentamos o potencial valor epistêmico do CIViS, considerando que seu uso pode instigar discussões e reflexões importantes durante o processo de *design* de visualizações narrativas interativas customizáveis, atendendo tanto *designers* quanto *co-designers* em seu caráter epistêmico.

Palavras Chave: Análise de dados, Visualização narrativa, Engenharia Semiótica, *End-User Development*, Interação Humano-Computador.

CIVIS: CUSTOMIZABLE INTERFACE MODEL TO SUPPORT BUILDING DATA VISUALIZATIONS CONTAINING NARRATIVES OF DATA EXTRACTED FROM SOCIAL MEDIA

ABSTRACT

The analysis of data from social media has been a growing practice, making data visualization an important aspect in the context of presenting the large volumes of data typically extracted from these media. One of the challenges in this area is the efficient presentation of these large volumes of data, providing techniques that support the design processes, analysis and understanding of these data. Moving forward in this direction, this thesis presents CIViS, a customizable interface design model to support the construction of interactive narrative visualizations of data extracted from social media. It is a descriptive model, based on the theory of Semiotic Engineering and its view on End-user Development, as well as on narrative visualization concepts. The model considers that the social-media data analyst has interest in organizing his/her data presentation in such a way to tell a story with them. Thus, CIViS proposes a set of components that structure the design space, with the purpose of assisting designers and data analysts as co-designers to reflect on different aspects that influence the construction of customizable interactive narrative visualizations. We carried out four studies and, based on the results triangulation, we obtained the bases for consolidating the idea of the model, continuing with its creation. After conceiving the CIViS model, we conducted a study focusing on the use of the model by social-media data analysts. With the conclusion of the study of use of the CIViS, we obtained results that indicate pertinent ways of representing relevant narrative aspects. We have also verified that the model is satisfactorily descriptive to assist in the design of visualizations, whether narrative or not. Finally, we present the potential epistemic value of the CIViS, considering that its use can instigate important discussions and reflections during the process of designing customizable interactive narrative visualizations, assisting both designers and co-designers in its epistemic character.

Keywords: Data analysis, Narrative Visualization, Semiotic Engineering, End-User Development, Human-Computer Interaction.

LISTA DE FIGURAS

Figura 1.1 – Processo de análise de mídia social [FG14].	15
Figura 1.2 – Etapas da tese ao longo do tempo.	20
Figura 2.1 – Direitos dos homossexuais nos EUA, estado por estado. Fonte: The Guardian - https://goo.gl/K8Z7jP	25
Figura 2.2 – Processo de criação de história em visualização proposto por Lee et al. [LRIC15].	26
Figura 2.3 – Gêneros de visualização narrativa identificados por Segel e Heer [SH10].	27
Figura 2.4 – Adaptado do artigo de Fischer [FGY ⁺ 04] sobre custo de aprendizagem em linguagens de programação.	30
Figura 2.5 – Modelo de Comunicação de Jakobson	32
Figura 2.6 – Processo de Metacomunicação.	33
Figura 2.7 – Processo de Metacomunicação/ Espaço de <i>Design</i> em EUD.	34
Figura 3.1 – Porcentagem de <i>tweets</i> com as três <i>hashtags</i> mais usadas.	39
Figura 3.2 – <i>Tweets</i> com sentimentos sobre os jogadores atacantes do time brasileiro, clas- sificados como positivos, neutros e negativos.	39
Figura 3.3 – Média dos <i>tweets</i> por sentimento com <i>timeline</i> interativa.	40
Figura 3.4 – <i>Tweets</i> por sentimento com <i>timeline</i> interativa, focando no recurso de interação.	40
Figura 3.5 – Estrutura básica comum a todas as visualizações.	46
Figura 3.6 – Gráfico de linhas interativo após o <i>zoom</i> (Gráfico 1).	46
Figura 3.7 – Gráfico multilinhas interativo com a quantidade de <i>tweets</i> positivos e negativos ao longo do tempo (Gráfico 2).	46
Figura 3.8 – Gráfico de círculos interativo (Gráfico 3).	47
Figura 3.9 – Gráfico de linhas interativo integrado com notícias (Gráfico 4).	47
Figura 3.10 – <i>String</i> de busca.	51
Figura 3.11 – Distribuição dos artigos ao longo dos anos.	53
Figura 3.12 – Artigos com mais de 20 citações segundo o Google Scholar Citations (em 11 de fevereiro de 2016).	54
Figura 3.13 – Artigos por categoria de contexto.	57
Figura 3.14 – Técnicas de visualização mais utilizadas.	59
Figura 3.15 – Necessidades dos analistas em relação a ferramentas de Visualização de Da- dos.	71
Figura 4.1 – Representação das mensagens de metacomunicação.	77
Figura 4.2 – Espaço de <i>Design</i> do CIViS.	80
Figura 4.3 – Estrutura do 1º nível do CIViS: comunicação <i>designer - co-designer</i>	81

Figura 4.4 – Exemplo de elementos de narrativa extraídos de um <i>post</i>	82
Figura 4.5 – Relacionamento entre componentes da cena de visualização narrativa no CIViS.	83
Figura 4.6 – Estrutura do 2º nível do CIViS: comunicação <i>co-designer</i> - usuário final.	83
Figura 4.7 – Fluxo de ações/decisões do <i>co-designer</i> sugerido pelo CIViS.	87
Figura 5.1 – Protótipo do espaço de <i>design</i> do 2º nível do CIViS: comunicação <i>co-designer</i> - usuário.	91
Figura 5.2 – Materiais organizados para realização da tarefa pelo analista de dados. (a) Cenário, dados, tarefa, fluxo de passos e tutorial do CIViS; (b) Cena; (c) Tabela de registro de decisões; (d) cartas que representam componentes do CIViS (havia mais de uma instância de cada carta; e (e) <i>post-its</i> e canetas coloridas.	92
Figura 5.3 – Modelo de carta do protótipo do CIViS em frente e verso. A frente contém o nome e ícone do componente ao qual a carta corresponde - neste exemplo a técnica de visualização gráfico de linha - e o verso contém a explicação desse componente, como parte do sistema de ajuda.	92
Figura 5.4 – Foto do momento em que P1 finalizava as suas decisões sobre a visualização narrativa interativa a ser criada.	96
Figura 5.5 – Foto do momento em que P2 finalizava as suas decisões sobre a visualização narrativa interativa a ser criada.	98
Figura 5.6 – Foto do momento em que P3 finalizava as suas decisões sobre a visualização narrativa interativa a ser criada.	99
Figura 5.7 – Foto do momento em que P4 finalizava as suas decisões sobre a visualização narrativa interativa a ser criada.	100
Figura 5.8 – Ideia inicial da proposta do modelo.	107
Figura D.1 – Carta que representa o processo de análise dos dados.	130
Figura D.2 – Cartas que representam os elementos de narrativa: (a) evento, (b) ator, (c) localização e (d) tempo.	131
Figura D.3 – Cartas que representam os modelos híbridos de visualizações narrativas do CIViS: (a) <i>Martini Glass Structure</i> , (b) <i>Interactive Slideshow</i> e (c) <i>Drill-Down Story</i>	132
Figura D.4 – Opções de estruturas de <i>frames</i> de uma cena.	133
Figura D.5 – Cartas que representam os gêneros de visualizações narrativa: (a) <i>magazine style</i> , (b) <i>annotated chart</i> , (c) <i>partitioned poster</i> , (d) <i>flow chart</i> , (e) <i>comic strip</i> , (f) <i>slide show</i> e (g) <i>film/vídeo/animation</i>	134
Figura D.6 – Cartas que representam as técnicas de visualização de dados: (a) gráfico de linha, (b) gráfico de barras, (c) gráfico de pizza, (d) grafo, (e) gráfico de área, (f) mapa de calor, (g) gráfico de coordenadas paralelas, (h) gráfico de bolhas, (i) tabela e (j) mapa.	136

Figura D.7 –Cartas que representam os elementos visuais: (a) estruturas visuais, (b) realce e (c) orientação de transições.	137
Figura D.8 –Cartas que representam os elementos estruturais: (a) ordenação, (b) interatividade e (c) mecanismos de mensagem.	137
Figura D.9 –Carta que representa o sistema de ajuda.	138
Figura E.1 –Fluxo de passos para customização de visualização narrativa interativa.	142
Figura E.2 –Fluxo de passos para customização de visualização narrativa interativa após análise de representação visual das possibilidades de retornar para qualquer etapa anterior.	143

LISTA DE TABELAS

Tabela 2.1 – Modelos híbridos de visualizações narrativas por Segel e Heer [SH10]	28
Tabela 3.1 – Perfil dos participantes do estudo 1 - grupos focais.	40
Tabela 3.2 – Perfil dos participantes do estudo 2 - <i>survey online</i>	45
Tabela 3.3 – Resultado do processo de seleção.	52
Tabela 3.4 – Principais técnicas de visualização narrativa classificadas por categoria.	58
Tabela 3.5 – Perfil dos participantes do estudo 4.	63
Tabela 3.6 – Persona 01	73
Tabela 3.7 – Persona 02	73
Tabela 3.8 – Persona 03	74
Tabela 3.9 – Persona 04	74
Tabela 5.1 – Perfil dos participantes do estudo do uso do CIViS.	93

SUMÁRIO

1	INTRODUÇÃO	15
1.1	OBJETIVOS	18
1.2	CONTRIBUIÇÕES	19
1.3	METODOLOGIA	20
1.4	ORGANIZAÇÃO DA TESE	21
2	FUNDAMENTAÇÃO TEÓRICA	22
2.1	VISUALIZAÇÃO NARRATIVA	22
2.2	<i>END-USER DEVELOPMENT</i> - EUD	29
2.3	ENGENHARIA SEMIÓTICA E EUD	31
2.4	CONSIDERAÇÕES SOBRE OS TRABALHOS ANALISADOS	35
3	ESTUDOS SOBRE ANÁLISE DE DADOS DE MÍDIAS SOCIAIS	37
3.1	ESTUDO 1 - GRUPOS FOCAIS COM JORNALISTAS	37
3.1.1	METODOLOGIA DE PESQUISA	38
3.1.2	RESULTADOS	41
3.1.3	CONCLUSÕES	43
3.2	ESTUDO 2 - SURVEY ONLINE COM JORNALISTAS	43
3.2.1	METODOLOGIA DE PESQUISA	44
3.2.2	RESULTADOS	48
3.2.3	CONCLUSÕES	49
3.3	ESTUDO 3 - MAPEAMENTO SISTEMÁTICO DA LITERATURA SOBRE VISUALIZAÇÃO DE DADOS E <i>STORYTELLING</i>	50
3.3.1	METODOLOGIA	50
3.3.2	RESULTADOS	53
3.3.3	DISCUSSÃO	60
3.3.4	CONCLUSÕES	62
3.4	ESTUDO 4 - ENTREVISTAS COM ANALISTAS DE DADOS DE MÍDIAS SOCIAIS	62
3.4.1	METODOLOGIA	63
3.4.2	RESULTADOS	64
3.4.3	CONCLUSÕES	69
3.5	TRIANGULAÇÃO DOS ESTUDOS E CRIAÇÃO DE PERSONAS	70

4	CIViS - MODELO DE <i>DESIGN</i> DE INTERFACE CUSTOMIZÁVEL PARA APOIAR A CONSTRUÇÃO DE VISUALIZAÇÕES NARRATIVAS DE DADOS EXTRAÍDOS DE MÍDIAS SOCIAIS	76
4.1	ESTRUTURA DO CIViS	79
4.1.1	DESCRIÇÃO DO ESPAÇO DE <i>DESIGN</i>	79
4.2	O PAPEL DO CIViS NO PROCESSO DE DECISÕES DE <i>DESIGN</i> DE VISUALIZAÇÕES NARRATIVAS	84
4.2.1	DECISÕES DO <i>DESIGNER</i>	85
4.2.2	DECISÕES DO <i>CO-DESIGNER</i>	86
5	ESTUDO DO USO DO CIViS	90
5.1	PROVA DE CONCEITO DO CIViS	90
5.2	METODOLOGIA	91
5.3	PERFIL DOS PARTICIPANTES	93
5.4	ANÁLISE DA CONSTRUÇÃO DA NARRATIVA	94
5.5	ANÁLISE DO CIViS	104
5.6	RESPONDENDO ÀS QUESTÕES DE PESQUISA	107
6	CONSIDERAÇÕES FINAIS	109
6.1	CONTRIBUIÇÕES	112
6.2	TRABALHOS FUTUROS	113
	REFERÊNCIAS	114
	APÊNDICE A – Roteiro do Grupo Focal	121
	APÊNDICE B – Lista dos artigos selecionados no mapeamento sistemático	122
	APÊNDICE C – Roteiro de Entrevista com Analistas de Dados	129
	APÊNDICE D – Tutorial do CIViS	130
D.1	ANÁLISE DA BASE DE DADOS	130
D.2	SELEÇÃO DOS ELEMENTOS DE NARRATIVA	131
D.3	SELEÇÃO DO MODELO DE VISUALIZAÇÃO NARRATIVA	131
D.4	DEFINIÇÃO DAS CENAS	132
D.5	SELEÇÃO DOS GÊNEROS	133
D.6	SELEÇÃO DOS TÉCNICAS DE VISUALIZAÇÃO	134
D.7	SELEÇÃO DOS ELEMENTOS VISUAIS	135

D.8	SELEÇÃO DOS ELEMENTOS ESTRUTURAIS	136
D.9	SISTEMA DE AJUDA	137
	APÊNDICE E – Material utilizado no estudo de uso do CIViS	139
E.1	APRESENTAÇÃO DO CIViS AOS PARTICIPANTES	139
E.2	TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO	140
E.3	QUESTIONÁRIO PERFIL	141
E.4	FLUXO DE PASSOS PARA CUSTOMIZAÇÃO DE VISUALIZAÇÃO NARRATIVA INTE- RATIVA	142
E.5	FLUXO DE PASSOS PARA CUSTOMIZAÇÃO DE VISUALIZAÇÃO NARRATIVA INTE- RATIVA REFINADO	143
E.6	DADOS PARA O ESTUDO	144
E.7	EXECUÇÃO	145
E.8	TABELA DE REGISTRO DAS DECISÕES DO <i>CO-DESIGNER</i>	146
E.9	ROTEIRO DE ENTREVISTA SEMIESTRUTURADA SOBRE A TAREFA	147

1. INTRODUÇÃO

A análise de dados de mídias sociais é uma prática que vem crescendo e demandando o desenvolvimento de soluções computacionais para apoiar todo o seu processo. O volume de dados que é coletado facilmente por meio de software e *crawlers*¹ é muito grande, o que gera a necessidade de soluções computacionais também para o processo de tratamento desses dados e para a sua apresentação. A coleta e o tratamento de dados são as etapas definidoras do escopo de análise, permitindo análises diversas, tais como identificação de padrões de comportamento, tópicos de interesse dos usuários, entre outros. São, portanto, etapas fundamentais e de grande impacto no processo de análise. Não menos importante, contudo, é a apresentação dos resultados dessa análise de forma que possam ser facilmente lidos e suscitar novas análises e releituras. Assim, no processo de análise, faz-se necessário definir quais dados se deseja coletar, como será feita a coleta, como esses dados serão processados e, também, como serão apresentados. Esses passos, ou etapas, (figura 1.1) são a base para a obtenção de dados, informações e conhecimentos que auxiliam a análise e a tomada de decisão.

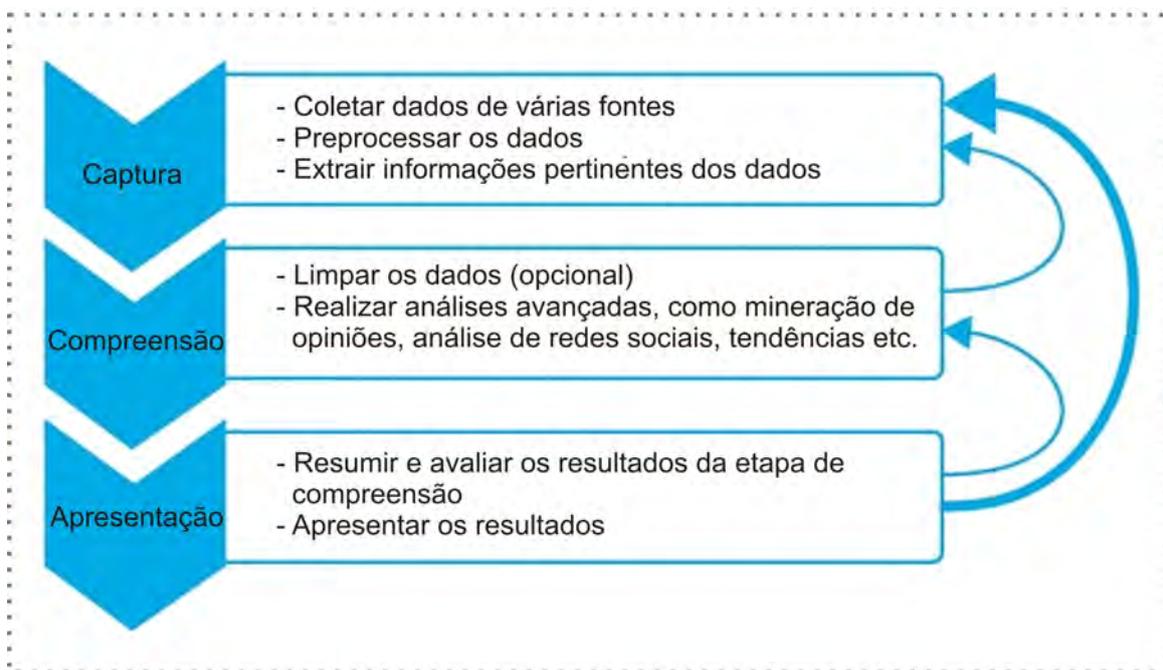


Figura 1.1 – Processo de análise de mídia social [FG14].

Para Russell [Rus11], a variedade das formas de compartilhamento de conteúdo na Internet abriu uma oportunidade de entrelaçar elementos da web (como usuários, conteúdo compartilhado, relações) com técnicas computacionais de coleta, processamento e visualização para auxiliar a análise de dados. O trabalho envolvido nisso abrange mineração de dados, processamento de linguagem natural, aprendizado de máquina, entre outras técnicas e processos analíticos

¹Crawlers são aplicações de software que percorrem a Web capturando e armazenando informações de acordo com um conjunto de regras pré-definidas. Também são chamados de robôs de Internet.

complexos. Esse arcabouço computacional naturalmente sugere cooperações interdisciplinares, e o esforço dos pesquisadores, nesse contexto, está em medir “como as coisas acontecem” nas mídias sociais, por exemplo, e buscar inspiração no conhecimento extraído dos dados para propor novos algoritmos para o mesmo contexto ou propor novos sistemas computacionais [Ben15].

Com o advento do *Big Data*, a análise de dados extraídos de mídias sociais, como o Twitter, o Instagram e o Facebook, tem sido o foco de muitas pesquisas em ciência da computação, ciências humanas e ciências sociais, tendo gerado novas áreas de pesquisa, como a *Web Science*, dedicada a estudar pessoas e tecnologias, aplicações, processos e práticas, que são moldados pela Web [BLHHW06, HSH⁺08, Shn07]. A alta capacidade de coleta e armazenamento de dados exige o uso de técnicas efetivas para recuperá-los e apresentá-los. Assim, torna-se pertinente aos profissionais da computação e de outras áreas que lidam mais especificamente com apresentação de dados, o estudo de como (ou de que formas) informações contextualizadas podem ser apresentadas ao usuário, considerando um grande volume de dados. Tendo em vista este nicho, o foco deste trabalho está na apresentação ou visualização de dados, uma área que tem ganhado cada vez mais importância no processo de análise.

Segundo a literatura, o objetivo da visualização de dados é auxiliar no processo de compreensão dos dados, aproveitando o sistema visual humano para identificar padrões, tendências e anomalias [HVW07]. Visualizações bem concebidas podem substituir cálculos cognitivos por simples inferências de percepção e melhorar a compreensão, memória e tomadas de decisão [HBO10]. Ao tornar os dados mais acessíveis e atraentes, as visualizações podem envolver diferentes públicos na sua exploração e análise. Desta forma, podemos considerar a visualização como uma espécie de tradutor de dados para uma linguagem de mais fácil compreensão, utilizando gráficos, tabelas, cores, entre outros recursos visuais. O desafio é criar visualizações eficazes e envolventes, apropriadas para o propósito de cada análise a ser realizada [HBO10] [HVW07].

Embora as ferramentas de visualização de dados tais como Tableau², Gephi³, por exemplo, ofereçam aos seus usuários diversos recursos para a geração de representações gráficas, elas não oferecem alternativas de personalização, como, por exemplo, customização da paleta de cores. Essas ferramentas também não permitem a livre ordenação das visualizações conforme o desejo de seu autor, que, se assim tivesse liberdade, poderia priorizar um ponto de vista, realçar um argumento e apresentar à sua audiência uma nova forma de ver esses mesmos dados, valendo-se de sua habilidade de “contar histórias” por meio de visualizações narrativas.

De fato, a técnica de “contar histórias” com dados tem sido considerada por pesquisadores da área de visualização há algum tempo. A narrativa⁴, como processo de registro de experiência, tem sido um importante meio para ensinar, entreter, convencer e registrar algo. Para Elias e Bezerianos [EAB13], visualização de dados é um ponto importante na construção da análise e na divulgação de descobertas, sendo a narrativa uma poderosa abstração para identificar ameaças e

²<http://www.tableausoftware.com/>

³<https://gephi.org/>

⁴As palavras narrativa e *storytelling* são usadas neste estudo como sinônimos, tendo o significado de contar histórias.

entender padrões. Isso gera a necessidade de desenvolver e melhorar as ferramentas de análise visual com suporte a narrativas [EAB13, SH14] para criação de visualizações narrativas.

Alinhados com essa perspectiva, Fan e Gordon [FG14] observam que análises visuais sofisticadas vão além da simples exibição de informações, fazendo uso de técnicas de visualização capazes de realçar aspectos importantes da análise. Dessa forma, possibilitar a customização de visualizações para diferentes usuários pode ajudá-los a entender grandes quantidades de informações, incluindo padrões que são mais evidentes para as pessoas do que para as máquinas. Analistas de dados e estatísticos podem oferecer suporte adicional a esse processo.

A partir desse contexto e da observação da existência de uma grande variedade de demandas dos usuários em relação a soluções que apoiem a análise de grande volume de dados, tecnologias customizáveis tornam-se cada vez mais necessárias, pois permitem que o usuário crie ou altere um ambiente de acordo com as suas necessidades para obter os resultados esperados. No entanto, usuários leigos em computação ainda não contam com sistemas eficientes que os permitam modificar e/ou personalizar visualizações de dados, existindo, ainda, muitas barreiras na interação dos usuários com esses sistemas [HVW07, EB11]. Alguns pesquisadores, como Elias et al. [EB11], Elias e Bezerianos [EAB13] e Heer et al. [HVW07, HBO10], têm se envolvido com a evolução dessas técnicas, desenvolvendo visualizações interativas e explorando novos recursos para melhoria da interação e envolvimento dos usuários.

Além de estarmos atentos a estas demandas, nossa pesquisa alinha-se com o 12^o dentre os grandes desafios para pesquisadores de IHC [SPC⁺16], que busca acelerar a clareza analítica de grandes volumes e variedade de dados com o apoio de interfaces visuais bem integradas e técnicas estatísticas, dando melhor respaldo a decisões de âmbito individual, comunitário ou global. Também alinha-se com o 2^o desafio (Acessibilidade e Inclusão Digital) dos “Cinco Grandes Desafios de Pesquisa na área de Interação Humano-Computador (IHC) no Brasil: 2012 - 2022” [BdSP12], que diz ser “necessária a construção de sistemas que possam ser (...) especializados para os diferentes usuários com diferentes necessidades e sugere a utilização de estruturas ontológicas (...) para representar tais diferenças e inferir sobre elas”. Assim, inserida neste contexto e alinhada às pesquisas previamente citadas, nossa pesquisa busca também investigar o processo de análise de dados de mídias sociais para desenvolver formas de apoiar analistas em uma melhor compreensão dos dados e uma melhor tomada de decisão.

Antes de prosseguir, é importante notar que, nesta pesquisa, fazemos distinção entre os papéis de **analista de dados** e **cientista de dados**. Embora sejam conceitualmente similares, cientistas de dados precisam ter habilidades em ciência da computação - como escrever códigos para coletar e limpar dados (incluindo conhecimento sobre aprendizado de máquina ou mineração de dados), matemática, estatística e domínio de negócios [PD12], enquanto que analistas de dados não precisam necessariamente possuir qualquer destas habilidades [FD12]. Aqui, focamos as atividades do analista de dados, que entendemos como pessoas que olham para um conjunto de dados e descobrem nele significados, tendo em vista seus negócios ou pesquisa como um todo, sejam eles jornalistas, cientistas sociais, historiadores, comunicadores, ou quaisquer outros. No

contexto deste trabalho, tivemos a chance de realizar estudos com jornalistas, profissionais de mídia e profissionais que trabalham com a análise de dados de mídias sociais.

Considerado o público alvo de nossa pesquisa, cujas necessidades desejamos atender, a questão de pesquisa que norteia o trabalho é definida da seguinte forma: *Como apoiar analistas de dados na customização de visualizações narrativas de dados extraídos de mídias sociais?* A fim de refletir e responder esta questão, realizamos estudos para entender o contexto do problema, tomar ciência do estado da arte da pesquisa que envolve agregar narrativas à visualização de dados, conhecer as práticas e as demandas dos analistas de dados de mídias sociais e também entender o quanto a visualização de dados está presente nas atividades desses profissionais. Além disso, compreender o que poderia ser desenvolvido nesse aspecto para apoiar o processo de análise de dados.

Em seu conjunto, a realização desses estudos levaram à criação de um modelo para o *design* de visualizações narrativas, que visa auxiliar o analista de dados na customização do *design* de suas visualizações. Este modelo foi desenvolvido, aplicado e avaliado com base na teoria da Engenharia Semiótica [Sou05], apresentada no capítulo 2. A expectativa é que o modelo auxilie os usuários analistas de dados na reflexão sobre as decisões a serem tomadas na co-autoria do *design* de visualizações de dados, de forma a criar visualizações que lhes sejam significativas a partir dos “diálogos” estabelecidos com a interface, que os auxiliará na definição de “o que” e “como” se deseja visualizar.

O contexto de aplicação é o *design* da apresentação de dados (coletados e processados) de mídias sociais por analistas de dados. O usuário analista de dados atuará como co-autor do *design* de visualizações, sobre as quais ele poderá refletir e decidir quais dados e quais recursos são mais adequados segundo o que deseja apresentar e analisar. O principal desses recursos a ser aqui considerado são as narrativas que, associadas às visualizações, possibilitarão uma maior exploração dos dados pelo usuário final. Os analistas poderão customizar o *design* da apresentação dos dados utilizando-se de técnicas de visualização já existentes, como gráficos de linha ou de dispersão, *timeline* interativa, grafos, entre outras, além de recursos visuais e estruturais para contar a história dos seus dados.

Na próxima seção, apresentamos o objetivo principal e os objetivos específicos dessa pesquisa. Na seção 1.2, resumimos as contribuições a serem trazidas pela pesquisa, seguida da seção 1.3, em que apresentamos a metodologia de pesquisa adotada. Por fim, na seção 1.4, apresentamos a organização desta tese.

1.1 Objetivos

O objetivo principal desta pesquisa é ajudar analistas de dados de mídias sociais a compreender os dados que possuem e a projetar visualizações narrativas interativas para apresentar esses dados para sua audiência. Para isso, criaremos um modelo descritivo de visualização narrativa in-

terativa customizável, permitindo que o analista de dados atue como co-autor das visualizações projetadas.

Os objetivos específicos são:

- compreender o estado da arte sobre o uso de narrativas associadas a visualizações de dados para entender o que os pesquisadores da área têm produzido;
- compreender como analistas de dados de mídias sociais (leigos em computação) realizam análises de grandes volumes de dados;
- compreender o quanto o uso de ferramentas de visualização de dados está inserido na prática dos analistas de dados de mídias sociais;
- compreender o que os analistas de dados de mídias sociais necessitam em relação a ferramentas de visualização de dados;
- explorar a engenharia semiótica no *design* de visualizações narrativas interativas customizáveis;
- criar uma prova de conceito do modelo, pelo qual os analistas de dados poderão customizar visualizações narrativas interativas de dados de mídias sociais;
- elaborar uma proposta para auxiliar o usuário no aprendizado da execução dessa customização.

1.2 Contribuições

A principal contribuição desta pesquisa é a criação de um modelo descritivo de interface de visualização narrativa interativa em que o usuário pode participar da definição de quais dados serão apresentados e de que forma, considerando o contexto de análise de dados de mídias sociais. Em outras palavras, o *designer* oferece ao usuário analista de dados recursos para customizar sua visualização narrativa interativa, assumindo, então, o papel de *co-designer* da visualização em questão.

Esse modelo é baseado na teoria da Engenharia Semiótica [Sou05] para aplicações que envolvem *End-User Development* (EUD). Assim, uma segunda contribuição da pesquisa é a exploração da Engenharia Semiótica (em relação à sua aplicabilidade) para visualizações customizáveis e interativas contendo narrativas de dados. Outra contribuição, do ponto de vista prático, refere-se à implementação de uma prova de conceito do modelo proposto e, do ponto de vista experimental, sua avaliação e análise.

Enfim, esta tese apresenta contribuições para a área de IHC e para a teoria da Engenharia Semiótica (em relação à sua aplicabilidade neste domínio), na produção de novos conhecimentos e novas questões de pesquisa.

1.3 Metodologia

Para iniciar esta pesquisa, foram realizados estudos preliminares: *i)* levantamento bibliográfico e mapeamento sistemático da literatura sobre visualização de dados e narrativas, *ii)* revisão da literatura sobre EUD; *iii)* grupos focais sobre visualização de dados com jornalistas e alunos de jornalismo; e *iv)* uma *survey* com jornalistas e profissionais de mídias.

Os estudos preliminares possibilitaram uma melhor compreensão do domínio de estudo e do problema em questão. A partir deles, refletimos sobre os seus resultados e realizamos a etapa de requisitos, que consistiu em entrevistas com analistas de dados de mídias sociais e na definição de cenários e de personas para apoiar a construção do modelo. Por fim, criamos o modelo CIViS, realizamos um estudo para prova de conceito e discutimos os resultados obtidos.

As etapas da pesquisa estão ilustradas na figura 1.2, que pode auxiliar em uma melhor compreensão do trabalho realizado.

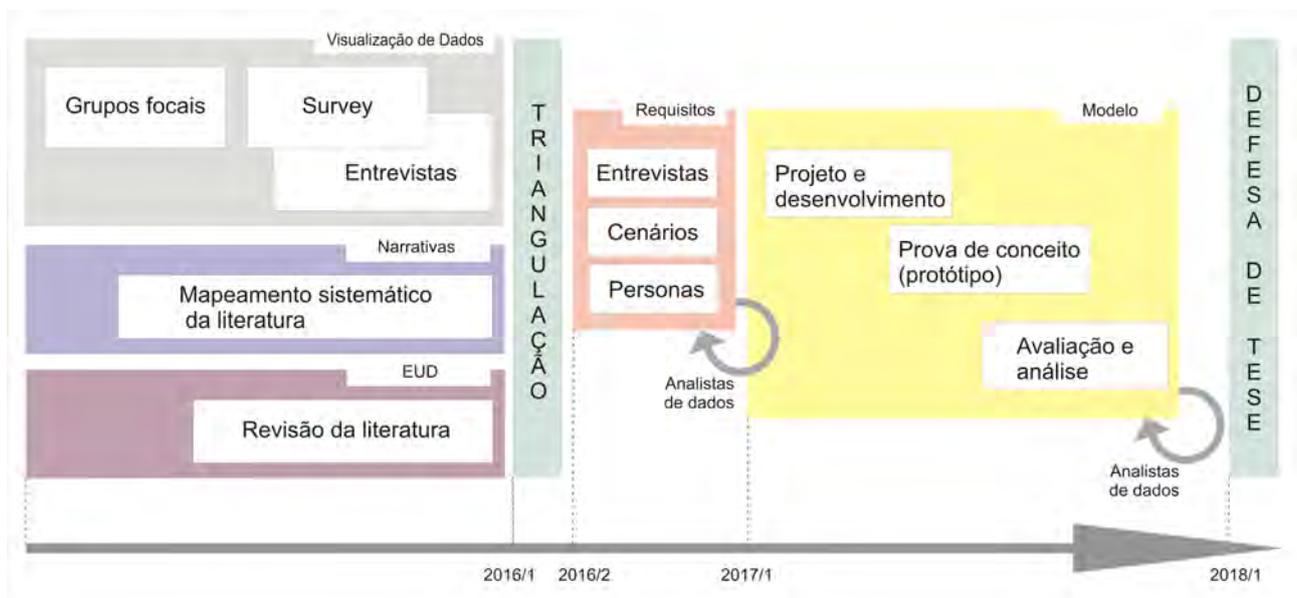


Figura 1.2 – Etapas da tese ao longo do tempo.

Esta pesquisa de doutorado foi realizada totalmente com abordagens qualitativas interpretativas [DL13], com enfoque exploratório, cujo propósito foi observar, analisar, questionar e compreender o que era e como se dava o processo de análise de dados de mídias sociais por analistas. Segundo Creswell [Cre14], a pesquisa qualitativa é uma pesquisa interpretativa, pois o pesquisador geralmente está envolvido em uma experiência sustentada e intensiva com os participantes. A pesquisadora desta tese participou de todas as coletas e análises de dados qualitativos envolvidos nos estudos preliminares e no estudo principal da pesquisa, bem como da triangulação desses dados, conforme será apresentado ao longo desta tese.

1.4 Organização da Tese

Este trabalho está organizado da seguinte forma: no próximo capítulo, apresentamos fundamentação teórica sobre os principais assuntos relacionados à tese (Visualização Narrativa, *End-User Development* (EUD) e Engenharia Semiótica) e, também, os trabalhos relacionados. Em seguida, no capítulo 3, apresentamos três estudos preliminares para melhor compreender o problema em questão. Ainda neste capítulo, apresentamos um quarto estudo, considerado o principal, do qual extraímos o nosso conhecimento sobre as atividades de análise de dados de mídias sociais e as necessidades dos analistas. No capítulo 4 apresentamos o CIViS - modelo de *design* de interface customizável para apoiar a construção de visualizações narrativas de dados extraídos de mídias sociais. O capítulo 5 apresenta um estudo de uso do CIViS por analistas de dados atuando como *co-designers* e as nossas análises sobre os dados coletados neste estudo. Por fim, no capítulo 6, apresentamos as considerações finais, que envolvem as considerações, contribuições, as limitações e os trabalhos futuros em relação a esta pesquisa.

2. FUNDAMENTAÇÃO TEÓRICA

O CIViS foi elaborado com base nos conceitos de Visualização Narrativa, de *End-User Development* (EUD) e da Engenharia Semiótica. Além disso, baseou-se também em estudos realizados para compreender o contexto do problema em questão, bem como o perfil e as necessidades dos analistas de dados de mídias sociais, que serão apresentados no capítulo 3. Neste capítulo, apresentaremos as teorias e trabalhos relacionados à Visualização Narrativa, ao EUD e à Engenharia Semiótica e sua visão sobre EUD.

2.1 Visualização Narrativa

Como abordado na introdução, em linhas gerais, o objetivo da visualização é auxiliar o processo de compreensão dos dados [HVW07], podendo envolver os mais diferentes públicos na sua exploração e análise [HBO10]. A visualização de dados pode ser considerada uma espécie de tradutor de um grande volume de dados para uma linguagem de mais fácil compreensão. Essa linguagem poderia ser uma representação gráfica dos dados ou de conceitos visuais que indicam o objetivo da visualização [War12, WGK15]. Para os autores [War12, WGK15], a visualização tem um impacto positivo no processo de tomada de decisão e também auxilia na análise e na comunicação dos dados. Por isso, ferramentas de visualização têm aumentado as possibilidades de analistas de dados explorarem mais facilmente grandes conjuntos de dados, alcançarem novos *insights* e uma melhor compreensão desses dados [Bro15].

Utilizar visualização para facilitar a comunicação de informações requer, inicialmente, a compreensão dos tipos de dados que podem ser visualizados. Alguns pesquisadores propõem a classificação dos dados sob diferentes aspectos. Geralmente, os dados são classificados como ordinais ou nominais [WGK15], sendo que: *i)* ordinais tomam os dados em valores numéricos - como dados binários, discretos e contínuos; e *ii)* nominais tomam os dados em valores não-numéricos - como dados categóricos, classificados e arbitrários. Ware [War12] identificou os dados como entidades e relacionamentos, sendo definidos como: *i)* entidade é o objeto a ser visualizado, que possui dados como o tamanho, a forma, o valor, a cor, a orientação e a textura; *ii)* relacionamento é a associação entre entidades, como a causalidade, por exemplo. Tanto as entidades como os relacionamentos possuem atributos. O autor classificou esses atributos em dados de categoria, dados inteiros e dados de números reais. Além de Ware, Shneiderman [Shn96] propôs uma taxonomia de tipo de dados de visualização com sete tipos de dados (um, dois, dados tridimensionais, dados temporais e multidimensionais e dados de árvore e redes) e sete tarefas (visão geral, *zoom*, filtro, detalhes *on-demand*, relacionar, história e extratos).

Embora alguns pesquisadores de visualização assumam que as ferramentas que eles usam para análise também sejam adequadas para apresentação dos dados, de acordo com Kosara e Mackinlay [KM13], os objetivos e abordagens da análise são diferentes daqueles da apre-

sentação. Na visualização (ou apresentação), o objetivo principal é esclarecer/convencer sobre um determinado ponto ou explicar algumas descobertas.

Para isso, *designers* de visualizações precisam planejar a melhor forma de representação gráfica dos dados e das informações para facilitar a compreensão dos usuários, bem como possibilitar a manipulação e interação dos usuários com os gráficos para sua melhor exploração dos dados [FCLC01]. Em outras palavras, o *designer* precisa definir quais técnicas de visualização deverão ser utilizadas para representar cada tipo de dados e quais tarefas o usuário poderá realizar naquela visualização.

Existe uma grande variedade de técnicas de visualização, sendo as mais populares e utilizadas as técnicas: tabelas, gráficos de linha, gráficos de barras, gráficos de dispersão, gráficos de pizza, gráficos de área, histogramas, gráficos de bolhas, gráficos de coordenadas paralelas, mapas de calor, *treemaps*, grafos e mapas. A escolha da técnica ou a decisão de combinar mais de uma técnica é fundamental para uma boa exploração de dados e obtenção de conhecimento [FCLC01].

Há pouco mais de 15 anos, pesquisadores [GP01, SH10, Dia10, EAB13, KM13, Fig14a, Fig14b, BDF15] começaram a sugerir que *storytelling* poderia ser usado para estimular o engajamento e a interação do usuário na exploração de dados. *Storytelling* refere-se à expressão “contar uma história” e significa, resumidamente, uma sequência ordenada de fatos que pode conter palavras, imagens, visualizações, vídeos ou qualquer uma das suas combinações [KM13]. O ato de contar uma história é uma das formas mais antigas de comunicação, e a área de pesquisa de visualização de informações vem se beneficiando disso. Gershon e Page [GP01] foram os primeiros pesquisadores a perceber que a narrativa poderia contribuir para a visualização da informação e sua principal contribuição foi conceituar a ideia. Em 2010, o tópico apareceu novamente com Segel e Heer [SH10], que afirmaram que histórias de dados são um tipo emergente de visualizações. Kosara e Mackinlay [KM13] afirmam que a narrativa pode oferecer uma maneira efetiva de apresentar dados.

Há mais de 20 anos, Shneiderman [Shn96] já sinalizava que os mecanismos da história poderiam desempenhar um papel importante no processo de visualização, apoiando a análise iterativa, permitindo que os usuários revejam, recuperem e revisem os estados de visualização. Mesmo que a taxonomia de Shneiderman tenha sido proposta para a área de visualização de forma geral, ela tem semelhança com as propostas da visualização narrativa. De acordo com Figueiras [Fig14b], a combinação de narrativas com visualização de dados tem sido apontada como uma forma eficiente de representar e dar sentido aos dados. Elementos narrativos em visualização de dados são capazes de fornecer explicações sobre o assunto e, muitas vezes, apoiar uma forma estruturada de interpretação que normalmente não existe na visualização de dados tradicional. Visualizações de dados já foram (e ainda são) utilizadas para apoiar formas tradicionais de contar histórias, como informação extra ou como elementos de prova [Dia10]. Hoje em dia, há um grande esforço para transformar as visualizações em uma forma independente de narrativa, que pode existir sem o apoio das formas tradicionais, como vídeo ou texto [Fig14a].

Na área de visualização narrativa, o principal trabalho que motivou nossa pesquisa foi o de Segel e Heer [SH10], em que os autores usaram a analogia de contar histórias para discutir uma série de estudos de caso, de novas mídias, até pesquisas em visualização, comparando os pontos fortes e fracos de diferentes técnicas de visualização. Alguns resultados deste trabalho [SH10] foram uma das bases para a criação do nosso modelo (Capítulo 4).

Ainda sobre narrativas em visualizações, muitos pesquisadores têm pesquisado o uso de técnicas de narrativas na construção de visualizações de dados com o objetivo de verificar se estas técnicas ajudam o usuário a se engajar mais na análise dos dados. Nesse contexto, Aylett et al. [ALW11] apontaram para a existência de conflito entre interatividade e conteúdo da narrativa, problemas, questões e ferramentas relativos a autoria, e direções em avaliação, com o propósito de explorar o equilíbrio entre o controle do autor e do leitor. Eles argumentam que boas visualizações de dados, nessa área, são aquelas que fornecem a seus usuários um breve panorama da narrativa para facilitar a compreensão. Satyanarayan e Heer [SH14] criaram uma ferramenta de visualização de dados para *storytelling* jornalístico e um modelo de abstrações de *storytelling*. Eles entrevistaram jornalistas e estudantes de jornalismo, que possuíam experiência em criação de narrativas de dados, com o objetivo de obter *insights* de como a ferramenta proposta poderia se adaptar aos processos do jornalismo e atender as possíveis necessidades que os jornalistas tivessem. Os participantes conseguiram criar as visualizações narrativas na ferramenta e utilizar os principais recursos disponibilizados.

Visualização narrativa é uma expressão que Segel e Heer [SH10] utilizaram para se referirem a visualizações que incorporam narração de histórias em seu *design*. Eles desenvolveram um *framework* que sugere estratégias de *design* para visualização narrativa, a fim de evitar algumas dificuldades quando o usuário receber liberdade e controle sobre a narrativa. Os autores apontam que as visualizações narrativas diferem das formas tradicionais de contar histórias de maneiras importantes. Segundo eles, enquanto histórias em textos ou filmes geralmente apresentam os eventos em uma progressão bem controlada, contar histórias com visualização de dados pode ser, similarmente, tradicional, em uma sequência linear, porém, pode também ser interativo, convidar os leitores para verificação dos dados, apresentar novas questões e buscar explicações alternativas [SH10].

O uso de visualização narrativa é uma prática muito comum na área de jornalismo e, atualmente, jornalistas têm integrado visualizações às histórias que irão contar. Com isso, a visualização narrativa tem sido muito usada no jornalismo digital para facilitar a narração de notícias e comunicar informações de forma eficiente e intuitiva. Mídias como o *The New York Times*¹, *The Guardian*², *The Washington Post*³ já incorporam visualizações em notícias. Na figura 2.1 podemos ter uma ideia geral de como os meios de comunicação estão usando a visualização em suas matérias.

A partir das nossas leituras sobre o assunto, percebemos que uma visualização narrativa pode ser estruturada a partir de um conjunto de elementos e processos que envolvem: *i*) elemen-

¹<https://www.nytimes.com/>

²<https://www.theguardian.com/international>

³<https://www.washingtonpost.com/>

Gay rights in the US, state by state

Gay rights laws in America have evolved to allow — but in some cases ban — rights for gay, lesbian and transgender people on a range of issues, including marriage, hospital visitation, adoption, housing, employment and school bullying. The handling of gay rights issues vary by state and follow trends by region

Updated 26 June 2015

• Supreme court: gay marriage legal across the US – live updates

Guardian US interactive team

theguardian.com, Tuesday 8 May 2012 16.12 BST

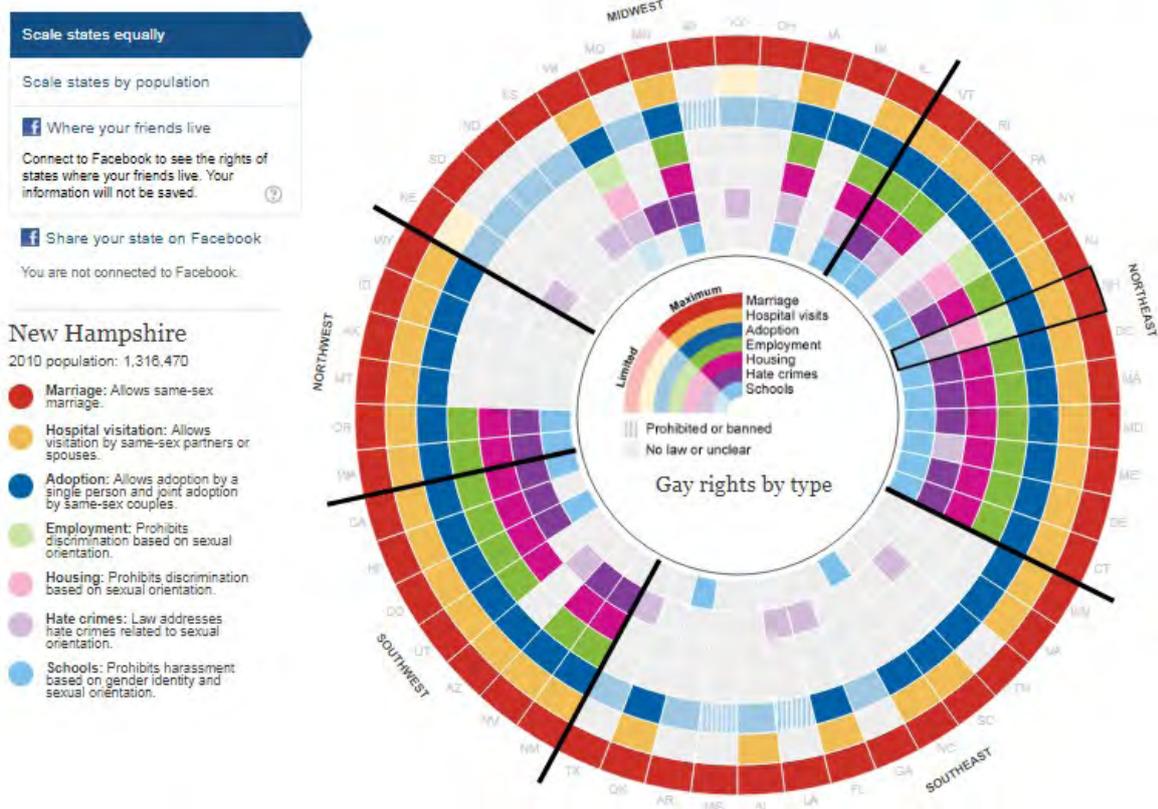


Figura 2.1 – Direitos dos homossexuais nos EUA, estado por estado. Fonte: The Guardian - <https://goo.gl/K8Z7jP>

tos de narrativas (partes que compõem uma história); *ii*) processo de criação de uma história em visualização; e *iii*) o espaço de *design* (elementos presentes na interface).

Elementos de narrativas envolvem Tema, Enredo, Tempo e Narrador. O Tema diz respeito ao que trata a história. O Enredo é o conjunto de acontecimentos ordenados protagonizados pelos personagens. O Tempo está relacionado à evolução do Enredo, e o Narrador é quem estrutura a história [Gan04, Pin06]. Na área de visualização, existem alguns conjuntos de elementos de narrativa utilizados na literatura, como o proposto por Cruz e Machado [CM11], que se basearam em elementos de fábulas inerentes a narrativas de dados. Os elementos são eventos, atores, tempo e localização. Outro caminho para identificar elementos de narrativa é o 5W1H (*who, what, when, where, why, how*) [YHYO01] cuja proposta é de que as respostas a estas perguntas levem à

identificação do contexto da narrativa, como os participantes, o evento, o momento, a localização, o propósito e as circunstâncias em que ocorreram os fatos.

O **processo de criação de uma história em visualização** requer a definição do contexto, seleção das informações, seleção da técnica de visualização e a escolha da ordem de apresentação dos dados da narrativa. Este processo foi descrito no trabalho de Lee et al. [LRIC15], que também serviu de base para a criação do nosso modelo, consistindo na melhoria da definição de o que constitui uma história de dados visual. Neste trabalho, Lee et al. [LRIC15] apresentaram uma descrição detalhada das atividades, artefatos e papéis envolvidos no processo de contar histórias em visualização. O *Visual Data Storytelling Process* foi proposto como um modelo de trabalho que pode conter muitos *loops* e várias ocorrências de cada componente: explorar dados, criar uma história e contar uma história. A progressão desses três componentes não precisa ser linear. Na Figura 2.2, podemos ver esses componentes e os papéis que os contadores de histórias assumem. Todo este processo é o modelo que os autores propuseram para transformar dados em histórias que podem ser visualmente compartilhadas.

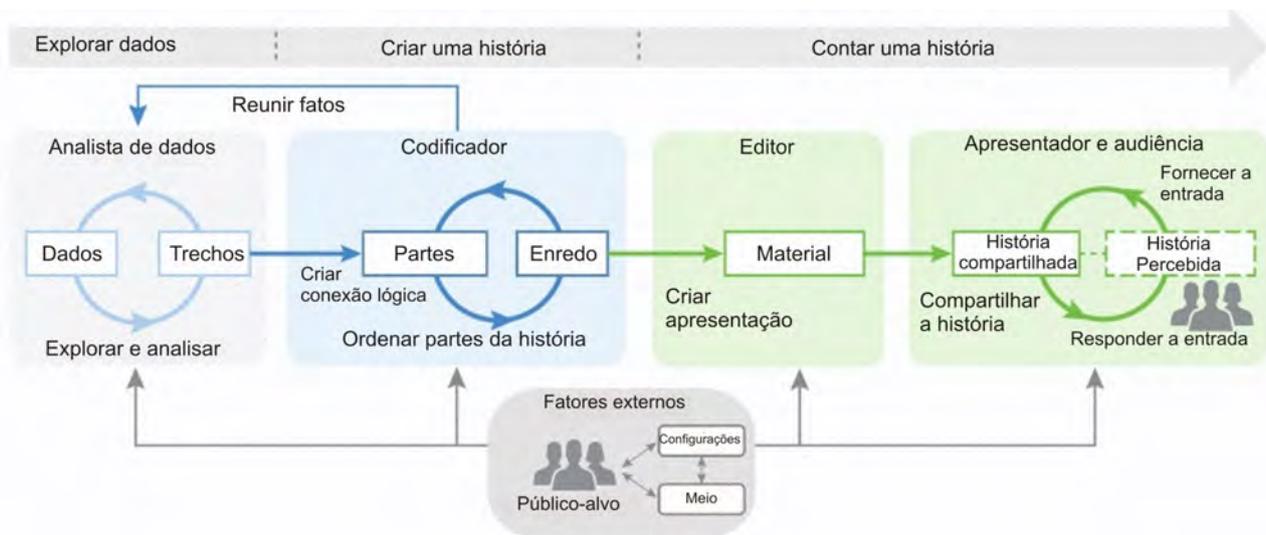


Figura 2.2 – Processo de criação de história em visualização proposto por Lee et al. [LRIC15].

O **espaço de design** evidencia as características que uma visualização narrativa pode conter. Neste trabalho, usaremos como base o espaço de *design* proposto por Segel e Heer [SH10], que consiste em um *framework* com três categorias: gêneros, elementos visuais e elementos estruturais, descritos brevemente a seguir.

Os autores identificaram sete gêneros (formas de apresentar a história) diferentes: *magazine style*, *annotated chart*, *partitioned poster*, *flow chart*, *comic strip*, *slide show* e *film/ vídeo/ animation*, como ilustrado na Figura 2.3. De acordo com os autores, esses gêneros variam principalmente em termos do número de *frames* que cada um contém e da ordenação de seus elementos visuais, além de não serem mutuamente excludentes. Cada gênero descrito por Segel e Heer [SH10] funciona melhor para diferentes tipos de história. Cabe ao *designer* escolher o gênero apropriado para a história que pretende contar, e isso envolve fatores como a complexidade dos dados, a complexidade da história, o público-alvo, e o meio.

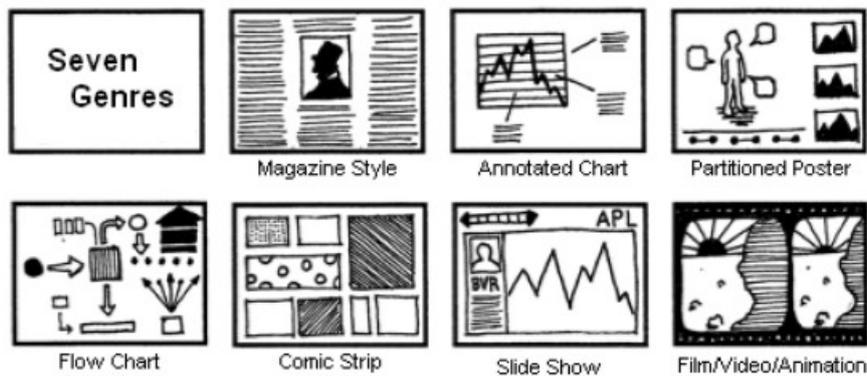


Figura 2.3 – Gêneros de visualização narrativa identificados por Segel e Heer [SH10].

Segel e Heer [SH10] informaram que sua abordagem de classificação de gêneros não considera a experiência cognitiva e emocional dos leitores. Assim, um outro estudo desenvolvido por Figueiras [Fig13], apresentou uma tipologia para a visualização de dados online com o objetivo de atender ao maior número possível de casos de visualização e que consiste em dez diferentes tipos de visualização de dados, não excludentes entre si. Esses tipos incluem o Gráfico Sequencial, *Slideshow*, Gráfico/Diagrama, Mapa, *Tag Cloud*, Modelo, Desenho, Vídeo/Animação, Cartaz e Jogo. Essa tipologia não é tão somente mais exaustiva como também identifica possíveis combinações de elementos que podem melhorar determinados tipos de visualização.

Seguindo com a descrição do espaço de *design* proposto pelos autores [SH10], os elementos visuais são os recursos que auxiliam na visualização narrativa, que podem ser estruturas visuais, realce e/ou orientação de transições. Estruturas visuais são os mecanismos - como *splash screen*, padrões visuais consistentes ou barras de progresso, com o histórico das mudanças - que possibilitam ao leitor compreender a narrativa e identificar em que parte da história ele se encontra.

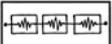
Já os elementos estruturais auxiliam e facilitam o uso da narrativa em visualizações por meio de ordenação, interatividade e mecanismos de mensagem. A ordenação, como o próprio nome diz, tem relação com a forma como as cenas são ordenadas, podendo ser linear, randômica ou direcionada pelo usuário-leitor. A interatividade é o conjunto de possibilidades de interação - como filtro, seleção e navegação - para manipular o conteúdo da visualização. E os mecanismos de mensagem envolvem a comunicação de observações e comentários, geralmente de forma textual, que podem ser anotações, resumos, textos introdutórios, títulos [SH10].

De acordo com os autores, os gêneros, a interação e as mensagens devem se equilibrar entre a narrativa pretendida pelo autor e a descoberta da história pelo leitor. Com isso, eles criaram um espectro de abordagens orientadas por autores e orientadas por leitor, em que as abordagens guiadas pelo leitor tendem a ser mais interativas e menos ordenadas, e as guiadas pelo autor tendem a ser pouco interativas e com muito uso de mensagens [SH10].

Como, geralmente, as visualizações narrativas possuem modelos que combinam as duas abordagens (guiada pelo autor e guiada pelo leitor), Segel e Heer [SH10] apresentaram três mo-

delos contendo as combinações mais usadas: *Martini Glass Structure*, *Interactive Slideshow* e *Drill-Down Story*, descritas na Tabela 2.1.

Tabela 2.1 – Modelos híbridos de visualizações narrativas por Segel e Heer [SH10]

Modelo	Descrição
	<i>Martini Glass</i> : inicia com a abordagem guiada pelo autor com o uso de perguntas, observações ou textos para introduzir a visualização. Assim que a visualização narrativa guiada pelo autor estiver completa, ela se abre para um estágio guiado pelo leitor, em que o usuário fica livre para explorar os dados de forma interativa. Na estrutura em forma de copo de Martini, a haste representa a narrativa guiada pelo autor, e a taça representa os caminhos disponíveis e possíveis de interação guiada pelo leitor.
	<i>Interactive Slideshow</i> : segue um formato típico de apresentação de slides, mas incorpora interação do leitor com a narrativa, dentro dos limites de cada slide. Esta estrutura permite ao usuário explorar mais detalhadamente pontos específicos da visualização antes de avançar para a próxima etapa da história. No entanto, os slides, individualmente observados, geralmente funcionam como o <i>Martini Glass</i> , com a narrativa guiada pelo autor antes de permitir ao usuário interagir.
	<i>Drill-Down Story</i> : apresenta um tema geral e, em seguida, permite ao usuário escolher entre instâncias particulares desse tema para revelar mais detalhes. Esta estrutura possui mais ênfase na abordagem guiada pelo leitor, possibilitando ao usuário escolher quais histórias são contadas e quando. No entanto, ela ainda requer uma quantidade significativa de autoria do <i>designer</i> para determinar os possíveis tipos de interação do usuário, quais histórias serão incluídas, e os detalhes a serem incluídos para cada história.

A partir da análise de trabalhos sobre visualização de dados com narrativas e entendimento dos elementos que os compunham, foram elencados sete componentes considerados fundamentais para a construção de visualizações narrativas: os dados, os elementos da narrativa, as informações, os gêneros, os elementos visuais, os elementos estruturais e as visualizações. Esses elementos foram uma proposta feita por Ghidini [Ghi17] com a justificativa de que são componentes com características próprias e que possuem relação entre si. O autor construiu um modelo de representação de narrativas através de visualizações interativas, com dados extraídos de redes sociais. Esse trabalho vem ao encontro da nossa pesquisa contribuindo para a proposta de um *co-designer* construir suas visualizações, escolhendo seus elementos e os seus significados. Além desses componentes, na nossa pesquisa elencamos outros, que julgamos necessários: a análise da base de dados, os modelos de visualização narrativa, a cena e o sistema de ajuda. No capítulo 4 retomaremos essa discussão e apresentaremos as adaptações feitas para esta tese.

Nesse sentido, para construção de um modelo que oriente o analista de dados *co-designer* na criação de visualizações narrativas, é necessário entender os conceitos que envolvem todo esse processo de *design*. Com isso, será possível traçar os caminhos percorridos na construção da narrativa e tentar representá-los na criação de visualizações, considerando que o próprio usuário (analista de dados) irá definir a forma de apresentação dos dados, em tempo de interação. Na próxima seção, apresentaremos os conceitos que norteiam o desenvolvimento pelo usuário final.

2.2 *End-User Development - EUD*

As necessidades e contextos dos usuários mudam com o tempo, com a demanda de adaptações ou mesmo novas funcionalidade ou comportamentos diferentes do que foi especificado inicialmente para o sistema [Fis07]. Identificar as necessidades do usuário é uma das dificuldades encontradas pelos *designers*. No entanto, metodologias como EUD amenizam essa questão permitindo que os usuários criem seus próprios programas (não, necessariamente, tendo que codificá-los) [LPKW06, WH07]. Ou seja, o usuário pode adequar e adaptar o sistema de acordo com as necessidades que surgirem com o tempo.

Muitas aplicações oferecem mecanismos de customização ou programação, tornando o usuário co-responsável pela eficiência da aplicação na resolução de seus problemas [LPKW06, ES08]. Essas aplicações são reflexo de uma tendência a querer aumentar as habilidades dos usuários dentro do paradigma de EUD, permitindo que o usuário personalize ou customize a aplicação [Bar99]. Essas aplicações são chamadas de aplicações extensíveis.

Segundo Lieberman [LPKW06], “EUD é um conjunto de métodos, técnicas e ferramentas que permitem aos usuários de software agirem como desenvolvedores de software, criando, modificando ou estendendo um artefato de software em algum momento, quando necessitar de algo que não está disponível”. Pode parecer difícil que um usuário leigo consiga aprender a programar durante a realização de uma tarefa, mas estudos indicam que não são necessários pré-requisitos extraordinários para se escrever ou descrever procedimentos semelhantes a programas [Bar99, LPKW06, ES08]. Com isso, é possível que usuários leigos em computação ou com pouco conhecimento na área possam criar, modificar ou estender uma aplicação. Mørch [Mør97] apresenta três diferentes formas para adaptação e alteração de uma aplicação, a saber:

- Customização: o usuário consegue modificar a aparência dos objetos ou os valores de seus atributos a partir de um conjunto de configurações pré-definidas.
- Integração: o usuário consegue ir além da customização e acrescentar funcionalidades ou conectar componentes sem acessar o código diretamente.
- Extensão: o usuário consegue acrescentar novas funcionalidades, inclusive adicionar código, em pontos definidos pelo *designer*.

Fischer et al. [FGY⁺04] mapearam o *trade-off* de linguagens EUD em relação ao escopo de sua aplicação e os custos de aprendizagem, variando os valores entre “alto” e “baixo”, como mostrado no diagrama apresentado na Figura 2.4. O mais alto custo (de escopo e aprendizado) envolve linguagens de programação como Java, C++, Python, que geralmente são usadas por usuários finais que estão muito motivados, principalmente em pesquisas científicas. No contraponto, o mais baixo custo envolve linguagens de EUD específicas de domínio, que atendem uma área de aplicação específica. Neste contexto se encontram a customização de pacotes de software, fazendo com que a programação seja reduzida à entrada de parâmetros em um formulário [FGY⁺04].

Segundo os autores, o contexto ideal de EUD é o que envolve um escopo alto e um baixo custo de aprendizado.

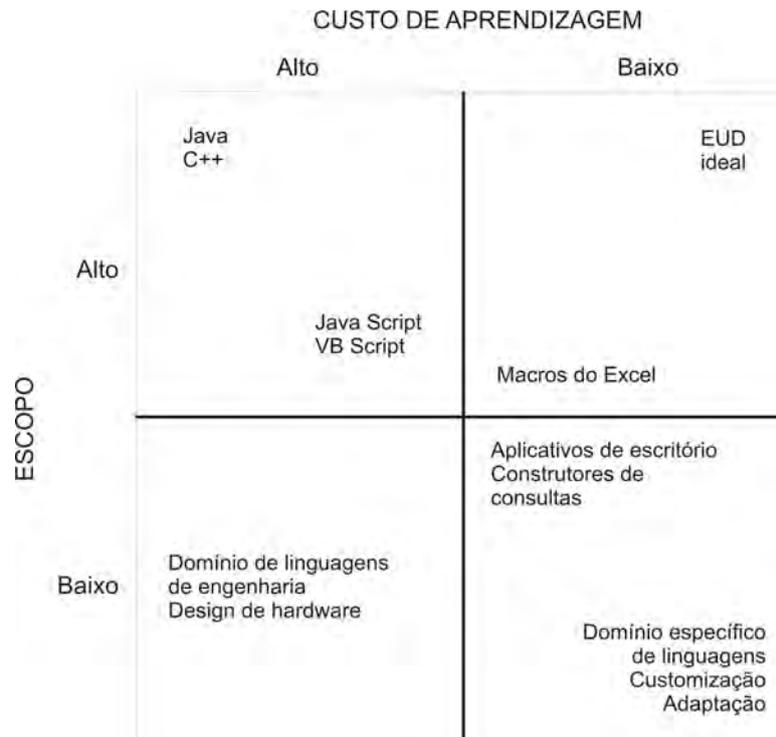


Figura 2.4 – Adaptado do artigo de Fischer [FGY⁺04] sobre custo de aprendizagem em linguagens de programação.

De acordo com a proposta de meta-*design* de Fischer [Fis07], um conjunto de objetivos, técnicas e processos para criação de novas mídias e contextos permite ao usuário atuar como *designer* ou co-autor do sistema. Essa é a solução EUD apresentada pelo autor. Ao estender uma aplicação, o usuário passa a atuar como *designer*, pois assume o papel de projetista da aplicação de acordo com a vontade e necessidades de seus usuários. Já a teoria da Engenharia Semiótica, que iremos discutir na próxima seção, dá ênfase maior às dimensões sociais da programação, definindo que as atividades realizadas pelos usuários finais estão relacionadas a um tipo de comunicação mediada pelo computador, mais social e com uma importância maior em relação à codificação [Sou17]. Segundo de Souza [Sou17], ambos, o meta-design e a Engenharia Semiótica, propõem que o principal desafio para o uso produtivo de sistemas é que a interpretação dos usuários sobre os sistemas, os contextos e as oportunidades de ação estejam em constante evolução. Dessa forma, devemos fazer *design* para uma evolução significativa e constante.

Neste estudo, iremos considerar o processo de *design* baseado na Engenharia Semiótica e sua visão sobre EUD, em que o usuário atua como co-autor da mensagem a ser transmitida pela aplicação aos demais usuários, comunicando a eles, por meio da interface, o que é necessário fazer para melhor utilizar o sistema.

2.3 Engenharia Semiótica e EUD

Semiótica é a ciência dos signos e dos processos significativos que têm por objetivo a investigação de qualquer fenômeno de produção de significado e sentido [San83]. Seu campo de atuação é extenso, pois qualquer signo produzido ou interpretado por seres vivos é matéria semiótica [San83]. Signo é tudo aquilo que representa algo para alguém [Pei77]. Hoje, com o surgimento de novas estruturas e novos meios de disseminação da linguagem humana, torna-se necessário “ler os signos com a mesma naturalidade com que respiramos, com a mesma prontidão que reagimos ao perigo e com a mesma profundidade que meditamos” [San00]. A pessoa, ao perceber um signo e interpretá-lo, gera um significado sobre ele. Esse significado pode dar origem a outras ideias, iniciando o processo de interpretação (semiose).

A Engenharia Semiótica (EngSem) [Sou05, Sou13] traz a perspectiva da semiose na comunicação durante o processo de interação humano-computador, em que o *designer* se comunica com os usuários por meio da interface. De acordo com De Souza [Sou05], na perspectiva dessa teoria, o *designer* participa do processo de comunicação que acontece entre o usuário e o sistema no momento da interação. Mais especificamente, a EngSem trata a interação humano-computador como um processo de comunicação humana mediada por computador, em que os signos são representados na interface por botões, mensagens, ajuda, a serem interpretados de acordo com o que foi projetado. Dessa forma, o processo de significação consiste na codificação e decodificação dos signos utilizados pelo *designer* para comunicar ao usuário o quê, como, onde, por que e para que ele pode/deve se comunicar com o sistema desenvolvido para ele. Assim, neste processo, a interface funciona como uma espécie de “proxy” ou preposto do *designer*, consistindo em um sistema de significação especificado e codificado por ele [Sou13].

Na EngSem, a comunicação entre usuários e sistemas é parte de um processo chamado de metacomunicação (comunicação sobre uma comunicação), que é iniciado pelo *designer* sobre *como, quando, onde e porquê* o usuário pode se comunicar com o sistema que ele projetou [Sou13]. Nesta perspectiva, a interface é vista como uma mensagem enviada do *designer* para o usuário. Esta visão da EngSem baseia-se na perspectiva dos autores Jakobson [Jak60] e Peirce [Pei92, Pei98] e é formada por 3 principais conceitos: ontologia, epistemologia e metodologia. Nesta pesquisa, nos atentaremos apenas ao conceito da ontologia.

A ontologia da EngSem é composta de 4 categorias: processos de comunicação, processos de significação, interlocutores e espaço de design. No processo de comunicação estão a intenção, o conteúdo e a expressão da mensagem trocada entre os interlocutores. O processo de significação está relacionado aos conceitos de signo e semiose tratados anteriormente e, no contexto de IHC, este processo de significação entre o *designer* e o usuário é único. Os interlocutores são o *designer*, a interface do sistema e o usuário. Finalmente, o espaço de design é baseado no modelo de comunicação de Jakobson [Jak60] que é composto pelos elementos: emissor, receptor, mensagem, código, canal e contexto, conforme representado na figura 2.5.

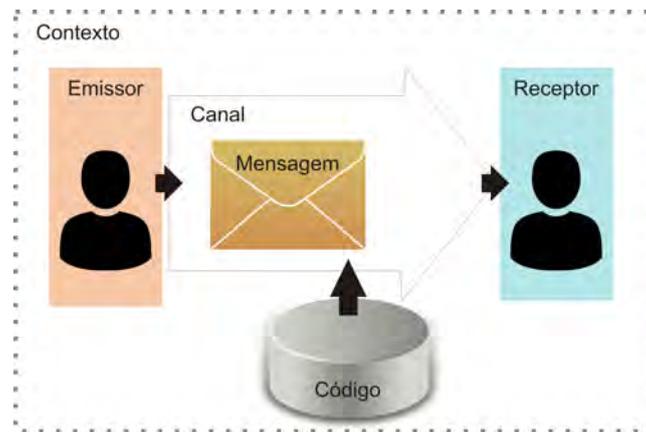


Figura 2.5 – Modelo de Comunicação de Jakobson

O espaço de design da EngSem pode ser mapeado para os elementos do modelo de Jakobson da seguinte forma: o emissor é o *designer*, o receptor é o usuário, o canal é o computador, a mensagem é a interface do sistema, o código são as mensagens codificadas computacionalmente, e o contexto é onde ocorre a comunicação [Sou05]. Na EngSem, é responsabilidade do *designer* ter consciência de cada um desses elementos e saber tomar as decisões sobre cada um dos aspectos do espaço de design. Ele precisa compreender bem quem é o receptor da mensagem que irá enviar, bem como qual é o código e o contexto dessa mensagem, para uma melhor tomada de decisões de *design* [Sou05].

Segundo De Souza [Sou05], na EngSem a interface é considerada um artefato de metacomunicação, ou seja, a interface é composta por mensagens enviadas do *designer* para o usuário e cada mensagem, por sua vez, pode enviar e receber mensagens do usuário, conforme representado na figura 2.6. Nesse processo, a interface cumpre dois papéis: i) o de comunicar a funcionalidade da aplicação (o que a interface representa, que tipos de problemas está preparada para resolver) e o modelo de interação (como se pode resolver um problema), e ii) o de possibilitar a troca de mensagens entre o usuário e a aplicação. A mensagem enviada pelo *designer* para o usuário por meio da interface (chamada de *metamensagem* ou *mensagem de metacomunicação*) pode ser interpretada da seguinte forma [Sou05]:

“Esta é a minha interpretação sobre quem você é e o que eu entendi que você quer ou precisa fazer, de que formas prefere fazê-lo e por quê. Eis, portanto, o sistema que concebi para você, e este é o modo que você pode ou deve utilizá-lo para realizar os objetivos associados com esta (minha) visão”.

Ambos, *designer* e usuário, estão envolvidos na metacomunicação, em que a interface do sistema assume o papel de preposto dos seus *designers*, representando-os no momento da interação. Para isso, o *designer* precisa compreender a intenção do usuário e, em uma alternativa para essa compreensão, possibilitar aos usuários atuarem como *designers*.

A visão da Engenharia Semiótica sobre EUD envolve os pilares considerados na comunicação: o sistema de significação e o de comunicação em si [Sou05]. O ser humano se comunica de

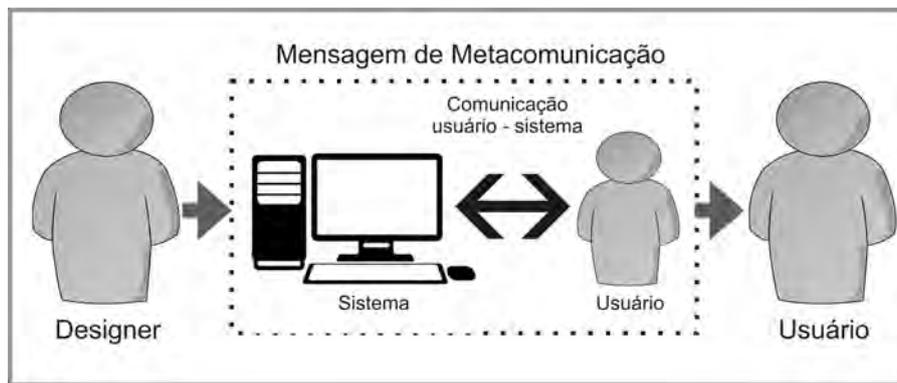


Figura 2.6 – Processo de Metacomunicação.

forma ilimitada e faz uso de várias formas de expressão como humor, ironia, novos termos e novos significados [Alb08] mas, na comunicação humano-computador, os computadores não conseguem, ainda, interpretar a real intenção humana e, por isso, uma boa alternativa pode ser possibilitar que os usuários também atuem como *designers*.

Há mais de 20 anos, iniciaram-se estudos relacionando a EngSem ao EUD [dSBdS01]. Segundo De Souza [Sou17], o que foi proposto naquela época ainda é aplicável hoje, mesmo com grandes mudanças na tecnologia. Antes, o foco das pesquisas eram os tipos mais técnicos de metacomunicação e agora o foco está em investigar aspectos subjetivos da metacomunicação, ou seja, “a presença de crenças e valores individuais, coletivos, afetivos e intencionais no projeto de software e programação de computadores” [Sou17]. As pesquisas com foco nos aspectos mais técnicos propuseram que os sistemas apoiassem atividades de usuários como projetistas, possibilitando a inspeção e elaboração de modificações e extensões através de técnicas de EUD. Com isso, a mensagem de metacomunicação (Figura 2.7) a ser transmitida através do sistema tem a co-autoria do usuário [Sou05], ficando da seguinte forma:

“Esta é a minha interpretação sobre quem você é, e o que eu entendi que você quer ou precisa fazer, de que formas prefere fazê-lo e por quê. Eis, portanto, o sistema que concebi para você, e este é o modo que você pode ou deve utilizá-lo para realizar os objetivos associados com esta (minha) visão. Mas sei que você pode querer modificar a minha visão, objetivando realizar coisas (de alguma forma) que eu não havia pensado. Posso lidar com mudanças que você pode querer fazer, desde que você me fale o que quer nesse código específico” [Sou05].

A evolução dos estudos de EUD na visão da EngSem vem acontecendo desde a publicação da teoria [Sou05] e, mais recentemente, em uma análise sobre o impacto científico da pesquisa em EUD, De Souza [Sou17] aponta que a Engenharia Semiótica é uma alternativa promissora para estimular a pesquisa de EUD baseada em teoria. Alinhados com a EngSem, dois trabalhos inspiraram, inicialmente, a ideia desta pesquisa: o de Albergaria [AMPMJ08], que é um modelo de interface extensível para ampliar o uso de ambientes de mineração de dados na web; e o de Sousa e Barbosa [SB13, dSB14a], que é um sistema de recomendação para apoiar a construção

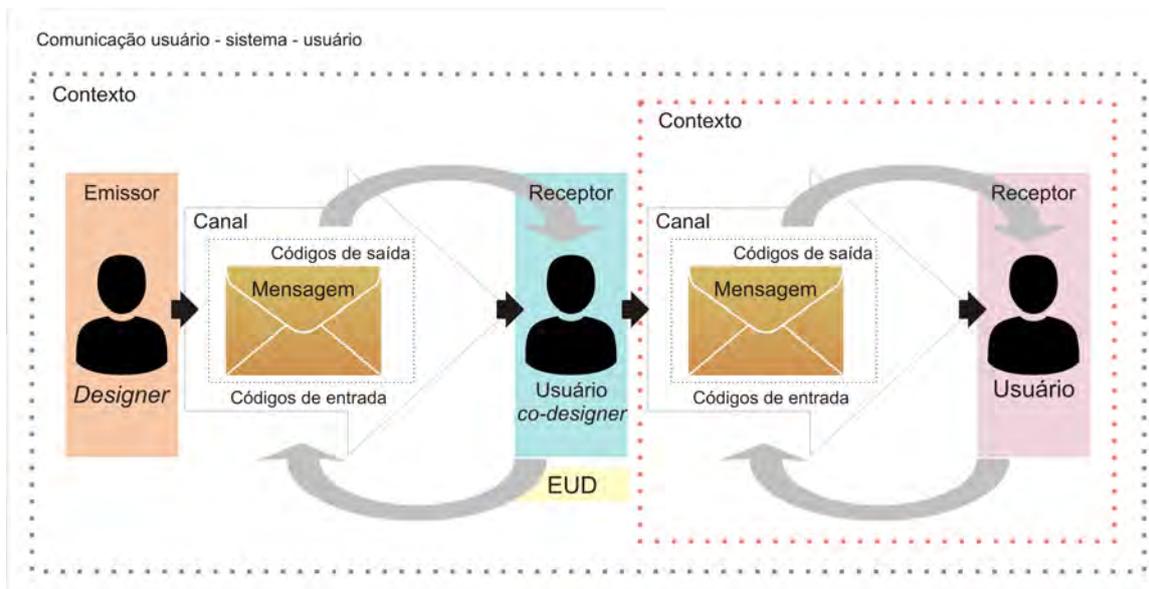


Figura 2.7 – Processo de Metacomunicação/ Espaço de *Design* em EUD.

de gráficos com dados estatísticos por meio de uma série de recomendações baseadas nos dados selecionados e na interação do usuário com a ferramenta. O primeiro, pelo fato de se tratar de um modelo customizável pelo usuário final, e o segundo, por se tratar de uma visualização criada pelo usuário a partir de recomendações do sistema.

O trabalho Albergaria [AMPMJ08, Alb08] envolveu EUD e Engenharia Semiótica. A autora propôs e implementou um modelo de interface extensível para ampliar o uso de ambientes de mineração de dados na web. O modelo considerou dois tipos de usuários: os especialistas e os leigos. Os especialistas, que conhecem o domínio da aplicação e o sistema, criam um nível de abstração para o usuário leigo usar o sistema em contextos e problemas específicos. O modelo foi baseado na teoria da Engenharia Semiótica e permitiu ampliar a aplicação de técnicas de mineração de dados em relação aos contextos de uso e ao público alvo. Neste modelo, diferentemente do que propomos em nossa pesquisa, os usuários leigos são dependentes dos usuários especialistas para criarem abstrações de acordo com as suas necessidades. Nossa proposta é um modelo em que o usuário leigo consiga customizar a interface de visualização de dados sem a necessidade do auxílio de um usuário especialista.

No segundo trabalho que serviu de inspiração inicial para a nossa pesquisa, as autoras Sousa e Barbosa [SB13, dSB14a] desenvolveram o ViSC, um sistema de recomendação para apoiar a construção de gráficos com dados estatísticos através de uma série de recomendações baseadas nos dados selecionados e na interação do usuário com a ferramenta. O sistema oferece um conjunto de gráficos que ajudam a responder questões relacionadas aos dados exibidos no gráfico. Ao percorrer os gráficos recomendados, através das questões relacionadas, o usuário adquire conhecimento tanto do domínio quanto dos recursos de visualização que melhor representam os conceitos do domínio de interesse. As autoras concluíram que as questões eram importantes

para a geração de gráficos eficientes e, por isso, a ferramenta é satisfatória em relação a apoiar os usuários novatos na construção de gráficos.

Em 2015, Monteiro [Mon15] propôs uma extensão da Engenharia Semiótica chamada de EUME – *End-User Semiotic Engineering*, que consiste em uma forma de olhar a Engenharia Semiótica na perspectiva do usuário final que atua como *designer*. Com isso, ele poderá realizar “a construção da metacomunicação através da criação, escolha e disposição dos signos na interface, engajando-se num genuíno processo de Engenharia Semiótica” [Mon15].

Assim como o estudo de Monteiro, nosso estudo não possui foco em investigar os ambientes ou ferramentas que possibilitem EUD, mas investigar os possíveis usuários de ambientes de visualizações customizáveis, neste caso os analistas de dados de mídias sociais, e propor uma solução, baseada na Engenharia Semiótica, para guiar os *designers* e analistas de dados *co-designer* na reflexão e no planejamento das decisões de *design* de suas visualizações narrativas interativas.

2.4 Considerações sobre os trabalhos analisados

Como apresentado no começo deste capítulo, o modelo CIViS foi elaborado com base nos conceitos de Visualização Narrativa, de EUD e da Engenharia Semiótica, além de estudos realizados com foco nas necessidades dos analistas de dados de mídias sociais. Outros trabalhos se relacionam, também, com o nosso propósito, em diferentes aspectos.

Na linha de observar e apoiar analistas de dados em suas análise, Brooks [Bro15] fez um estudo de processos e práticas de análise de dados coletados a partir de mídias sociais. Este estudo foi baseado em entrevistas e etnografia sobre as práticas de analistas de dados de mídia social. O autor discutiu dois projetos de *design* centrado no usuário focados em apoiar o processo de análise de dados: exploração visual de dados do Twitter, e codificação qualitativa colaborativa de mensagens de bate-papo. O trabalho apresenta uma descrição de como os cientistas sociais trabalham com conjuntos de dados obtidos em mídias sociais, as implicações para a concepção de uma melhor técnica computacional de tratamento dos dados (aprendizagem de máquina), a importância da análise visual (visualização de dados) e observações sobre ferramentas livres para a análise online de dados sociais.

Recentemente, um trabalho que se aproximou bastante de um dos propósitos do nosso modelo, que é o de possibilitar a reflexão do usuário analista de dados sobre os dados e como transformá-los em visualizações narrativas interativas, é o de Zhang [Zha17]. O autor estudou e explorou a forma como a visualização é aplicada no jornalismo digital no New York Times (NYT). Como a visualização tem sido muito utilizada no jornalismo para facilitar a construção de narrativas, Zhang investigou como as visualizações são projetadas e incorporadas em notícias. Segundo o autor, há um padrão de escolha de técnica de visualização em relação ao assunto da notícia. Ele mapeou também os princípios de concepção de visualizações em notícias e os componentes mais

utilizados. O objetivo deste trabalho foi servir de referência para profissionais de mídia que tenham interesse em utilizar visualizações narrativas em seus trabalhos.

Não somente estes trabalhos, mas os demais também citados neste capítulo, se relacionam com esta tese em algum aspecto, trazendo subsídios para a criação do modelo de visualização narrativa customizável pelo usuário analista de dados de mídias sociais, que foi especificado a partir das demandas apontadas por eles. Além disso, muitos outros trabalhos [BdP03, BPdS05, Bar06, dSB14b, Vil16] que tratam da criação de ferramentas epistêmicas⁴ para auxiliar *designers* na reflexão sobre os problemas que possuem e as possíveis soluções estão, também, relacionados a esta pesquisa e apoiaram as nossas reflexões sobre a criação do CIViS.

⁴Ferramenta epistêmica pode ser definida como uma ferramenta que apoia o *designer* nas reflexões sobre o problema que possui, sobre como é possível solucionar esse problema e sobre o *design* que está sendo feito, ajudando-o, dessa forma, a obter conhecimento sobre tudo isso [Sou05].

3. ESTUDOS SOBRE ANÁLISE DE DADOS DE MÍDIAS SOCIAIS

Neste capítulo, apresentaremos os estudos realizados nos estágios iniciais da nossa pesquisa, que nos permitiram compreender como as visualizações podem auxiliar profissionais que trabalham com mídias sociais em suas atividades diárias. Foram conduzidos quatro estudos, sendo que o estudo 1 [STT⁺16] e o estudo 2 [SCT⁺17] foram realizados no âmbito de um projeto de pesquisa pertencente ao Programa de Apoio a Integração entre Áreas/PRAIAS, da PUCRS (Editais 07/2014 e 07/2015), com a colaboração de pesquisadores das áreas de Ciência da Computação e Comunicação Social. Estes dois primeiros estudos (empíricos) foram conduzidos por meio de grupos focais e *survey online*, dos quais obtivemos resultados qualitativos que contribuíram em nossas primeiras decisões para a criação do modelo CIViS.

O estudo 3 consistiu em um mapeamento sistemático da literatura sobre visualização de dados e narrativas. Buscamos, com esse mapeamento, obter uma visão do estado da arte da pesquisa nessa área e descobrir quais técnicas, elementos e processos poderiam ser considerados na construção de visualizações narrativas. Ao fim desse estudo, ampliamos nossa compreensão sobre como os pesquisadores da área vêm definindo o espaço de *design* de suas visualizações narrativas.

O estudo 4 (empírico) foi focado mais especificamente em auxiliar no refinamento das decisões na construção do modelo CIViS, consistindo na condução de entrevistas semi-estruturadas com analistas de dados de mídias sociais, com o propósito de compreender melhor suas atividades e demandas, bem como o processo de análise. Os resultados obtidos nos permitiram identificar aspectos importantes relacionados ao uso de visualizações, aos desafios e às dificuldades existentes no processo de análise de dados de mídias sociais, e às demandas existentes em relação a visualização de dados.

O detalhamento desses estudos e seus resultados são apresentados nas seções 3.1, 3.2, 3.3 e 3.4, respectivamente.

3.1 Estudo 1 - Grupos Focais com Jornalistas

O primeiro estudo, como mencionado anteriormente, faz parte de um projeto de pesquisa maior, realizado no âmbito do Programa de Apoio a Integração entre Áreas/PRAIAS, da PUCRS, (Editais 07/2014 e 07/2015). Ele foi realizado com o objetivo de entender como técnicas de visualização podem apoiar estudantes de jornalismo e jornalistas em uma melhor compreensão sobre o sentimento dos usuários nas redes sociais em relação a um determinado assunto [STT⁺16]. Para isso, realizamos um trabalho interdisciplinar, envolvendo pesquisadores da ciência da computação e da comunicação social, em que foram coletados e processados *tweets* durante a Copa do Mundo de 2014, e em seguida foram criadas visualizações para análise de sentimento dos brasileiros *on-*

line no Twitter durante o dia do jogo Brasil vs. Alemanha¹. As visualizações foram apresentadas a jornalistas e estudantes de jornalismo em dois grupos focais heterogêneos (jornalistas e estudantes), em que eles puderam relatar suas percepções e opiniões sobre os gráficos e os dados apresentados. A questão que guiou esse trabalho foi “*como a visualização de dados pode apoiar a obtenção de informações que seriam difíceis de se conseguir e/ou se perceber de outras formas?*”

3.1.1 Metodologia de pesquisa

Esta pesquisa descritiva [LFH10] teve seu foco na construção de uma descrição das opiniões de jornalistas (profissionais ou estudantes) sobre algumas técnicas de visualização e quanto isso pode ajudar em seus trabalhos. Segundo Lazar et al. [LFH10] (pág. 21) pesquisas descritivas, como observações, *surveys* e grupos focais, se concentram na construção de uma descrição exata do que está acontecendo.

Para fazer a pesquisa, dividimos o trabalho em quatro etapas: *i*) coletar dados do Twitter; *ii*) fazer a limpeza e anotação dos dados; *iii*) criar visualizações dos dados; e *iv*) realizar sessões de grupo focal, seguindo de suas análises. Descreveremos aqui as etapas de visualizações e dos grupos focais, que foram o nosso objetivo maior neste estudo.

Na etapa de criar as visualizações, com base nos dados coletados e nos resultados obtidos após a limpeza e anotação dos dados, foram criadas visualizações com foco em *hashtags* e análise de sentimentos. Algumas das *hashtags* mais utilizadas pelos brasileiros online sobre a Copa do Mundo antes do evento foram #vaitercopa, #vaitercopasim e #naovaitercopa. Muitas vezes, essas *hashtags* foram usadas para enfatizar a opinião do usuário sobre os preparativos para o evento esportivo com elogios, críticas de gastos públicos em estádios, atrasos em trabalhos urbanos e de mobilidade, e assim por diante. Por isso, criamos alguns gráficos simples relacionados à visualização de *hashtags* e dos sentimentos dos usuários durante o dia 8 de julho de 2014, quando o jogo da semifinal ocorreu. A figura 3.1 apresenta um gráfico de barras com a porcentagem do uso das três *hashtags* durante esse dia e a figura 3.2 apresenta um gráfico de linhas representando os sentimentos sobre os jogadores atacantes do time brasileiro, classificados como positivos, neutros e negativos.

A figura 3.2 mostra o gráfico dos sentimentos dos brasileiros online no Twitter (classificados como positivos, neutros e negativos) sobre os jogadores atacantes do time brasileiro. O eixo X contém a mudança de placar do jogo, incluindo também a diferença entre o primeiro e o segundo tempo do jogo.

Além dessas visualizações, também desenvolvemos uma visualização interativa da análise de sentimentos obtida com a anotação de polaridade. As figuras 3.3 e 3.4 mostram as duas visualizações interativas da análise de sentimentos sobre o jogo da semifinal, com os *tweets* classificados como positivos, neutros ou negativos. Esses protótipos foram desenvolvidos usando o

¹Foi uma partida de semifinal, realizada em 8 de julho de 2014, que terminou com o placar: Alemanha 7 x 1 Brasil.

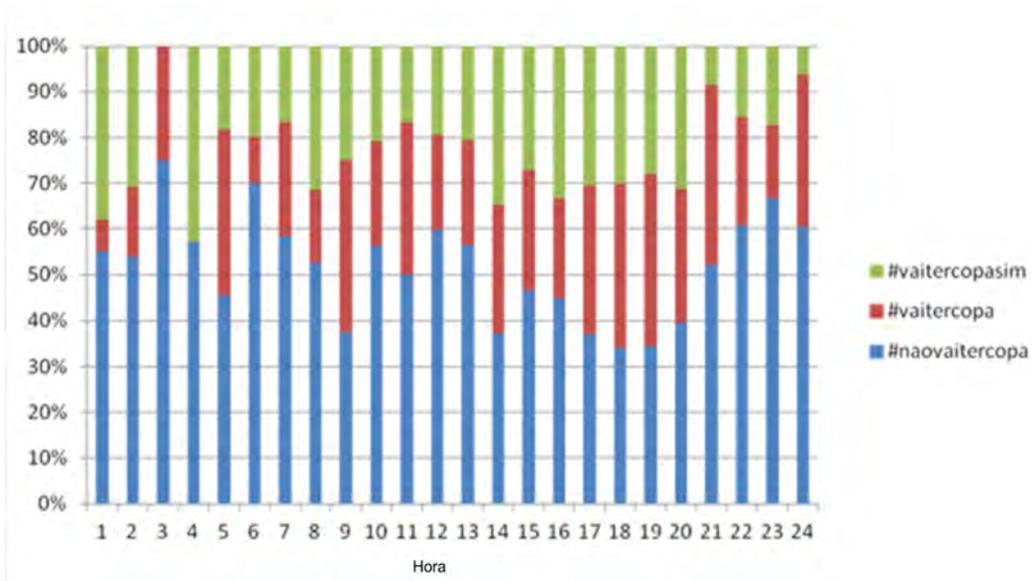


Figura 3.1 – Porcentagem de *tweets* com as três *hashtags* mais usadas.



Figura 3.2 – *Tweets* com sentimentos sobre os jogadores atacantes do time brasileiro, classificados como positivos, neutros e negativos.

D3.js [Mur13]. Na figura 3.3, a visualização corresponde à média dos *tweets* classificados pelos três tipos de sentimentos durante o jogo. O usuário pode interagir na linha de tempo que aparece abaixo do gráfico, ampliando o período de tempo a ser mostrado no gráfico através da seleção de uma parte da linha de tempo. Na figura 3.4, a visualização mostra a quantidade de *tweets* por tipo de sentimento durante o jogo, e a interação também ocorre através da seleção de um período de tempo na linha do tempo para aumentar ou diminuir o *zoom*.

Após a criação das visualizações, partimos para a realização dos grupos focais. As sessões de grupo focal foram conduzidas em dias diferentes em uma sala privada da Universidade, com duração de cerca de 2 horas e seguindo o roteiro planejado (Apêndice A). Na primeira sessão, 4 pessoas (do sexo feminino) participaram do estudo, jornalistas e estudantes de pós-graduação. Na segunda sessão, participaram 5 pessoas (2 do sexo feminino e 3 do sexo masculino), todos es-

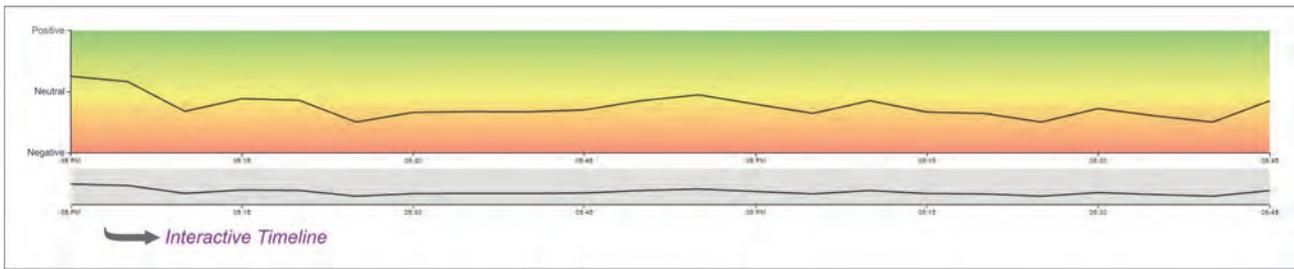


Figura 3.3 – Média dos *tweets* por sentimento com *timeline* interativa.

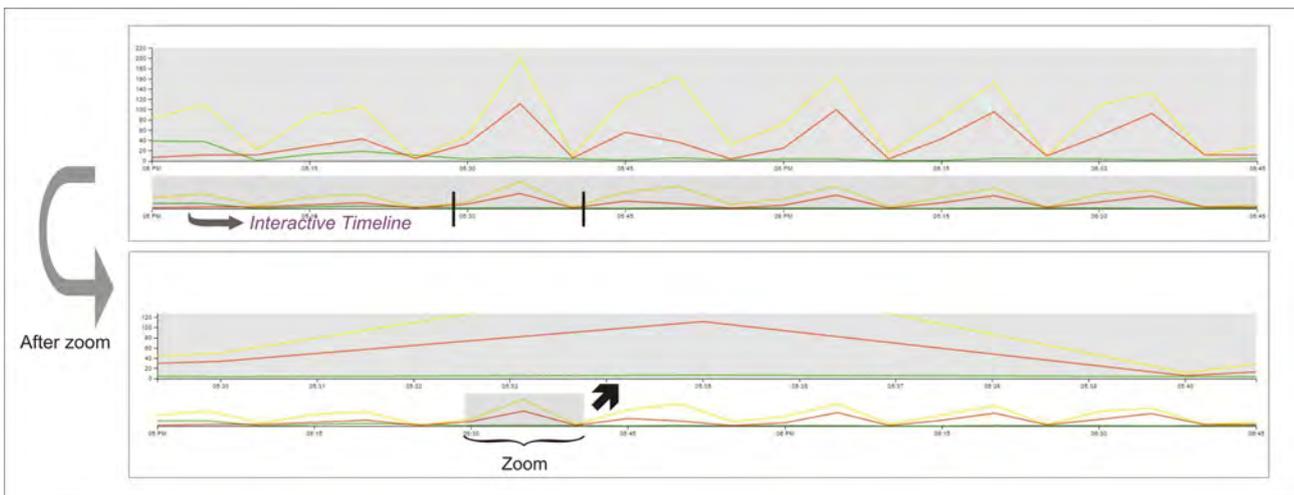


Figura 3.4 – *Tweets* por sentimento com *timeline* interativa, focando no recurso de interação.

tudantes de graduação em Jornalismo com participação ativa em redações de mídia experimental. A tabela 3.1 apresenta o perfil dos participantes.

Tabela 3.1 – Perfil dos participantes do estudo 1 - grupos focais.

Nº	Idade	Sexo	Graduação	Possui conta no Twitter?	Utiliza o Twitter para obter informação?	Seção de Grupo Focal
P1	25	F	Pub. e Propaganda	Sim	Sim	1
P2	26	F	Comunicação Social	Sim	Não	1
P3	28	F	Comunicação Social	Sim	Sim	1
P4	30	F	Comunicação Social	Sim	Sim	1
P5	18	M	Jornalismo	Sim	Não	2
P6	20	M	Jornalismo	Sim	Sim	2
P7	22	M	Jornalismo	Sim	Sim	2
P8	19	F	Jornalismo	Sim	Sim	2
P9	19	F	Jornalismo	Sim	Sim	2

Inicialmente, foi explicado aos participantes os objetivos da pesquisa e da sessão de grupo focal e, em seguida, eles foram convidados a preencher um questionário sobre o seu perfil e a assinar um termo de consentimento livre e esclarecido, caso concordassem em participar. De todos os participantes, apenas 2 não utilizaram as redes sociais para acompanhar a repercussão dos acontecimentos durante a Copa do Mundo de 2014. Os demais utilizaram, principalmente o Twitter e o Facebook. No questionário, perguntamos a opinião dos participantes sobre quais seriam as vantagens do uso de redes sociais para a atividade jornalística. As respostas variaram entre

“possibilidade de maior interação e proximidade com o receptor” e “divulgação de amplo alcance” a “agilidade na obtenção de fontes”, “complementação de informações, levantamento de pautas e busca de *cases*”.

Durante as sessões, após uma breve explicação de cada uma das quatro técnicas de visualização apresentadas, o mediador encorajou os participantes a falar sobre suas ideias e a discutir abertamente suas percepções e experiências, bem como as características pertinentes às visualizações. Para as duas visualizações interativas, o mediador convidou os participantes a irem ao computador e a interagir com elas. As sessões foram gravadas usando duas câmeras de vídeo e um dispositivo de gravação de áudio. As transcrições dessas gravações e anotações sobre tópicos de discussão foram o principal componente da análise deste estudo.

3.1.2 Resultados

Os resultados foram extraídos a partir da análise das discussões das sessões de grupo focal e foram organizadas sob dois aspectos: critérios jornalísticos e técnicas de visualização. Os resultados sugerem que a visualização de dados, em geral, fornece a capacidade de identificar tendências, padrões e anomalias em conjuntos de dados que não seriam visíveis a olho nu. Como tal, elas têm um valor intrínseco para complementar a evidência coletada pelos jornalistas que informam sobre um assunto. Percebemos que os jornalistas estudantes tendem a ser mais críticos com as visualizações, sugerindo muitas vezes mudanças de *layout* de estilo e cor e possíveis usos com outras áreas de comunicação, como publicidade. Atribuímos essa diferença ao ambiente interdisciplinar do trabalho na escola de comunicação, mas também à rapidez e objetividade que o jornalismo profissional exige de seus funcionários, algo que pode ofuscar soluções especulativas facilmente.

Na visualização de *hashtags* (Figura 3.1) os participantes criticaram o esquema de cores, em que a *hashtag* positiva #vaitercopa foi plotada em vermelho, uma cor geralmente associada a algo negativo e não como uma emoção otimista. Eles questionaram as múltiplas leituras possíveis das *hashtags*, sendo que o mesmo texto pode ser usado com uma intenção irônica ou satírica com a adição de outras palavras, algo que a preparação do conjunto de dados e a visualização não contemplaram. Uma sugestão recorrente para esta visualização foi a superposição de uma linha de tempo que indica os momentos-chave do dia do jogo, facilitando tirar conclusões a partir da visualização.

A visualização contendo a análise de sentimentos (Figura 3.2) em relação aos jogadores atacantes da seleção brasileira foi elogiada, pela clareza na percepção das mudanças nas linhas positiva e negativa associada às lembranças do clima generalizado no país no dia da partida. Os participantes manifestaram a curiosidade de continuar seguindo a mudança de sentimento após o jogo, para observar como os usuários brasileiros *online* no Twitter lidavam com aquele placar desprezível. Por outro lado, um participante do primeiro grupo comentou: “*O neutro e positivo estão*

bem dentro do esperado, mas, o que são esses dois grandes picos negativos? Eu não consigo saber apenas olhando para o gráfico". Na interpretação do participante, a visualização mostrou anormalidades que poderiam ser investigadas por diversas estratégias: a capacidade de transmitir uma gravação de vídeo da partida, passar o mouse sobre as linhas e ver as mudanças na frequência da palavra durante a partida e uma análise de conteúdo da os *tweets* considerados negativos, por exemplo, poderiam qualificar uma compreensão aprofundada do sentimento.

Outra crítica foi o espaço constante usado entre os pontos no eixo X, o que sugeriu que os gols aconteceram em intervalos regulares de tempo, o que não aconteceu. Conforme relatório da FIFA sobre a partida, quatro dos sete gols aconteceram dentro de seis minutos, entre o vigésimo terceiro e o vigésimo nono minutos do primeiro tempo². As linhas refletem a surpresa, caindo drasticamente no segundo e terceiro gols alemães, com um pico negativo vindo logo depois. Os pesquisadores consideraram que a crítica é válida, pois uma proporção clara entre a representação do espaço e do tempo poderia informar uma interpretação mais realista dos dados e do evento desse dia.

Em geral, os participantes consideraram as visualizações interativas (figuras 3.3 e 3.4) como mais interessantes entre todas apresentadas. Sobre a visualização com a média dos sentimentos, os participantes discutiram o fato de que sentimentos irônicos durante o jogo ocorreram, na realidade, o que pode ter influenciado os sentimentos classificados como neutros, já que a ironia provavelmente não teria sido considerada um sentimento positivo ou negativo. Um dos participantes afirmou que, se não tivesse sido apresentada primeiro a visualização dos sentimentos em linhas separadas e depois a opção de visualizar a média dos sentimentos, não conseguiria imaginar que a média dos sentimentos fosse o neutro, mas pensaria que o sentimento geral durante o jogo teria sido o negativo. Essa discrepância entre as percepções geradas pelas duas opções de visualização a partir dos mesmos dados aponta para a importância de levar em consideração as habilidades e limitações do sistema visual humano para a identificação de padrões [HBO10] ao criar visualizações de dados. Nesse sentido, a interatividade é uma estratégia produtiva para facilitar a compreensão de grandes volumes de dados, porque permite ao usuário observar e comparar diferentes tipos de padrões e, assim, observar a mesma base de dados sob diferentes perspectivas.

Como sugestões de melhorias nas visualizações apresentadas, os participantes sugeriram maior possibilidade de interação como *"quando você passa o mouse sobre a linha do gráfico, mostra o número exato de tweets"*, *"sobrepor a linha de tempo do jogo com uma gravação de vídeo do jogo, para facilitar a compreensão dos picos de tweets no gráfico e relacioná-los com o tempo do jogo em que esse pico ocorreu"*. Além disso, outra sugestão foi *"criar uma conexão da linha do tempo com notícias do momento, para uma melhor compreensão do comportamento dos usuários ao longo do tempo"* e as possíveis conexões entre a mídia tradicional e Twitter.

Assim, as sessões de grupos focais mostraram que as técnicas de visualização são ferramentas relevantes para a identificação de padrões e desvio de padrões. Para os jornalistas, esses

²Relatório 2014 da Copa do Mundo da FIFA Brasil™: Brasil x Alemanha 1 : 7. Disponível on-line em: <http://goo.gl/AN2Rv7>

padrões são importantes, pois geralmente indicam eventos dotados de informações de valor, cujos atributos os tornam propensos a serem incluídos em notícias. A partir da identificação desses fenômenos, os jornalistas podem buscar documentos, fontes e pessoas que expliquem o contexto, a fim de informar o público sobre esse evento. Com isso, podemos inferir que, mesmo antes de serem produtos de notícias, as visualizações de dados são ferramentas para auxiliar na construção dessas notícias.

3.1.3 Conclusões

A análise dos grupos focais permitiu uma compreensão de pontos importantes a serem considerados em visualizações, como cores utilizadas para representar as informações, técnicas que poderiam ser adicionadas para uma melhor compreensão dos dados, como vídeo, e necessidade de maior interação nos gráficos. A partir dos relatos sobre as informações obtidas das visualizações apresentadas, percebemos que as visualizações de dados são ferramentas relevantes para apoiar a construção de notícias e uma profunda compreensão dos eventos, e os jornalistas apontaram recursos que poderiam melhorar as visualizações apresentadas, como o aumento das técnicas de visualização interativa.

Acreditamos que não é suficiente simplesmente ter essas análises para alcançar nosso objetivo final de estabelecer um ou mais modelos de interação em visualização de dados. No entanto, percebemos que essas análises são um passo inicial nessa direção. Acreditamos que a compreensão das necessidades dos usuários (analistas de dados), em relação às técnicas de visualização e de recursos como a interação, são fundamentais no processo de elaboração de um modelo que possa apoiar o *design* e a criação de visualizações interativas.

3.2 Estudo 2 - Survey Online com Jornalistas

O segundo estudo teve como objetivo conhecer a opinião de jornalistas e profissionais de mídia sobre técnicas de visualizações interativas, compreender se jornalistas e profissionais de mídia conhecem e usam ferramentas de visualização de dados em seu contexto de trabalho, bem como se eles consideram esses recursos importantes [SCT⁺17]. Assim como o estudo 1, este faz parte de um projeto de pesquisa maior, realizado no âmbito do Programa de Apoio a Integração entre Áreas/PRAIAS, da PUCRS, (Editais 07/2014 e 07/2015). Para este estudo, nós coletamos e processamos *tweets* durante a campanha presidencial no Brasil em 2014 e desenvolvemos visualizações interativas para mostrar os sentimentos dos eleitores *online* sobre os candidatos. As visualizações foram apresentadas a jornalistas e profissionais da mídia por meio de uma *survey online* e as respostas obtidas serviram de base para nossas análises e discussões dos resultados.

3.2.1 Metodologia de pesquisa

Seguimos a mesma metodologia do estudo 1 (apresentado na seção 3.1), realizando uma pesquisa descritiva [LFH10] em que fizemos uma descrição sobre as opiniões de jornalistas e profissionais de mídia sobre as visualizações de dados interativas que criamos, suas opiniões sobre visualização de dados em geral e o quanto elas poderiam auxiliar em seus trabalhos, obtidas por meio de uma *survey online*. Além disso, o nosso objetivo era atingir um número maior de jornalistas e refinar nossas observações sobre nossos estudos. A *survey* é um conjunto de questões que são solicitadas a uma pessoa responder. É um tipo de coleta de dados que geralmente não é aprofundada e abrange um grande número de respostas obtidas rapidamente de uma população que está geograficamente dispersa [LFH10].

A *survey* foi escrita em português, composta por 29 perguntas divididas em três seções: a primeira seção, com as primeiras doze questões, eram sobre o perfil dos participantes e suas experiências com redes sociais. A segunda seção apresentou perguntas (da 13 até a 24) para coletar informações sobre as 4 visualizações interativas de dados desenvolvidas e as percepções dos participantes sobre elas. Para cada visualização havia uma pergunta fechada e três perguntas abertas. As visualizações foram apresentadas por meio de gravações de tela com narração, disponibilizadas no Youtube. Finalmente, na terceira seção (da pergunta 25 até a 29) perguntamos - por meio de perguntas abertas - as opiniões dos participantes em relação a visualização de dados em geral, com perguntas abertas. O resultado pretendido da pesquisa era sintetizar as opiniões sobre os elementos apresentados em uma visualização interativa. A pesquisa foi distribuída por e-mail para pessoas que haviam sido pesquisadas previamente como tendo um perfil relacionado ao nosso público-alvo, e distribuída também em listas de discussão e/ou grupos específicos em redes sociais de profissionais que trabalham com jornalismo e com dados de mídias sociais. Um total de 50 pessoas responderam à nossa *survey*, conforme perfil apresentado na tabela 3.2. Quase todos os participantes (45) têm conta no Twitter e 37 usam este meio como fonte de informação. 30 deles já usaram infografia ou visualizações de dados em suas atividades de trabalho.

Em relação às visualizações apresentadas, em todas o eixo x representa o tempo durante os dias entre o primeiro e o segundo turno da eleição. Utilizamos a biblioteca D3.js³ para criação das visualizações interativas com base no nosso interesse em analisar o quanto a interação pode influenciar (ou não) a exploração dos dados por parte dos usuários. Basicamente, todas elas possuem uma estrutura mínima similar, como mostrado na figura 3.5: 1 e 2 - um conjunto de controles auxiliares, 3 e 4 - linha de tempo, e o componente de *brush*, respectivamente. No Gráfico 1 (figura 3.5 e 3.6), o eixo y representa sentimentos positivos (no topo do eixo) e negativos (na parte inferior). A linha laranja representa sentimentos sobre a candidata Dilma, e a linha azul representa sentimentos sobre o candidato Aécio. Pelo gráfico, podemos notar que existem mais *tweets* positivos sobre a Dilma do que sobre o Aécio. Uma característica importante é o componente *brush*,

³<https://d3js.org/>

Tabela 3.2 – Perfil dos participantes do estudo 2 - *survey online*.

ID	Idade	Sexo	Estado	Formação	Grau
P1	31	M	RS	Jornalismo	Mestrado/Doutorado
P2	38	F	RS	Comunicação social, Jornalismo e Direito.	Mestrado/Doutorado
P3	24	F	RS	Comunicação Social	Graduação
P4	39	F	RS	Jornalismo	Mestrado/Doutorado
P5	34	M	RS	Comunicação Social	Mestrado/Doutorado
P6	33	F	SC	Jornalismo	Mestrado/Doutorado
P7	26	M	RS	Jornalismo	Mestrado/Doutorado
P8	62	M	RS	Jornalismo	Especialização
P9	26	M	RS	Jornalismo	Graduação
P10	29	F	RS	Jornalismo	Mestrado/Doutorado
P11	27	M	SC	Jornalismo	Graduação
P12	35	M	RS	ciência da computação	Especialização
P13	25	M	RS	Jornalismo	Especialização
P14	27	F	RS	Comunicação Social	Mestrado/Doutorado
P15	25	M	RJ	Jornalismo	Graduação
P16	37	F	RS	Comunicação Social	Mestrado/Doutorado
P17	26	F	RS	Jornalismo	Especialização
P18	23	M	RS	Jornalismo	Graduação
P19	29	M	RS	Jornalismo	Mestrado/Doutorado
P20	52	F	RS	Jornalismo	Especialização
P21	22	M	RS	Jornalismo	Graduação
P22	28	M	RS	Jornalismo	Graduação
P23	22	F	RS	Jornalismo	Graduação
P24	44	F	RS	Jornalismo	Graduação
P25	25	M	SP	Jornalismo	Graduação
P26	33	F	GO	Jornalismo	Graduação
P27	62	M	PI	Comunicação Social	Mestrado/Doutorado
P28	39	M	SP	Jornalismo	Mestrado/Doutorado
P29	24	M	SP	Jornalismo	Especialização
P30	21	M	PR	Jornalismo	Graduação
P31	25	M	MG	Analista de Sistemas	Graduação
P32	26	F	RJ	Jornalismo	Especialização
P33	22	M	RJ	Jornalismo	Graduação
P34	27	M	MG	Jornalismo	Graduação
P35	24	F	MG	Sistemas de informação	Graduação
P36	35	M	RJ	Jornalismo	Mestrado/Doutorado
P37	27	F	MG	design gráfico	Mestrado/Doutorado
P38	37	M	MG	Jornalismo	Mestrado/Doutorado
P39	32	M	SC	Jornalismo	Mestrado/Doutorado
P40	37	M	MG	Jornalismo	Mestrado/Doutorado
P41	47	F	PR	Jornalismo	Mestrado/Doutorado
P42	53	M	BA	Jornalismo	Especialização
P43	34	F	SP	engenharia	Mestrado/Doutorado
P44	48	M	MG	Comunicação Social	Especialização
P45	39	F	MG	Jornalismo	Mestrado/Doutorado
P46	24	F	DF	Comunicação Social	Graduação
P47	33	F	CE	Jornalismo	Graduação
P48	21	F	SP	Jornalismo	Graduação
P49	42	M	SP	Jornalismo	Mestrado/Doutorado
P50	54	M	SP	engenharia	Mestrado/Doutorado

que permite focar em um intervalo de tempo específico e expandir a área do gráfico (aumentar ou diminuir o *zoom*). A figura 3.6 mostra o gráfico de linhas interativo após o *zoom*.



Figura 3.5 – Estrutura básica comum a todas as visualizações.

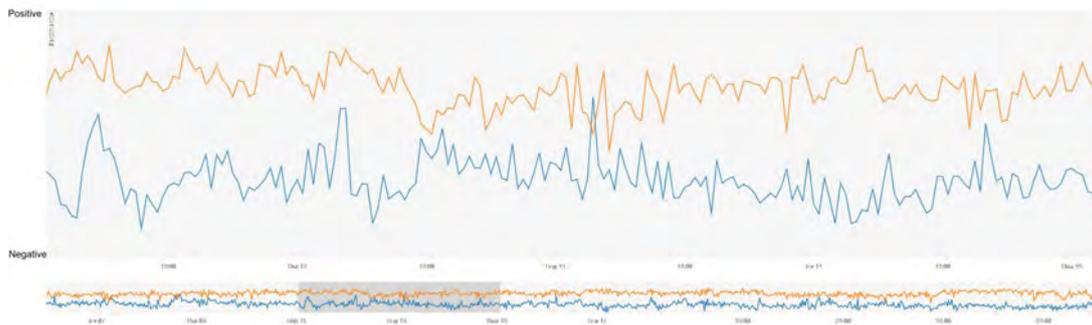


Figura 3.6 – Gráfico de linhas interativo após o zoom (Gráfico 1).

O Gráfico 2 (figura 3.7) é um gráfico multilinhas interativo em que o eixo y representa a quantidade (a soma) de *tweets* positivos e negativos coletados de ambos os candidatos, Dilma e Aécio. O conjunto de controles auxiliares permite escolher a frequência, variando de 5, 10, 15 e 30 minutos a 1 hora. É possível notar que picos mais baixos ocorreram durante o amanhecer, e há picos quando os debates entre os candidatos foram realizados na televisão.

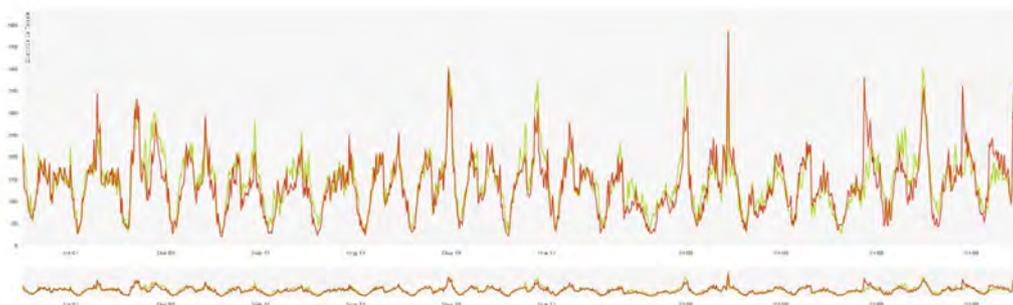


Figura 3.7 – Gráfico multilinhas interativo com a quantidade de *tweets* positivos e negativos ao longo do tempo (Gráfico 2).

O Gráfico 3 (figura 3.8) é um gráfico de círculos interativo que mostra a quantidade de *tweets* de ambos os candidatos, Dilma e Aécio. O sentimento é indicado por cor: verde para positivo e vermelho para negativo. O tamanho do círculo indica a quantidade de *tweets*, e os círculos maiores se posicionam no topo do gráfico. Nesta visualização, o conjunto de controles

auxiliares permite escolher uma representação visual por assunto, portanto, os *tweets* podem ser separados por candidato. Ao posicionar o mouse sobre um círculo, a quantidade equivalente de *tweets* é mostrada e, ao clicar em qualquer círculo, uma nova representação visual é aberta com *tweets* daquele momento, separados em dois grupos (positivo e negativo).

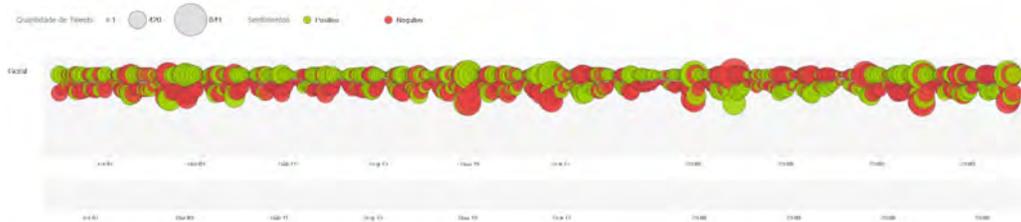


Figura 3.8 – Gráfico de círculos interativo (Gráfico 3).

A última visualização (Gráfico 4) é um gráfico de linhas interativo integrado com notícias (figura 3.9), fruto de uma sugestão levantada em nosso estudo 1 [STT⁺16]. O gráfico é o mesmo que o Gráfico 1, mas agora com notícias agregadas. A linha de tempo das notícias foi construída usando uma ferramenta chamada Timeline.js⁴ usando notícias (também coletadas do Twitter), sobre esse assunto, durante o mesmo período dos *tweets* usados nas visualizações. Posicionando o mouse sobre uma notícia, aparece uma tela mostrando o *tweet* original e, ao clicar sobre ele, essa publicação é aberta no Twitter em uma nova guia do navegador, para que os usuários possam ver mais detalhes. Esse tipo de visualização permite, por exemplo, especular ou criar hipóteses sobre momentos de picos altos no gráfico de linha e as notícias postadas naquele momento.

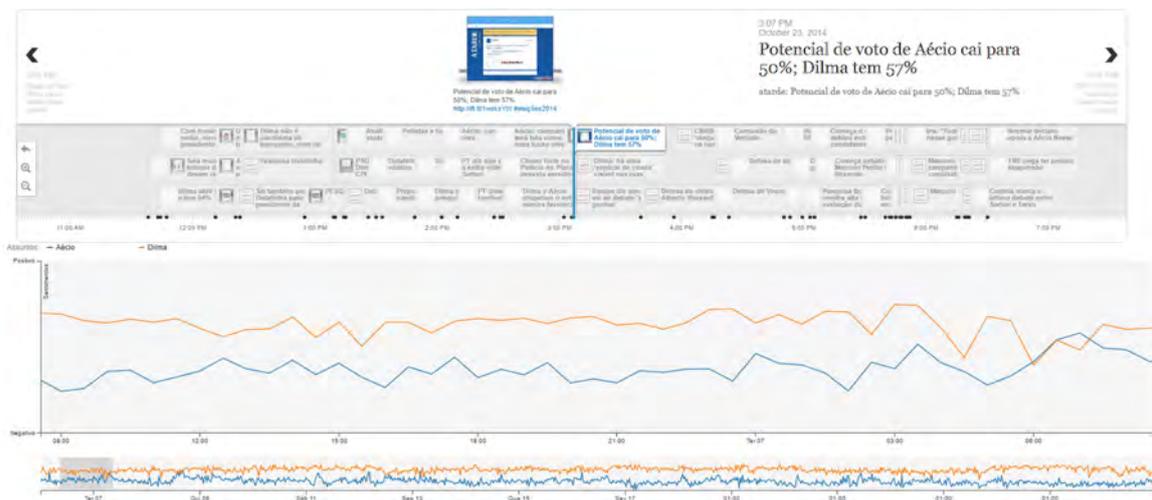


Figura 3.9 – Gráfico de linhas interativo integrado com notícias (Gráfico 4).

⁴<https://timeline.knightlab.com/>

3.2.2 Resultados

Em geral, as visualizações foram consideradas satisfatórias e os participantes disseram que, dessa forma, é possível interpretar um grande volume de dados de forma mais dinâmica. A visualização em forma de linhas (gráficos 1 e 2) foi a que os participantes mais gostaram. No entanto, também foi amplamente sugerido que o uso de gráficos de barras poderia ser capaz de apresentar uma melhor separação de camadas de dados. P12 sugeriu que, associado ao gráfico de linhas poderia ter outros tipos de gráficos, como o de pizza, ou infográficos, com imagens. P26 sugeriu adicionar ao gráfico de linhas fatos importantes ocorridos naquele período. Outra sugestão interessante foi a de P29, *“pesquisas de intenção de voto inseridas pontualmente no gráfico complementar as informações e mostrariam contradições ou semelhanças nos dados recolhidos”*. Muitas respostas convergiram em agregar mais dados e informações textuais e visuais, principalmente em momentos de pico, para ajudar na compreensão do contexto e dos acontecimentos. Muitos sugeriram acrescentar gráfico com menções, comparações com publicações em outras mídias, animação e movimento. P37 sugere que a *“marcação de alguns eventos relevantes relacionados às variações das curvas (acabaria sendo mais um infográfico do que gráfico), ajudaria a dar exemplos de leituras possíveis do gráfico e ilustraria parte do tipo de interesse nessas visualizações”*.

O Gráfico 3 foi considerado difícil de entender e confuso. P20 escreveu: *“Achei uma confusão, não consigo identificar nada.”* Não somente para este, mas para todos os gráficos, sugeriram agregar informações de acordo com as regiões do país, pois consideram que isso valorizaria informações obtidas por meio das visualizações. Em geral, os participantes não conseguiram compreender o Gráfico 3, mas, ressaltaram que o recurso de dar *zoom* na linha de tempo ajuda a, pelo menos, visualizar com mais detalhe os dados.

Outra questão importante foi a possibilidade de subdividir o gráfico em unidades menores. A ideia de um único gráfico que abrange todas as informações - mesmo com recursos interativos, como a capacidade de se concentrar em períodos - não parece ser suficiente para garantir uma compreensão adequada. Foi sugerido acrescentar seleção por região e por menções. O Gráfico 4 foi altamente elogiado pela possibilidade de visualização das relações entre os *posts* e as notícias. Para P2, foi a visualização *“mais completa”*, porém acha que *“poderiam usar blocos de cores e não linhas. Gosto do detalhamento, mas acho que complexifica a interpretação”*. Os comentários sobre o Gráfico 4 reforçam o quanto ele foi o que mais agradou os participantes. P16 escreveu: *“adorei a ligação com notícias! Seria possível destacar por cores os veículos ou grupos de veículos de comunicação?”* E ainda exemplificou como o agradaria essa solução, por exemplo visualizar *“blogs, versões online de jornais, portais de notícia, sites segmentados, etc. Seria legal para sabermos que tipo de publicação mais repercute em mídias sociais.”*

Sobre a influência da visualização de dados na prática jornalística, como uma possível forma (ou fonte) de gerar pautas, percebemos que as ferramentas de análise de dados ainda não são facilmente acessíveis (ou conhecidas) e, portanto, ainda não são tão influentes. Uma das res-

postas pode retomar bem isso: *“quando ferramentas de cruzamento de dados estiverem disponíveis para manipulação de repórteres de forma simples e prática serão universos interessantes para garantir pautas. A lei de acesso às informações tem sido de grande valia, mas falta investimento público para manter os sites mais acessíveis”* (P16). Já P30, acha que não, pois *“há que se ter os critérios de notícia. A visualização enriquece a apresentação mas não substitui a notícia”*. E P43 afirma que, sobre isso, acha que é *“possível gerar debate, pois a escolha inicial de pauta em si já foi feita quando se escolheu mexer com este e não qualquer outro conjunto de dados (ou combinação de dados).”*

3.2.3 Conclusões

Uma vez que as mídias sociais tendem a estar cada vez mais integradas às práticas jornalísticas, apresentamos os resultados de uma pesquisa que analisa a forma como as visualizações interativas dos dados do Twitter podem beneficiar os jornalistas e os profissionais de mídia em suas atividades diárias. Relatamos um estudo empírico qualitativo baseado em uma *survey online* com 50 participantes, para entender se as informações obtidas a partir das visualizações interativas apresentadas podem apoiar atividades de trabalho de jornalistas e profissionais de mídia. Exploramos o domínio da análise jornalística em resposta aos dados extraídos do Twitter, incluindo implicações para um melhor *design* de visualizações interativas. Nossa análise de dados revela que os participantes consideraram a visualização de dados um recurso valioso em suas atividades de trabalho, no entanto, ela precisa ser melhor introduzida nas práticas de trabalho desses profissionais, e seu uso precisa ser melhor estudado.

Talvez possamos fazer uma analogia com a introdução do CSCW (*Computer-Supported Cooperative Work*) nos processos de trabalho nas empresas. Assim como a comunidade de CSCW, a comunidade de visualização de dados deve realizar mais estudos com o objetivo de compreender como os trabalhos de profissionais de mídia, jornalistas e analistas de dados de mídias sociais ocorrem na prática. Alinhados com a pesquisa de Brooks [Bro15], acreditamos que deve ser dada maior atenção à compreensão dos usuários de visualização de dados e ao contexto em que o trabalho de análise de dados ocorre.

Neste estudo, percebemos sugestões que vão ao encontro da proposta da tese, em relação a visualizações narrativas, quando os participantes sugeriram visualizações similares a infográficos, com mais recursos para contextualização dos dados, como vídeo, animação, texto explicativo, entre outros.

3.3 Estudo 3 - Mapeamento sistemático da literatura sobre visualização de dados e *storytelling*

O estudo 3 foi realizado com o intuito de aprofundar os estudos sobre visualização de dados associadas a narrativas e mapear o estado da arte das pesquisas nesta área. Há mais de 15 anos, pesquisadores da área de visualização de dados começaram a discutir e, um pouco mais tarde, a implementar o uso de narrativas em visualizações, fornecendo maneiras, por exemplo, de demonstrar melhor as estruturas temporais e as relações causais entre fatos e eventos. Eles acreditam que essa integração pode estimular a exploração, análise e comunicação de dados.

Com base nisso, fizemos um mapeamento sistemático da literatura em visualização de dados e narrativas, com o objetivo de fornecer uma visão do estado da arte da pesquisa acadêmica neste campo. A nossa principal questão se refere a “como a narrativa tem sido utilizada na visualização de dados e qual é o estado atual da pesquisa acadêmica neste campo?”. Aplicamos o método de estudo de mapeamento de Kitchenham e Charters [KC07] para pesquisar e selecionar artigos, extrair e sintetizar os dados deles.

3.3.1 Metodologia

Para realizar este mapeamento sistemático, seguimos as diretrizes de Kitchenham e Charters [KC07] para estruturar e organizar nossa pesquisa. De acordo com Kitchenham [Kit04], uma revisão sistemática da literatura (SLR) é um método para avaliação e interpretação de tópicos relevantes para uma questão de pesquisa, assunto ou evento de interesse. Existem dois tipos de revisões sistemáticas: *i*) as reflexões convencionais [PR08] [KC07], que agregam resultados sobre a eficácia de um tratamento, intervenção ou tecnologia, e que estão relacionadas a questões específicas de pesquisa; e *ii*) estudos de mapeamento sistemático (MS) [AO05], que visam identificar todas as pesquisas relacionadas a um tópico específico para responder a questões mais amplas relacionadas às tendências da pesquisa na área.

De acordo com Kitchenham [Kit04], o mapeamento sistemático é uma ampla revisão de estudos em uma área específica que visa identificar evidências sobre o tema. Ele “*permite a identificação de clusters de evidências e mostra os gaps para direcionar o foco de futuras revisões sistemáticas e identificar áreas para mais estudos primários a serem conduzidos*”. Nosso trabalho envolve um mapeamento sistemático sobre o estado da arte da visualização de dados com narrativas. Entendemos que há uma lacuna importante na literatura, considerando que não encontramos artigos que mostrem o estado da arte da visualização de dados com narrativas de uma maneira abrangente. A seguir descrevemos as etapas metodológicas utilizadas, incluindo questões de pesquisa e o processo de seleção de artigos.

Questões de pesquisa

A construção do protocolo de revisão começou escolhendo as questões de pesquisa para orientar o processo. Nosso objetivo foi produzir um mapeamento entre visualização de dados e narrativas que respondesse as seguintes questões:

RQ1: Quais são os principais tópicos investigados na pesquisa sobre visualização de dados com narração de histórias?

RQ2: Quais são os métodos de pesquisa utilizados nos estudos?

RQ3: Em que contexto esses estudos foram aplicados?

RQ4: Quais são as técnicas de visualização mais comuns na visualização de dados com narração de histórias?

RQ5: Quais são as principais contribuições e desafios da visualização de dados com narração de histórias?

Busca e processo de exclusão

Nós pesquisamos publicações em periódicos e conferências em busca de artigos sobre visualização de dados e *storytelling* publicados até novembro de 2015. A busca automatizada foi realizada usando sistemas como ACM digital library⁵, IEEE Xplore digital library⁶, Elsevier Science-Direct⁷, Scopus⁸, and Wiley⁹. A *string* de busca foi construída a partir de dois termos de pesquisa: visualização e *storytelling*. Para aumentar a sensibilidade da pesquisa, foram fornecidos sinônimos para cada termo e adicionados plurais, nos casos em que os sistemas de busca não tratavam os caracteres como esperado. Para construir a *string*, os sinônimos foram agrupados e, dentro de seu grupo, foram separados por OR, e os grupos de sinônimos para cada termo foram agrupados com AND, como mostrado na figura 3.10. Durante o processo de busca, excluimos os artigos que preenchessem um dos seguintes critérios de exclusão:

- Escrito em um idioma diferente do inglês;
- Publicado em áreas fora da computação.

(visualization OR visual)
AND
(narrative OR storytelling OR story OR stories)

Figura 3.10 – *String* de busca.

⁵<http://portal.acm.org>

⁶<http://www.ieeexplore.ieee.org/Xplore>

⁷<http://www.sciencedirect.com>

⁸<http://www.scopus.com>

⁹<http://http://www.wiley.com>

Seleção de estudos

Os critérios de seleção de artigos buscaram identificar aqueles que forneciam relações diretas com a questão da pesquisa. A seleção de artigos relevantes foi realizada em duas etapas. Na primeira, foram analisados 3.471 artigos, resultantes do procedimento de busca, por meio da leitura do título, resumo e palavras-chave. Todos os trabalhos que claramente não eram relevantes para as questões de pesquisa eram excluídos, como por exemplo trabalhos com foco na criação de filmes e animações. Alguns artigos precisavam de uma análise mais cuidadosa porque estavam claramente focados no assunto, mas um dos termos da *string* de busca não estava presente no título, no resumo ou nas palavras-chave. Além disso, alguns artigos não tinham resumo ou palavras-chave, o que nos fez adaptar o processo.

No segundo passo, os critérios de exclusão foram aplicados ao conjunto de 365 artigos resultantes da primeira etapa. Dois pesquisadores (sendo a pesquisadora desta tese um deles) trabalharam de forma independente analisando todo o conjunto de artigos. Eles leram o resumo, a introdução e a conclusão de cada artigo e, em seguida, aplicaram os critérios de exclusão a artigos que não eram potencialmente relevantes ou que não estavam relacionados à nossa questão principal. Um terceiro pesquisador trabalhou para resolver discrepâncias e, depois disso, as diferenças e dúvidas foram resolvidas em uma reunião de consenso. Criamos uma lista de artigos excluídos no segundo passo com a explicação do motivo da exclusão. Essa etapa resultou na seleção de 73 trabalhos. Considerando os sistemas de busca usados e sistemas de indexação, o IEEEXplore foi a fonte que mais retornou artigos, contando com 48 dos 73 artigos selecionados (Tabela 3.3).

Tabela 3.3 – Resultado do processo de seleção.

Fonte	Recuperados	Pre-seleção	Seleção
ACM digital library	568	56	19
IEEEXplore digital library	541	82	48
Elsevier ScienceDirect	572	16	1
Scopus	1208	203	2
Wiley	582	8	3
TOTAL	3471	365	73

Extração de dados

De acordo com Kitchenham [KC07], o processo de extração de dados para estudos de mapeamento é amplo e pode ser denominado como uma classificação ou estágio de categorização. O objetivo é classificar documentos com detalhes suficientes para responder as questões de pesquisa e identificar documentos para revisões posteriores. Nesta extração de dados, cada pesquisador trabalhou de forma independente, com base em um formulário de extração (feito no MS Excel). A seguir, listamos os dados extraídos de todos os artigos selecionados:

- Identificação do artigo

- Número de citações no Google Scholar
- Ano de publicação do artigo
- Local de publicação do artigo
- Título do artigo
- Nomes de todos os autores
- Classificação como Teórico ou Apicado
- Tópicos principais investigados
- Métodos de pesquisa usados no estudo
- Contexto do artigo
- Técnicas de visualização usadas
- Contribuições principais

Semelhante ao processo que fizemos na seleção do estudo, dois pesquisadores trabalharam nesta tarefa para todos os campos de dados e um terceiro pesquisador analisou desentendimentos nos dados extraídos. As diferenças foram resolvidas em uma reunião de consenso.

3.3.2 Resultados

Nosso principal objetivo nesta subseção é resumir informações sobre os artigos revisados neste mapeamento. As referências completas dos artigos selecionados são apresentadas no Apêndice B. Nossa pesquisa recuperou mais de 3.000 trabalhos, dos quais selecionamos 73 artigos publicados entre 2001 e 2015, sendo quase 65% deles foram publicados depois de 2006 (figura 3.11). Além de classificar as técnicas de visualização usadas, agrupamos as estratégias de *design* para visualizações narrativas, incluindo técnicas de interação e formatos narrativos.

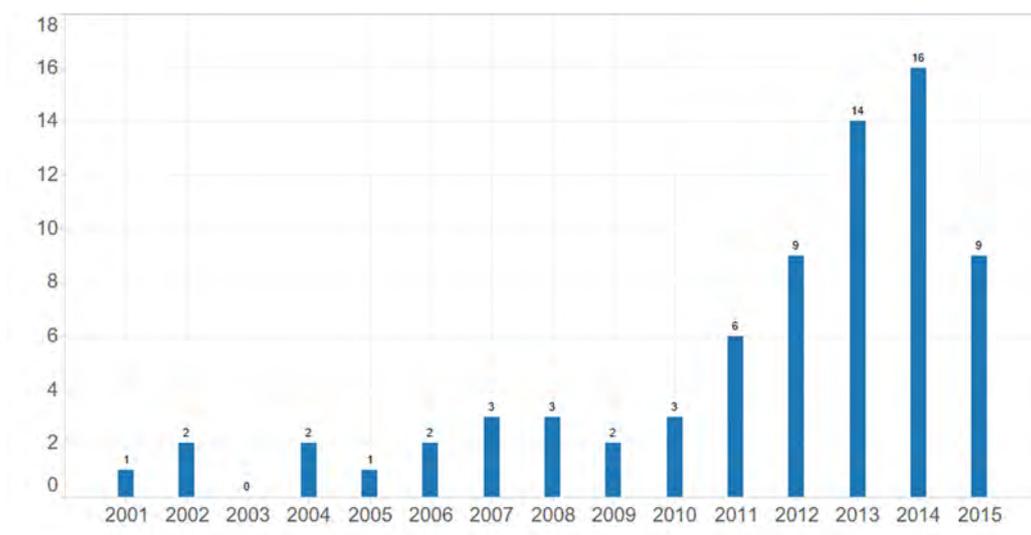


Figura 3.11 – Distribuição dos artigos ao longo dos anos.

Quanto à autoria dos 73 artigos, ao todo são 234 autores distintos¹⁰ identificados. Analisamos também o país de origem da instituição vinculada aos autores. É importante ressaltar que nem todos os trabalhos forneceram sua instituição ou país e, nesses casos, os autores não foram considerados em nossa análise. A maioria deles é dos EUA (110), seguido pelo Reino Unido (32) e outros países com menor número de autores.

Em relação aos autores que começaram a trabalhar com narrativa e visualização, os primeiros que colaboraram foram Gershon e Page [GP01] (S1), que discutiram a contribuição que a narrativa poderia dar à visualização. Algum tempo depois, Segel e Heer [SH10] (S17) reinventaram essa noção de contar história com dados, em visualizações, chamando-a de visualização narrativa. Eles estudaram elementos de visualização e identificaram alguns padrões e estruturas que a mídia usa para apresentar visualizações narrativas. Estes são os dois artigos mais citados na área, de acordo com as citações do Google Scholar, como mostrado na figura 3.12.

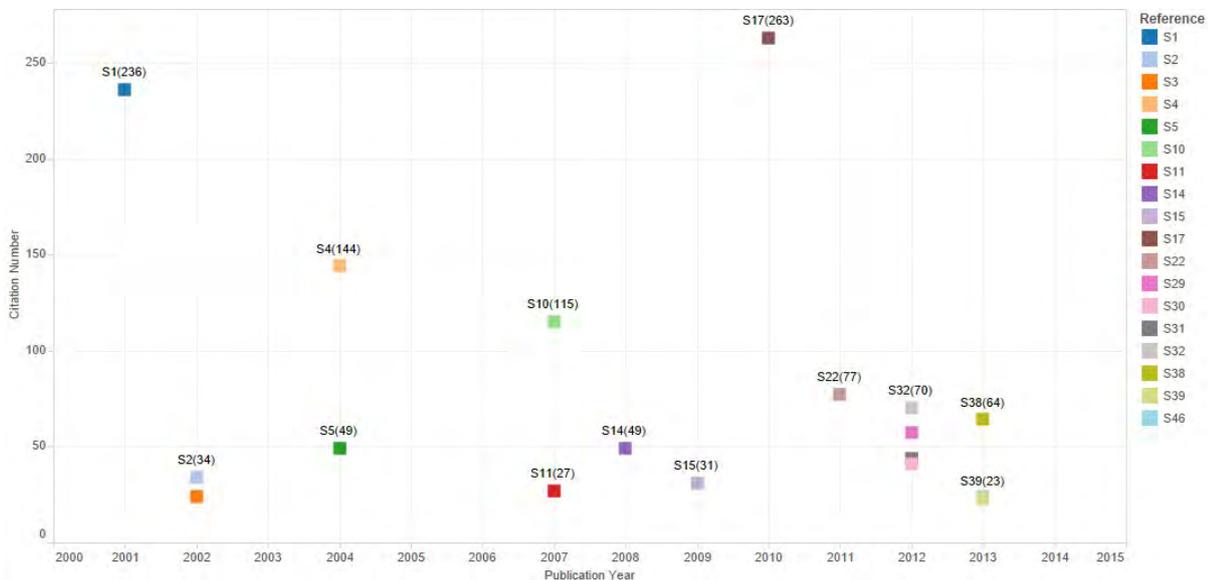


Figura 3.12 – Artigos com mais de 20 citações segundo o Google Scholar Citations (em 11 de fevereiro de 2016).

Outros artigos mais citados são Viegas et al. [VBN⁺04] (S4) e Eccles et al. [EKHW07] (S10), que apresentaram ferramentas para identificar padrões temporais em conjuntos de dados. Em S4, os autores criaram duas visualizações projetadas para revelar padrões sociais nos arquivos de e-mail. Essas visualizações fornecem aos usuários maneiras acessíveis de ver padrões de alto nível de troca de e-mails ao longo do tempo. De acordo com os autores, as ferramentas apresentam um retrato pessoal de um indivíduo através do contexto de suas interações de e-mail. No S10, os autores apresentaram um sistema que detecta padrões geo-temporais e integra a narração da história para aumentar a coesão analítica de fazer sentido. Eles apresentaram como o conceito de história se relaciona com o processo de análise e como ele pode ser apoiado na geovisualização. Segundo os autores, a narrativa é um meio efetivo de comunicação e colaboração e é usada para relatar eventos complexos aos tomadores de decisão. Ao contrário dos dois artigos mais citados,

¹⁰Todos os co-autores do mesmo artigo foram considerados também autores.

estes aplicaram o conceito de narrativa na visualização através da criação de ferramentas e coletaram *feedback* dos usuários (nas avaliações das ferramentas).

Respondendo as questões de pesquisa

Cinco questões de pesquisa nortearam o processo de extrair dados dos 73 artigos selecionados (Apêndice B). A seguir, descreveremos os resultados do nosso mapeamento que respondem às questões de pesquisa levantadas.

RQ1: Quais são os principais tópicos investigados sobre visualização narrativa?

Os tópicos de pesquisa foram categorizados usando categorias criadas manualmente a partir da análise dos artigos. Nós descrevemos brevemente cada tópico de pesquisa em relação a como eles tratam a narrativa na visualização nos artigos investigados. Alguns estudos estão relacionados a mais de uma questão de pesquisa e envolvem mais de um tópico de pesquisa.

O tópico de pesquisa de **Métodos e técnicas para narrativa** considera modelos, estruturas e ferramentas propostas para contar histórias com dados. Percebemos que vários trabalhos se concentraram em tempo, evento e localização. Esses trabalhos envolveram a compreensão do espaço de *design* em contextos específicos, como biologia, jornalismo, linguística e assim por diante, para as técnicas de visualização narrativa propostas. Permitiram também saber quais técnicas são mais apropriadas para cada contexto e que são usadas na maioria deles.

Análise, Engajamento e Significado é um tópico relacionado a artigos que tratam o processo de análise, a criação de sentido, o engajamento de usuários e o significado de dados. É investigado do ponto de vista de como apresentar dados para melhor compreendê-los. Esses artigos examinam as necessidades dos usuários no processo de análise e os comentários dos usuários sobre ferramentas, modelos e *frameworks*. Este tópico também depende do contexto, uma vez que cada analista de dados ou equipe de analista de dados possui suas próprias necessidades.

Processamento de dados e estrutura de dados é um tópico de pesquisa que discute estrutura de dados, técnicas de processamento de dados, ontologias e formas de preparar dados para apresentá-los. Este tópico também discute técnicas para extrair automaticamente informações de um conjunto de dados, como eventos, opiniões e relacionamentos. Alguns estudos focaram neste tópico como parte da construção de uma visualização narrativa.

Visualização narrativa na pesquisa é um tópico de pesquisa explorado a partir da perspectiva de como esse assunto é investigado. Os artigos consideram a pesquisa de narrativas com visualização de dados e seus limites em vários campos, incluindo gráficos computacionais, psicologia e ciências sociais, por exemplo. Todos os trabalhos têm contribuições para este tópico, mas aqui escolhemos aqueles que discutem práticas, tendências e desafios na visualização narrativa.

RQ2: Quais são os métodos de pesquisa usados nos artigos?

Responder a RQ2 foi um desafio, porque apenas alguns trabalhos analisados apresentaram uma metodologia explícita de pesquisa (ou métodos de pesquisa relacionados). Considerando este fato, destacamos que alguns dos métodos aqui discutidos foram inferidos por nós a partir da análise feita (e não foram explicitamente indicados pelos autores). Nós analisamos artigos pesquisando informações sobre as seguintes etapas básicas de pesquisa: Revisão de Literatura (se apresentou ou não no documento), *Design* (se foi apresentado ou não, ou métodos de pesquisa relacionados), Implementação (se foi apresentada ou não) e Avaliação (se foi apresentada ou não, ou métodos de pesquisa relacionados).

Quase todos os trabalhos analisados começaram por uma revisão da literatura, mostrando pesquisas usadas como base ou como trabalhos relacionados às suas pesquisas (apenas alguns apresentaram o sistema proposto ou abordagem diretamente). Alguns artigos não apresentaram qualquer descrição da fase de projeto, com foco direto na implementação do sistema. No caso de implementações, alguns deles apresentaram fases tradicionais de Engenharia de Software como a coleta de requisitos, a elaboração de casos de uso, ou o uso de histórias de usuários. Considerando aqueles que descreveram os métodos de pesquisa usados, encontramos uma referência explícita ao uso de abordagens centradas no usuário ou centradas no ser humano, bem como processos de projeto iterativo e *design* participativo. As entrevistas foram um método de pesquisa muito utilizado. Outros métodos centrados no usuário, como observações, grupos focais, pesquisa de usuários e *workshops* também foram citados.

Mais da metade dos trabalhos analisados apresentaram algum tipo de implementação. Levando esses artigos em conta, quase 70% apresentaram uma avaliação, algumas focadas na implementação em si, considerando aspectos como eficiência e eficácia, desempenho ou flexibilidade, e outras focadas na perspectiva dos usuários, descrevendo o uso de abordagens de avaliação como, por exemplo, *feedback* dos usuários, estudos de usuários, validação com especialistas e estudos de campo. Alguns artigos que não apresentaram implementações específicas também apresentaram avaliações das abordagens discutidas (ou ideias). Considerando os dois casos (implementações ou não), os métodos mais utilizados para avaliar (ou demonstrar) as propostas envolveram usuários, como comentado anteriormente, ou estudos de caso.

RQ3: Em que contexto esses artigos foram aplicados?

Para especificar contextos diferentes, analisamos trabalhos tentando identificar o público-alvo, a fonte de dados e as áreas de conhecimento envolvidas. Percebemos que os três contextos mais comuns foram:

- Documentos gerais, que incluem textos longos como livros, textos online e hipertextos, usados por historiadores ou economistas, por exemplo;

- Notícias e jornalismo online, que se concentram tanto no processo de produção quanto na divulgação;
- Redes sociais, usados para analisar linhas de tempo e informações de escala de cidades.

Os dados científicos, filmes, vídeos e dados de saúde também foram contextos em que alguns artigos foram aplicados. Vários trabalhos não apresentaram um contexto específico e outros foram projetados para fins bastante específicos, como ambientes de aprendizagem, redes e esportes. Figura 3.13 mostra o número total de artigos para cada categoria de contexto identificada.

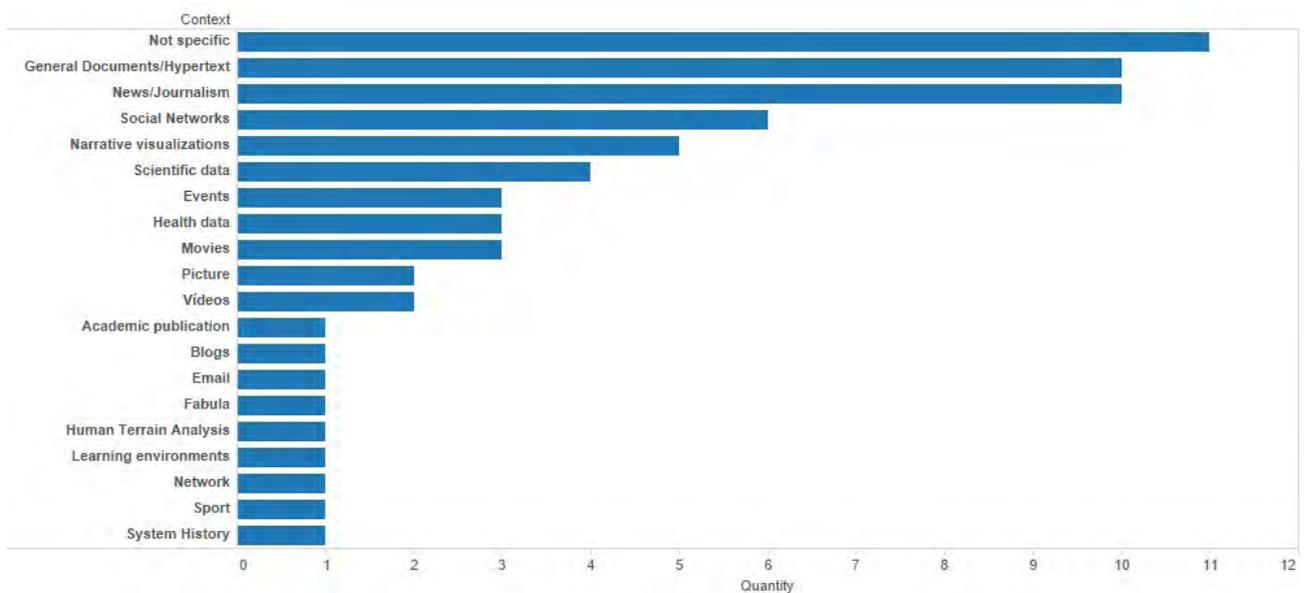


Figura 3.13 – Artigos por categoria de contexto.

RQ4: Quais são as técnicas de visualização mais comuns na visualização de dados com narrativas?

Para responder a RQ4, analisamos todos os trabalhos que forneceram ou usaram visualizações para identificar alguns tópicos ou realizar um estudo de avaliação. Em muitos casos, os autores não forneceram o nome explícito da técnica e apenas explicaram o sistema desenvolvido. Nestes casos, tentamos identificar a técnica utilizada.

Entre 73 artigos, 50 tiveram alguma visualização, o que nos permitiu extrair uma ou mais técnicas usadas. Organizamos a técnicas de visualização em categorias que contêm a subcategoria relacionada, a fim de manter uma organização e fornecer uma melhor visualização de dados. As técnicas e as categorias relacionadas podem ser visualizadas na Tabela 3.4, em que mantemos os nomes desses elementos em inglês, devido ao fato de terem sido extraídos dos artigos selecionados (que estão em inglês). Identificamos 36 técnicas de visualização diferentes. As categorias representam Relações, Imagens, Informações Textuais, Sequências, Proporções, Mapas, Temporal, Outros e Abordagem Híbrida. A categoria Outros representa técnicas de visualização que foram usadas para representar alguma entidade específica (como o caso de "Representação do corpo humano", que foi utilizado para mostrar em qual parte do corpo do paciente a doença foi manifestada).

A categoria de abordagem híbrida foi criada para visualizações que permitem a representação de mais de uma categoria.

As técnicas de visualização mais utilizadas foram as que aparecem no gráfico da figura 3.14 e outras técnicas com menor número de uso. No entanto, se mudarmos a perspectiva e analisarmos os números por categoria principal, 35 dos 50 trabalhos usaram representações de relacionamento e 20 representações temporais. O texto foi usado em 13 trabalhos, seguido de Proporção (6), Outros e Imagem (4), Abordagem Híbrida (2) e Sequência (1).

Tabela 3.4 – Principais técnicas de visualização narrativa classificadas por categoria.

Categoria	Sub-Categoria	Técnicas de Visualização	Total
Relação	Rede	GRAPH	10
		STREAMGRAPH	5
		TREES	3
		TREETABLE	2
		HAIRBALL	1
		HYPertext LINKED	1
		NODE-LINK	1
		SUNBURST	1
		LINE GRAPH	1
	Correlação	LINE CHART	6
TABLES		5	
SCATTER PLOT		3	
HEATMAP		2	
Q-Q PLOT		1	
ACF PLOT		1	
Proporção	BUBBLE CHART	1	
Distribuição	HISTOGRAM	2	
Imagem	Galleria	GALLERY	3
		ALBUM	1
Texto	Palavra	LONG TEXT	6
	Texto	WORD CLOUD	5
		LIST	3
Sequência	Fluxo	FLOW CHART	1
Proporção	Proporção	BAR CHART	4
		PIE CHART	1
		CIRCLES	1
		Other	1
Outros	Outros	MOLECULAR REPRESENTATION	2
		HUMAN BODY REPRESENTATION	1
		DOTS	1
Mapa	Mapa	MAP	16
Tempo	Linha de tempo	TIMELINE	18
		STORYLINE	3
		RADIAL LINES	1
Abordagem híbrida	Tempo + Relação	TOPIC MATRIX	1
	Tempo + Proporção	FLOWLIKE VISUAL METAPHOR	1

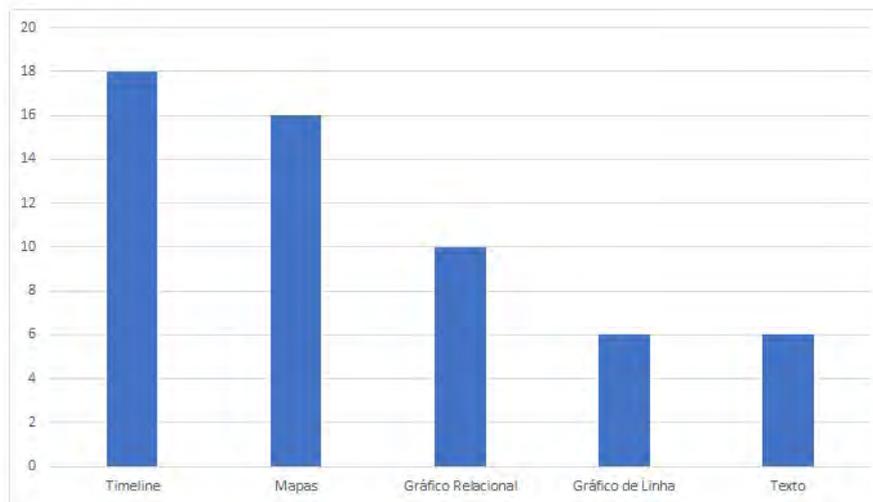


Figura 3.14 – Técnicas de visualização mais utilizadas.

RQ5: Quais são as principais contribuições e desafios da visualização narrativas?

Nossa resposta a esta questão de pesquisa baseia-se nas principais contribuições dos artigos. A introdução da narrativa na visualização não só pode ajudar a interpretação dos dados, mas também ajudar a visualização a se tornar mais atraente e ajudar os analistas no envolvimento da análise. Os estudos variaram desde o discurso teórico sobre esse tema até o desenvolvimento de ferramentas interativas, criando estratégias atraentes para o público.

Entre os 73 trabalhos, 12 apenas discutiram conceitos e estratégias para inserir histórias na visualização, bem como os seus benefícios. Todos os outros, além de discutir esse assunto, propuseram alguns métodos, estruturas, modelos ou ferramentas de visualização narrativa. Quase todos os artigos buscam maneiras de aumentar o envolvimento dos usuários na exploração de dados. Muitos autores afirmam que pensar na visualização em um contexto narrativo pode ajudar a tornar os dados mais compreensíveis, memoráveis e confiáveis. Portanto, é necessário conhecer o público, seu nível de conhecimento, de domínio e familiaridade com as técnicas de visualização. Para criar uma boa visualização, também deve ser dada atenção à coleta e processamento de dados.

As principais contribuições dos artigos são as estratégias usadas para contar histórias com dados. Os pesquisadores tentaram combinar a compreensão do domínio e as necessidades dos analistas de dados, com a ideia de um modelo ou ferramenta usando técnicas de visualização conhecidas, como *timeline*, *storyflow*, detecção de eventos, anotações de texto e gráficos, gráficos (bolha, barra, pizza, linha), etc. Assim, a maioria dos estudos se concentrou no desenvolvimento de técnicas e ferramentas que facilitam a obtenção de informações através de visualizações narrativas interativas.

Entre os desafios, a avaliação foi destacada como uma etapa importante. É necessário entender onde, quando e por que a visualização narrativa é bem sucedida. Além disso, é importante conhecer a eficácia de uma história ao transmitir sua mensagem. Outro desafio é fornecer

diretrizes para fazer uma boa história, permitindo que as pessoas (analistas de dados, por exemplo) conheçam vários gráficos e selecionem ou personalizem o mais atraente.

3.3.3 Discussão

Este estudo procurou contribuir para o domínio de visualizações narrativas. Encontramos algumas pesquisas sobre o assunto e percebemos que esse campo ainda é recente, com muitas discussões e reflexões em andamento. Nossas descobertas mostram uma grande concentração de pesquisas em métodos, técnicas, modelos e ferramentas para visualização narrativa. É importante ressaltar que artigos fortemente relacionados com desenvolvimento e melhoria de técnicas computacionais tiveram mais pesquisas do que estudos relacionados à análise de processos, engajamento e percepção do contexto de pesquisa nessa área (*awareness*).

A partir dos trabalhos analisados, descobrimos que não houve uma prática de descrever sistematicamente os métodos de pesquisa utilizados nos artigos. Percebemos que, nestes artigos, geralmente há apenas uma apresentação da abordagem, modelo ou ferramenta e, em alguns casos, sua avaliação. Na maioria dos trabalhos, os autores não forneceram informações com clareza e não forneceram a razão pela qual eles adotaram aquele formato de narrativa específico. Além disso, eles não descreveram quais os benefícios de usar um formato em vez de outro e quais variáveis devem ser consideradas para justificar essas escolhas. Os estudos também não mencionam a tecnologia usada para criar a visualização como o uso de bibliotecas ou ferramentas, por exemplo.

Acreditamos que os pesquisadores precisam se preocupar com técnicas de visualização e formas de narrativas separadamente. Em termos de *design*, acreditamos que os pesquisadores precisam refletir sobre técnicas de visualização e formas de narrativas, para estudar essas combinações, com vistas a compreender as combinações e restrições que são mais benéficas. Os autores geralmente não explicam os motivos do *design* adotado na narrativa. Por exemplo, apesar de a maioria dos artigos terem algumas técnicas de interação disponíveis para os usuários (26 não apresentaram nenhum), não podemos conhecer o ganho real deles. Além disso, os autores geralmente apenas apontam as interações presentes na interface, sem justificá-las. Em geral, os benefícios ou ganhos não estão nesta granularidade, e são apresentados de forma mais geral. Apontamos que isso deve ser melhor discutido pelos autores, a fim de fornecer um contexto para os leitores, permitindo, com isso, saber se esse *design* pode ser aplicado para os problemas do leitor.

Em sintonia com Andrew Moere [BBC⁺11], concordamos que a relevância da narrativa em relação aos seus objetivos científicos ainda é relativamente inexplorada. Ele argumentou que o uso da narrativa na visualização era considerado como um conjunto de técnicas que tem a capacidade de dirigir a atenção de um espectador através de uma narrativa sequencial e que isso parecia perder o papel do contexto na prática de visualização. No entanto, nesta revisão encontramos artigos focados em contexto, com pesquisadores interessados em aumentar o processo de fazer sentido, o que contradiz seu argumento. Porém, não encontramos visualizações narrativas que

possam ser aplicadas ou adaptadas a diferentes contextos. Assim, podemos levantar algumas questões como "o contexto é relevante para a compreensão da visualização narrativa?", ou "existe alguma relação entre o contexto e as técnicas utilizadas na visualização narrativa?". Outro fator é que muitas técnicas são comuns às visualizações criadas em vários contextos diferentes.

Existe uma característica inerente desta área para promover a multidisciplinaridade, exigindo diferentes disciplinas para trabalharem em conjunto. Os estudos analisados combinaram principalmente pesquisadores de computação e de humanidades ou de ciências sociais, mas também é possível identificar trabalhos envolvendo outras áreas, como saúde e biologia. Como o trabalho envolvido na criação de visualizações narrativas está fortemente relacionado ao contexto, a colaboração através de equipes multidisciplinares é muito importante na pesquisa neste campo. Diferentes contextos encontrados nos estudos também demonstraram seu potencial para apoiar pesquisadores e profissionais de áreas distintas.

Em consonância com Figueiras [Fig14a] e de acordo com os resultados da pesquisa, concordamos que esta área de pesquisa ainda é muito nova e ainda tem pouca informação sobre como introduzir histórias nas visualizações. Há oportunidades extensas para causar impacto neste campo de pesquisa. Embora esta revisão seja um estudo inicial, acreditamos que pode fornecer algumas informações sobre o que o público se preocupa.

É importante destacar o interesse dos pesquisadores da ciência da computação em entender como as pessoas dão sentido aos dados. Isso mostra a necessidade e a importância de entender as necessidades das pessoas e como elas realizam suas análises. Este é certamente um grande passo para melhorar a experiência do usuário na visualização de dados. Uma oportunidade de pesquisa poderia ser melhor estruturar técnicas de visualizações narrativas de dados. Outra oportunidade de pesquisa poderia ser a combinação de muitas funcionalidades em uma única ferramenta ou um conjunto de ferramentas que ofereça autonomia aos usuários para análise, edição ou customização, e apresentação. No entanto, como discutido por Lee et al. [LRIC15], seria necessário refletir sobre se isto é uma boa alternativa.

Em resumo, primeiro destacamos os principais problemas que encontramos em nosso estudo de mapeamento sistemático. Acreditamos que, para entender melhor as ideias e abordagens apresentadas, os *designers* precisam:

- especificar os métodos de pesquisa usados;
- compreender as necessidades dos usuários finais;
- esclarecer o raciocínio de suas decisões de *design*;
- destacar benefícios para usar alguns formatos de narrativa em vez de outros.

Apesar dessas questões, encontramos várias oportunidades de pesquisa. Descobrimos que é necessário:

- verificar se o contexto é relevante para entender a visualização de dados com narrativa;
- verificar se existe uma relação entre o contexto e as técnicas usadas em uma visualização narrativa;

- entender as necessidades dos analistas e como eles fazem suas análises;
- fornecer a combinação de muitas funcionalidades em uma ferramenta ou um conjunto de ferramentas;
- fomentar a multidisciplinaridade, permitindo que diferentes disciplinas trabalhem juntas;
- estruturar melhor as técnicas de contar histórias em visualização de dados.

3.3.4 Conclusões

Consideramos que este mapeamento foi uma grande contribuição para aprofundar os nossos estudos na área e melhorar a nossa compreensão sobre o assunto. De mais de 3.000 documentos encontrados, selecionamos 73 relevantes. Esses estudos investigaram temas amplos em relação à visualização narrativa: métodos e técnicas de narração, análise, engajamento e significado, processamento de dados, estrutura de dados, detecção de eventos, e visualização narrativa em pesquisa. Todos os artigos trataram a importância da interação neste campo.

Levando em consideração o nosso objetivo maior, que permeou todos os estudos, de compreender o trabalho dos analistas de dados de mídias sociais e suas necessidades em relação à visualização, percebemos que é necessário promover uma maior integração entre as ciências sociais e humanas com a ciência da computação, para que as soluções computacionais se adaptem melhor aos objetivos das análises que irão apoiar.

Posterior ao mapeamento, fizemos uma análise mais aprofundada das estratégias de *design* de visualizações narrativas que gerou um artigo sobre esse assunto. A principal contribuição deste artigo, segundo Ghidini et al. [GSMS17], foi a extensão das propostas originais do espaço de *design*, identificando as principais estratégias adotadas na academia e oportunidades de pesquisa, no que se refere ao *design* e à interação do espectro de visualizações narrativas.

3.4 Estudo 4 - Entrevistas com Analistas de Dados de Mídias Sociais

O estudo 4 consistiu em entrevistas com profissionais habituados a fazer análise de dados extraídos de mídias sociais. Muitas pessoas, de diferentes áreas do conhecimento, estão extraindo e analisando dados das mídias sociais com diversos objetivos, como observar o comportamento ou entender a opinião da população online, por exemplo. No entanto, para isso, essas pessoas (que aqui chamaremos de analistas de dados) precisam usar ferramentas de coleta e apresentação de dados para conseguirem analisar grande volume de dados e, para isso, precisam aprender uma série de novas habilidades. A proposta deste estudo foi conhecer o trabalho envolvido neste tipo de análise e entender melhor este domínio, compreendendo as atividades e necessidades dos analistas.

Com este objetivo, formulamos nossas questões de pesquisa para este estudo de forma alinhada com o trabalho de Chin Jr. et al. [CJKW09], focando nos aspectos sobre como os analistas organizam os dados para a análise e para a apresentação. Também alinhamos com pesquisas concentradas nas práticas de análise de dados de mídias sociais [DNKS10, DDCN12, Bro15, Zha17]. Neste contexto, nossa principal questão de pesquisa é “*como os analistas lidam com a visualização de dados no processo de análise de dados de mídias sociais?*” Além disso, pretendemos responder as seguintes questões:

RQ1: Como os analistas de dados de mídias sociais entendem as visualizações de dados?

RQ2: Quais as necessidades dos analistas de dados de mídias sociais em relação à visualização de dados?

3.4.1 Metodologia

Neste estudo, utilizamos a entrevista para tentar obter a percepção de analistas de dados sobre visualização de dados e o quanto ela pode contribuir (ou não) no processo de análise. O roteiro que serviu de base para a entrevista semiestruturada realizada neste estudo está no Apêndice C. Alguns resultados preliminares já foram discutidos [SSM16] e são reforçados nos resultados aqui apresentados. Foram entrevistados oito participantes, cujos perfis são descritos na tabela 3.5. Os participantes têm formação acadêmica em diferentes áreas e todos trabalham com dados de mídias sociais, mas apenas três deles não usam ferramentas de visualização para suas pesquisas e análises, conforme mostrado na tabela.

Tabela 3.5 – Perfil dos participantes do estudo 4.

Nº	Idade	Formação	Usa ferramentas de visualização?	Trabalha com dados de mídias sociais?	Possui conta no Twitter?	Possui conta no Facebook?
P1	58	História	Não	Sim	Sim	Sim
P2	42	História	Não	Sim	Sim	Sim
P3	28	Jornalismo	Não	Sim	Sim	Sim
P4	37	Jornalismo	Sim	Sim	Sim	Sim
P5	32	Publicidade	Sim	Sim	Sim	Sim
P6	25	Análise de Sistemas	Sim	Sim	Sim	Sim
P7	28	Administração	Sim	Sim	Sim	Sim
P8	29	Jornalismo	Sim	Sim	Sim	Sim

As entrevistas foram realizadas com uma amostra de conveniência, por meio de contatos pessoais com profissionais que trabalham com a análise de dados de mídias sociais. Elas variaram de 50 a 60 minutos cada, foram gravadas em áudio e depois transcritas e analisadas usando o método de análise de conteúdo [LFH10]. Utilizamos o RQDA¹¹ para categorizar os dados e extrair as informações que serão discutidas aqui.

Neste estudo, foi possível também ampliar nossa compreensão sobre os desafios neste campo de pesquisa. A maioria dos participantes usa ou usou alguma ferramenta de visualização,

¹¹RQDA é um pacote do aplicativo R para análise de dados qualitativos.

e aqueles que não utilizaram justificaram que consideram as ferramentas complexas e difíceis de aprender. As ferramentas de visualização mencionadas nas entrevistas requerem alguns esforços de aprendizagem dos analistas em termos de utilização e combinação para gerar visualizações que atendam plenamente às suas necessidades.

A partir deste estudo, descobrimos que os pesquisadores da Visualização de Dados e da Interação Humano-Computador precisam aumentar os esforços para compreender o processo de análise de dados antes de projetar novas visualizações e interações. Discutiremos os resultados a seguir.

3.4.2 Resultados

Em relação ao entendimento de visualização de dados pelos analistas de dados de mídias sociais (**RQ1**), perguntamos aos participantes sobre suas habilidades em ferramentas de visualização e suas opiniões sobre o uso de visualização na análise de dados. Observamos que os participantes conhecem o conceito de visualização de dados e acreditam que seu uso na análise pode melhorar significativamente o processo de análise. Acreditamos que existe uma lacuna nesse aspecto, uma vez que, mesmo que eles conheçam o conceito e enxerguem os benefícios do seu uso, eles não apontaram a necessidade e o esforço em aprender a usar as tecnologias de visualização, consideradas por eles mais difíceis, mas que resolveriam algumas das suas necessidades. Ainda assim, todos relataram que gostariam de usar visualizações, mas alguns consideram as ferramentas disponíveis complexas e difíceis de aprender. Isso pode ser ilustrado pelos discursos dos participantes a seguir:

“Hoje, para além de todo esse esforço, visualizar aumenta mais ainda a possibilidade da gente de discussões e de análises, muito mais que antes. Traz muito mais dados e informações. A informação é o começo pra gente. Hoje algumas visualizações já avançam da informação e chegam na ideia da comunicação. Por isso é fundamental hoje em dia.” (P1).

“Eu usei algumas ferramentas de visualização, mas nunca usei o Gephi, por exemplo, porque acho muito difícil. Eu não consegui instalar o Gephi, para você ter ideia.” (P3).

“as visualizações não são meras ilustrações. Elas efetivamente tornam mais fácil ver algo que eu não conseguiria apenas explicar verbalmente. Elas revelam algumas coisas que eu não conseguiria dizer, interpretar sem ver.” (P4).

“Eu considero a visualização essencial na análise de dados. Na verdade, na análise de dados em si, não, porque aí o importante é a inteligência analítica que tu tá colocando em cima disso. (...) A visualização é importante no processo de engajamento e compreensão dos dados. Ela aumenta o entendimento daquilo.” (P7).

“A visualização de dados é um pouco a interface entre o trabalho que o jornalista faz e a informação que chega pro usuário, pro leitor, internauta. (...) O analista que não se beneficia das ferramentas de análise de dados está um pouco atrasado.” (P8).

De uma forma geral, os participantes reconhecem a importância e a necessidade de aprender a manipular dados e gerar visualizações interessantes sobre eles. Percebemos que, embora alguns analistas não utilizem ferramentas computacionais (3 participantes), eles sabem quais dados eles querem coletar e analisar. Assim, as entrevistas permitiram que entendêssemos como as pessoas estão realizando análises de dados de mídias sociais e o que elas consideram importante nesse contexto, principalmente no que se refere à visualização de dados.

Como falamos, todos os participantes trabalham com dados extraídos de mídias sociais, sendo que 6 deles se consideram analistas de dados, P5 não se considera e P8 se considera jornalista de dados, pois, segundo ele, *“eu tenho muito trabalho de análise e não sei se é exatamente o mesmo tipo de análise que um analista de dados faz”*.

Todos os participantes que se consideram analistas de dados não se sentiram completamente seguros para reconhecer esse papel, visto que eles acreditam ser uma denominação que envolve muitas habilidades, como pode ser resumido na fala de P7: *“para ser um analista de dados é preciso ter raciocínio lógico, estatístico, matemático, inteligência analítica, que é o principal, aí as outras habilidades vão depender muito do âmbito em que você trabalha”*. Por outro lado, como eles fazem análise de dados de mídias sociais, mesmo não possuindo todas as habilidades que imaginam ser necessárias ao analista, eles se intitularam analistas de dados, pela capacidade de analisarem grandes bases de dados.

Essa insegurança sobre desempenharem ou não o que os participantes entendem sobre as atividades típicas de um analista parece estar relacionada à falta de uma definição mais formal de o que é um analista de dados de mídias sociais e também de quais processos e atividades ele deve executar. O participante P4 acredita que a análise feita com base em alguma teoria já consolidada, como a Análise de Redes Sociais (ANS), fortalece o papel de analista de dados. Segundo P4, o que ele faz *“passa pela análise de dados, embora isso seja feito com um olhar bastante peculiar, que é um olhar interdisciplinar, um olhar pouco filiado a perspectivas conceituais mais estabelecidas, um olhar para discussões que não exatamente estão dentro de uma perspectiva consolidada dessas análises de dados.”*

Já o P2, que é um participante que não utiliza ferramentas computacionais para coleta e análise de dados, compreende a sua atividade de analista como algo feito em profundidade, pois *“costuma fazer apanhados, resumos, escrever coisas a partir do que vê nas mídias sociais, não nesse sentido quantitativo, mas na direção de tentar mapear um problema durante um determinado período de tempo, voltando sempre aos mesmos perfis e aos mesmos grupos etc. Assim, eu acho que isso configura uma ideia de análise de dados”*.

Do ponto de vista da Ciência da Computação, especificamente da Interação Humano-Computador, os resultados sugerem que os analistas de dados conhecem ou já ouviram falar de

algumas ferramentas de visualização, mas ainda não as dominam. No entanto, eles acreditam que a colaboração é uma boa maneira de resolver esse problema. Isso pode ser ilustrado, por exemplo, pela fala de P1: *“eu não sei como usar essas ferramentas, mas eu as conheço. Assisti a uma oficina sobre o Gephi e achei muito difícil. Eu uso visualizações no meu trabalho, feitas por meus alunos e por meus parceiros de trabalho, de acordo com minhas demandas. Também procuro visualizações em mídia ou com colaboradores. A visualização é muito importante porque permite compreender a dinâmica, ver que os dados estão em formas diferentes, ao mesmo tempo, em vários lugares. Isso mostra a complexidade das coisas, ao contrário de uma tabela.”*

Em geral, os participantes entendem que, não somente eles, mas os usuários da Internet, a comunidade digital, precisam entender essa realidade relacionada ao volume de dados que é compartilhado nas mídias sociais e o quanto de informação pode ser extraída desses dados. Segundo P3, *“as pessoas precisam desenvolver a cultura “dateira”, precisam entender que os dados trazem muita informação sobre as pessoas e o mundo. No ano 2000 o custo de armazenamento de dados era alto e a capacidade baixa, e hoje é o contrário!”*.

Com isso, um ponto importante identificado a partir das entrevistas foi a necessidade de se incentivar e promover a literacia digital das pessoas que querem fazer análise de dados de mídias sociais. Muitas pessoas não possuem habilidades de busca de informação, de manipulação de planilhas eletrônicas. Isso pode ser percebido nas falas de vários participantes como:

“Era um curso de jornalismo de dados, as pessoas que estavam participando eram, em sua maioria, jornalistas que não têm habilidade com softwares e internet, e todos acharam muito denso o conteúdo. Tipo: uma busca avançada no Google era muito difícil para as pessoas” (P3).

“A gente não domina a ferramenta, então efetivamente pode haver várias coisas lá que a gente não está conseguindo fazer ou não está explorando tanto, e que permitiriam estar mais próximo do nosso ideal” (P4).

“o Gephi é complicado, o NodeXL também é, porque em ambos tu tem que entender muito dos conceitos básicos de rede social, desde o nodo, o que ele é, como esses nós se relacionam, o que é um relacionamento entre aqueles pontos... então, assim, eu acho que hoje é impossível tu conseguir analisar dados sem uma ferramenta” (P6).

“A gente deu um curso na Folha de São Paulo de Análise de Dados com Google Sheets, que é uma ferramenta de planilha eletrônica. E, se você vai olhar, assim, numa redação, numa grande capital, numa das maiores redações do país, você vai olhar lá dentro... cara! quem é que sabe mexer em Excel? Meia dúzia! Então, pra mim, é um problema de literacia digital, sabe.” (P8).

É interessante perceber o reconhecimento relacionado a necessidade de literacia digital, principalmente nos dias atuais em que se tem disponível uma grande quantidade de ferramentas computacionais para auxiliar os analistas de dados na coleta, limpeza, análise e apresenta-

ção de dados. Como exemplos dessas ferramentas podemos citar o CartoDB¹², Gephi¹³, Google Analytics¹⁴, Google Sheets¹⁵, Infogr.am¹⁶, NodeXL¹⁷, OpenRefine¹⁸, Quadrigam¹⁹, Tableau²⁰, Web Scraper²¹, Timeline.js²², Story Maps²³ etc. Todas elas foram citadas pelos participantes, alguns mencionaram conhecer o propósito da ferramenta, mas não saber utilizá-la, outros já utilizaram algumas delas. Com isso, fica claro que os participantes sabem da existência de ferramentas para apoio às suas análises.

Sobre a **RQ2**, que busca levantar quais as necessidades dos analistas de dados de mídias sociais em relação à visualização de dados, conseguimos identificar algumas necessidades de forma bem direta. O participante P4, por exemplo, relatou que precisa combinar o uso de várias ferramentas para tentar obter o resultado desejado para suas análises, porém ainda não conseguiu representar totalmente as suas necessidades. Ele utiliza bastante o Gephi e tem feito alguns trabalhos usando o Tableau. Segundo P4, “o Gephi é a situação mais emblemática em que precisei de visualizações que atendam a minha necessidade e não tive”. Ele deseja criar visualizações dinâmicas que mostrem as redes se formando e se transformando ao longo do tempo. No entanto, ainda não encontrou ou não descobriu como as ferramentas existentes podem representar esse tipo de visualização:

“Por exemplo, no caso do 7x1, que é um estudo que eu fiz, eu quero ver exatamente as redes de tweets que aconteceram no intervalo do jogo. Eu quero ver o que as pessoas falaram no intervalo do jogo, entre o final do primeiro tempo, em que todo mundo, de alguma maneira, disse “uffa! acabou essa tortura durante alguns minutos” até o momento de recomeçar. Pra eu fazer isso com a linha do tempo do Gephi, hoje, ou pelo menos como a gente tem feito, “poutz!”Haja gambiarra!” (P4).

P4 relatou que, para tentar representar essa sua necessidade, ele utiliza o recurso de *timeline* do Gephi. Pelo *timestamp* ele ativa a linha de tempo do Gephi e pode ver a rede se formando. O Gephi cria embaixo da visualização uma *timeline* com um play. Ao apertar o play, ele vai montando e remontando a rede, como se você estivesse acontecendo ao vivo.

Esse foi um ponto importante que apareceu nas entrevistas: a necessidade de observar a mudança de contexto, relacionamentos e foco ao longo do tempo, em mídias sociais. Isso pode ser ilustrado também pelo discurso de outros participantes:

¹²<https://cartodb.com/>

¹³<https://gephi.org/>

¹⁴<https://www.google.com/analytics/>

¹⁵<https://www.google.com/intl/nl/sheets/about/>

¹⁶<https://infogr.am/>

¹⁷<https://nodexl.codeplex.com/>

¹⁸<http://openrefine.org/>

¹⁹<http://www.quadrigam.com/>

²⁰<http://www.tableausoftware.com/>

²¹<http://webscraper.io/>

²²<https://timeline.knightlab.com/>

²³<https://storymaps.arcgis.com/en/>

“Eu adoraria poder ver as controvérsias nas mídias sociais, especialmente ao longo do tempo. Se eu pudesse ver, retroagir em intensidade como as questões foram aparecendo ao longo do tempo e onde. Por exemplo: eu preciso distinguir aborto, enquanto aborto opção da mulher, de aborto política pública. Quando é o primeiro tem uma carga de moralismo muito grande, e vai ser respondido ali” (P1).

“No meu negócio, para eu melhorá-lo, eu preciso analisar historicamente os picos, por exemplo, e outras informações complementares (dos sites, por exemplo)” (P5).

“Eu acho que um recurso de versão resolveria a questão do tempo, que aí você teria a versão 1, foi a que você fez a coleta do dia x ao dia y, e naquele dia ali o movimento estava só no início; aí você vê aquela versão e depois você tem a versão 2, que você continua a coleta até um mês depois, por exemplo, e aí você tem a versão 3, que aquilo já está há 6 meses acontecendo, e aí você vê a mudança, então, esse controle de versão facilitaria” (P8).

Alguns participantes apontaram a importância das visualizações e da infografia na análise de dados, mencionando notícias da revista Time que resumiram tudo sobre a crise de emissões da Volkswagen em 2014 em um infográfico²⁴, em uma tentativa de relacionar isso ao que uma visualização de dados precisa conseguir informar aos seus leitores:

“Vi uma visualização sobre a crise da Volkswagen. Foi um resumo, mas também ajudou as pessoas a entender e entender, e não simplesmente se assustar com a manchete das notícias: “O presidente da Volkswagen renuncia por causa de (...)”. Então, a visualização de dados nos permite entender a complexidade de certas coisas”. (P1)

“Tem que ser algo que combina informações principais e que traz algo que você não lembre apenas se você ler o texto. Por exemplo, não lembro de ler quais modelos de carros estavam envolvidos no caso da Volkswagen, mas depois de ver o infográfico, eu mantive isso em mente.” (P2)

E outros participantes mencionaram diretamente sobre associar narrativas com visualizações:

“A forma que a história possibilita pra gente... ela chama a atenção para pensarmos que uma imagem tem toda uma série de contextos, formas de análise, texto, números e seus cruzamentos... e isso nenhum software possibilita.” (P1)

“Agregar narrativas nas visualizações tem uma grande importância porque as pessoas (público-alvo) querem receber a informação mastigada mas, se eu só jogar gráficos com números, não vai ter graça nenhuma. Então o contexto todo cria o ambiente e o entendimento da pessoa pra ela entender aquilo. É fundamental.” (P3).

²⁴<http://migre.me/uXWkk>

“Hoje em dia, não necessariamente, os infográficos precisam ser estáticos, eles podem ter níveis de interação ali dentro, mas o que eu acho que não perdeu, assim, na essência do infográfico é que existe essa ideia de você criar uma narrativa e guiar o usuário dentro dessa narrativa, pra entender um determinado assunto. No caso da visualização, você já tem uma narrativa um pouco menos sólida, sabe, você já deixa mais a cargo do leitor criar a própria narrativa dentro ali das interações que existem na visualização.” (P8)

Eles também estão preocupados com a customização. Alguns participantes relataram que muitas vezes eles precisavam usar *plug-ins*, baixar bibliotecas, importar pacotes ou escrever *scripts* para organizar as visualizações como eles queriam, e se isso fosse oferecido na ferramenta, seria muito bom: *“O Tableau, por exemplo, é uma ferramenta incrível. Mas, poderia nos dar a possibilidade de customizar algumas coisas com mais facilidade. Por exemplo, para que você criar sua própria paleta de cores, você precisa importar e criar um esquema inteiro complexo.” (P8)*

Assim, a partir das respostas dos participantes, resumimos o que foi mais comum em relação às necessidades apontadas por eles em relação a ferramentas de visualização para apoio à análise de dados de mídias sociais. Os analistas querem visualizações:

- associadas a dados de outros meios como vídeo, por exemplo;
- que mostrem ou permitam adicionar marcação de pontos ou momentos importantes;
- que expliquem os termos usados para coletar os dados;
- com narrativas;
- que mostre controvérsias nas mídias sociais;
- que mostre as mudanças dos dados ao longo do tempo.

Em relação à nossa principal questão de pesquisa: *“como é a relação entre analistas de dados e visualização de dados no processo de análise de dados?”*, nossos resultados sugeriram que a relação entre analista de dados e visualização está em evolução. Eles conhecem o conceito de visualização de dados e acreditam na importância do seu uso, mas alguns analistas ainda consideram as ferramentas complexas e difíceis de aprender. No entanto, quando eles precisam, pedem um parceiro para criar as visualizações.

3.4.3 Conclusões

Pretendíamos, com as entrevistas, nos aproximar do universo da análise de dados de mídias sociais e compreender como os analistas realizam todo o processo. Percebemos que há dificuldades em representar graficamente os dados da forma que desejam, por exemplo, a questão apontada sobre mudanças da rede ao longo do tempo. Na prática, os analistas fazem uso de várias ferramentas para gerar as visualizações que são apresentadas, geralmente, para complementar a

informação passada em forma de texto. As ferramentas mais utilizadas são o Gephi e o Tableau, mas, no entanto, a maioria acha o Gephi muito complexo e de difícil aprendizado.

Os analistas não planejam antecipadamente o que irão criar, em termos de visualizações. Na prática, eles coletam os dados e geram visualizações de vários tipos, até encontrarem aquela que melhor represente a informação que ele deseja apresentar. Com isso, percebemos que não há uma reflexão antecipada sobre os dados para decidir o que e como eles serão apresentados. No entanto, os analistas sentem falta de ferramentas de fácil uso e que permitam a eles construir as visualizações em um único ambiente, sem a necessidade de se utilizar várias ferramentas para isso.

A partir desse estudo, percebemos que existem diferentes perfis de analistas de dados de mídias sociais, principalmente no que se refere ao domínio de tecnologias. Assim, para apoiar a nossa reflexão sobre como ajudar a resolver os problemas dos analistas, fizemos uma triangulação dos estudos, em que explicitamos os requisitos identificados para uma ferramenta de visualização de dados, a partir das entrevistas, e apresentamos uma classificação dos perfis em *personas*, como será descrito a seguir.

3.5 Triangulação dos estudos e criação de personas

A partir dos dados analisados nos grupos focais, na *survey online* e nas entrevistas, conseguimos, a partir da triangulação desses resultados, construir a nossa compreensão de o quanto a visualização de dados está presente nas atividades de trabalho dos participantes e conhecer a opinião deles sobre a importância do uso de visualizações para apresentação de seus dados. Para os jornalistas, pudemos perceber que a combinação de técnicas de visualização podem ser úteis para a obtenção e a divulgação de informações em mídias sociais, podendo ser um recurso valioso em suas atividades de trabalho. Como estão familiarizados com o conceito de infografia, notamos que os jornalistas acreditam que a visualização pode ser utilizada como uma espécie de infográfico interativo, mas é algo que ainda precisa ser melhor introduzido nesta área.

Pudemos observar que, para os analistas de dados, a visualização é um conceito não apenas conhecido mas também necessário no processo de análise de dados de mídias sociais. No entanto, para eles, e principalmente para os leigos em computação, há ainda dificuldades relatadas em aprender a utilizar as ferramentas de visualização e há também demandas que eles acreditam que não sejam atendidas pelas ferramentas existentes.

Com isso, nesta seção, iremos apresentar: *i*) as necessidades/demandas dos analistas de dados para ferramentas de visualização, levantadas a partir das análises dos estudos realizados e sua triangulação; e *ii*) as **personas** criadas para representar os diferentes perfis de participantes dos estudos. Esses pontos foram fundamentais nas nossas decisões para criação do modelo CIViS, que será apresentado no próximo capítulo.

Tratando primeiramente das demandas/necessidades dos analistas de dados de mídias sociais para ferramentas de visualização de dados, nossas descobertas sugerem três pontos principais a serem considerados em uma ferramenta de visualização: (1) contextualização, (2) customização e (3) observação ao longo do tempo, conforme mostrado na figura 3.15. Por sua vez, esses pontos contêm as necessidades levantadas a partir da triangulação dos nossos estudos e que se tornaram importantes na reflexão para a criação do CIViS. Cada necessidade foi numerada, na figura, para serem relacionadas com as personas que apresentaremos ainda nesta seção 3.5.



Figura 3.15 – Necessidades dos analistas em relação a ferramentas de Visualização de Dados.

Tendo em vista estas necessidades e as informações obtidas a partir da identificação do perfil dos participantes em nossos estudos, criamos 3 *personas* dedicadas à representação de grupos de usuários segundo seu padrão de comportamento nas atividades de análise, uso de tecnologias, preferências, etc. Do ponto de vista formal, uma *persona* é a descrição de um usuário fictício, que não existe como uma pessoa específica, mas pode ser descrito de forma que o leitor consiga reconhecer a descrição e acredite que o usuário possa existir na realidade [Nie12]. Assim, a *persona* é descrita com base em informações relevantes de usuários potenciais e reais. Segundo Baxter et al. [BCC15], a *persona* deve ser fictícia, mas seus atributos devem ser de usuários reais, com detalhes para manter a autenticidade.

Uma vez tendo desenvolvido um perfil completo do usuário, podemos desenvolver *personas* (exemplares de seu usuário final) e cenários (um dia na vida do seu usuário final) [BCC15]. Segundo os autores, *personas* são projetadas para ajudar a manter o foco nos tipos de usuários

específicos durante as discussões de *design*. E cenários ajudam a testar o sistema e criar funcionalidades no produto que os usuários realmente desejam utilizar.

Neste trabalho, as *personas* são referentes a três grupos de perfis de usuários analistas de dados que irão atuar como *co-designers* de visualizações narrativas interativas. As *personas* serão utilizadas para análise e reflexão sobre como será a customização realizada por elas, como será explicado no capítulo 5. Supomos que, embora os analistas de dados de mídias sociais tenham perfis diferentes e sejam representados pelas diferentes *personas* criadas aqui, todos possuem, praticamente, as mesmas necessidades. Para criação das *personas*, nos baseamos nos elementos sugeridos por Baxter et al. [BCC15], que consistem de uma lista idealizada, mas que não necessariamente deve conter todos os itens:

- Identidade: dê à *persona* um nome e sobrenome. Forneça uma idade e outros dados demográficos que sejam representativos do perfil do usuário.
- Status: diz respeito a definir se a *persona* representa um perfil de usuário primário, secundário, terciário ou anti-usuário²⁵ do seu sistema.
- Objetivos: quais são os objetivos dessa *persona*, particularmente aqueles relacionados ao seu produto específico ou produtos concorrentes?
- Conjunto de habilidades: qual é o histórico e experiência de sua *persona*? Isso inclui educação, treinamento e habilidades específicas. Novamente, não se limite aos detalhes relacionados ao seu produto específico.
- Tarefas: quais são as tarefas básicas ou críticas que a *persona* conduz? Qual é a frequência, importância e duração dessas tarefas? Informações detalhadas sobre a tarefa devem ser incluídas nos cenários.
- Relacionamentos: compreender com quem a *persona* se relaciona é importante. A inclusão de relacionamentos na *persona* mantém você pensando em *stakeholders* secundários e terciários.
- Requisitos: o que sua *persona* precisa para usar seu produto ou ser bem-sucedido usando-o (por exemplo, alta velocidade de conexão à Internet, sistema operacional específico de telefone celular, treinamento específico ou formação)? Incluir citações ajuda a direcionar melhor essas necessidades.
- Expectativas: como a *persona* pensa que o produto funciona? Como ela organiza a informação em seu domínio de trabalho?
- Fotografia: inclua uma foto em sua *persona* para dar um rosto ao seu usuário final.

As *personas* aqui criadas foram derivadas de um processo de investigação que levantou as características dos potenciais usuários do modelo proposto (descrito no Capítulo 4). Esses

²⁵anti-usuário é alguém que não vai utilizar o produto e, por isso, não influencia nas decisões de *design*

usuários foram agrupados em perfis semelhantes que se adequavam a cada uma das *personas*. Apenas seus nomes e detalhes pessoais são inventados.

A *persona* 01, apresentada na tabela 3.6, foi pensada para representar os usuários (analistas de dados) que têm pouca ou nenhuma habilidade com ferramentas computacionais para coleta e apresentação de grande volume de dados. As demandas da *persona* 01 são representadas pelos itens 1, 3, 5, 6 e 7 da figura 3.15.

Tabela 3.6 – Persona 01

Foto	Descrição
	<p>Nome: João Chaves</p> <p>Idade: 50</p> <p>João tem formação nas áreas de humanas e de ciências sociais. Trabalha em universidade e suas pesquisas envolvem analisar fatos, acontecimentos, situações, contextos históricos, sociais e econômicos a partir de conteúdos de mídias sociais online, assim como observar e analisar as mudanças (e o que impactam essas mudanças) ao longo do tempo nesses aspectos. João não utiliza software para coletar os dados e apoiar a análise, mas ele procura parceiros para fazer isso. Ele gostaria de poder apresentar suas demandas ao parceiro que irá criar/executar as ferramentas de software necessárias para que este lhe forneça os resultados esperados.</p>

A *persona* 02, apresentada na tabela 3.7, foi pensada para representar os usuários (analistas de dados) com alguma habilidade com ferramentas computacionais para coleta e apresentação de grande volume de dados. Representa usuários que conhecem algumas ferramentas, mas não aprenderam a utilizá-las ou tiveram dificuldades. As demandas da *persona* 02 são referentes aos itens 2, 3 e 4 da figura 3.15.

Tabela 3.7 – Persona 02

Foto	Descrição
	<p>Nome: Maria Santos</p> <p>Idade: 28</p> <p>Maria tem formação na área de comunicação social. Trabalha em empresa e suas atividades envolvem promover o marketing digital da empresa, analisar as métricas das redes e mídias sociais online do perfil da empresa, criar relatórios a partir dessas análises, observar o comportamento dos usuários nas redes sociais, planejar estratégias de melhoria das métricas. Maria não utiliza software para coletar os dados e, para apoiar a análise dos dados, ela utiliza os recursos que os próprios sites oferecem e também o Google Analytics. Ela gostaria de ter tempo para aprender a usar ferramentas de visualização de dados como o Tableau e o Gephi, porque acredita que poderiam apoiar melhor a criação de gráficos sobre os dados que darão suporte às decisões a serem tomadas pela empresa.</p>

A *persona* 03, apresentada na tabela 3.8, foi pensada para representar os usuários (analistas de dados) com habilidades em ferramentas computacionais para coleta e apresentação de grande volume de dados. Representa usuários que conhecem algumas ferramentas, utilizam-nas e buscam o aprendizado de novas ferramentas. As demandas da *persona* 03 se referem aos itens 2, 3, 4, 5 e 7 da figura 3.15.

Tabela 3.8 – Persona 03

Foto	Descrição
	<p>Nome: José Pereira</p> <p>Idade: 31</p> <p>José tem formação na área de comunicação social/ jornalismo. Trabalha com pesquisas que envolvem investigar e analisar fatos, acontecimentos, situações, contextos históricos, sociais e econômicos a partir de conteúdos de mídias e redes sociais online, assim como observar e analisar as mudanças (e o que impactam essas mudanças) ao longo do tempo nesses aspectos. José utiliza software para coletar e fazer limpeza dos dados e para apoiar a análise, e também conhece e utiliza ferramentas de visualização de dados. Ele gostaria que essas ferramentas tivessem recursos para apresentar a mudança de contexto ao longo do tempo mas, que não fosse apenas uma espécie de “fotografia” de cada momento.</p>

Já a *persona* 04, apresentada na tabela 3.9, foi pensada para representar o *designer* participante do processo de comunicação que serviu de base para o modelo proposto nesta Tese. Ele é da área de Ciência da Computação e conhece algumas ferramentas de coleta, tratamento e apresentação de grande volume de dados. Seu objetivo é ajudar as demais *personas* nas suas necessidades.

Tabela 3.9 – Persona 04

Foto	Descrição
	<p>Nome: Carol Martins</p> <p>Idade: 39</p> <p>Carol tem formação na área de ciência da computação. Ela é <i>designer</i> de interação e tem interesse em ajudar os analistas de dados de mídias sociais a realizarem suas análises com o apoio de ferramentas computacionais. No entanto, ela precisa compreender o que um analista de dados de mídias sociais faz e também entender como funciona o processo de análise. Além disso, quer entender a importância da visualização de dados no trabalho desses analistas. Ela pretende ajudar os analistas a planejarem as visualizações de dados que gostariam de criar. Carol acredita que os analistas, ao conhecerem uma forma de planejar e organizar a apresentação dos seus dados, ampliarão o seu entendimento sobre as possibilidades de criar visualizações desses dados.</p>

Como dito anteriormente, além das *personas*, é interessante criarmos cenários, que são, muitas vezes, referidos como histórias sobre as *personas* criadas. Um bom cenário começa com

uma *persona* e, em seguida, acrescenta mais detalhes com base nos requisitos de seu usuário. Ele descreve como uma *persona* completa uma tarefa ou se comporta em uma determinada situação. Também fornece uma configuração que possui atores, objetivos, uma sequência de eventos e fecha com um resultado [BCC15]. Os cenários, neste trabalho, serão apresentados no decorrer dos próximos capítulos.

4. CIVIS - MODELO DE *DESIGN* DE INTERFACE CUSTOMIZÁVEL PARA APOIAR A CONSTRUÇÃO DE VISUALIZAÇÕES NARRATIVAS DE DADOS EXTRAÍDOS DE MÍDIAS SOCIAIS

Neste capítulo será apresentado o modelo **CIVIS** (*Customizable Interface model to support building narrative Visualizations of data extracted from Social media*), uma estrutura descritiva criada para apoiar *designers* atuando em níveis distintos de responsabilidade na produção de visualizações narrativas. Em um primeiro nível, o modelo pretende guiar o *designer* de sistemas customizáveis de visualização na reflexão sobre as necessidades do usuário imediato desses sistemas, o analista de dados de mídias sociais. Este, por sua vez, age como *co-designer* ao fazer uso do sistema e seus recursos de customização para criar visualizações narrativas de dados extraídos de mídias sociais, configurando assim o segundo nível de *design* que o modelo pretende apoiar. Note que, ao apresentar o analista de dados como *co-designer*, deixa-se explícita a visão de que este age, por meio de customização, como projetista e desenvolvedor de um artefato de software.

Desta forma, o modelo CIVIS, tendo sido fundamentado na teoria da Engenharia Semiótica e sua visão sobre EUD, busca guiar *designers* e *co-designers* no planejamento do *design* de suas visualizações. Não é objetivo do modelo fornecer uma solução definitiva para visualização narrativa de dados, mas sim proporcionar um maior entendimento sobre o problema em questão (a criação de representações gráficas para um conjunto de dados) e sobre as suas implicações, bem como a reflexão sobre as possíveis soluções a serem criadas. A inspiração para criação do modelo surgiu dos estudos iniciais realizados e de situações (cenários) que serão usadas como exemplos ao longo do capítulo. O cenário 1, a seguir, apresenta o contexto que originou este trabalho.

CENÁRIO 1

Carol é *designer* de interação e participa de um grupo de pesquisa sobre novas mídias, comportamentos e fenômenos observados nas mídias sociais. Os participantes do grupo são profissionais de diversas áreas: comunicação social, história, jornalismo e sociologia. Apenas Carol é da área de ciência da computação. O grupo já fez muitos trabalhos de análise de dados extraídos de mídias sociais e, com frequência, encontrou dificuldades em estabelecer parcerias com pesquisadores da área de tecnologia para criar soluções que atendessem completamente os seus interesses em relação à apresentação dos dados. Diante deste contexto, Carol se interessou em saber o que os participantes do grupo desejavam, em termos de soluções tecnológicas, para auxiliar os seus trabalhos. Inicialmente, a questão que guiou esta investigação foi:

- Como (de que forma) você gostaria de apresentar os seus dados?

Considerados os níveis de *design* que pretendemos apoiar, a partir deste cenário e dos demais que serão apresentados ao longo do capítulo e, também, com base na teoria da Engenharia Semiótica, concebemos a existência de duas diferentes mensagens de metacomunicação: uma emitida pelo *designer* e outra pelo *co-designer*, como é representado, de uma forma geral, na figura 4.1.

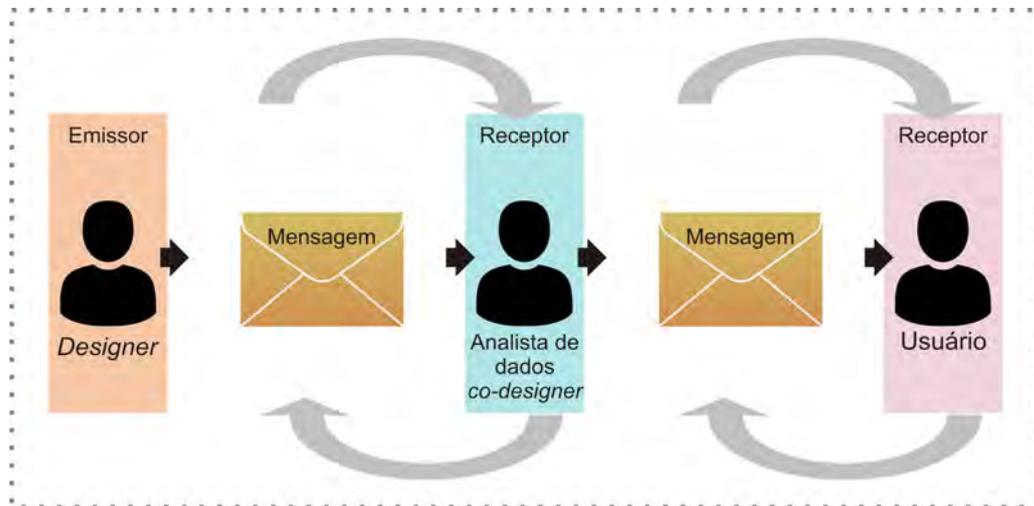


Figura 4.1 – Representação das mensagens de metacomunicação.

É parte deste trabalho definir as características de cada uma dessas mensagens, indicando sua estrutura fundamental. Isto dito, para a construção da primeira metamensagem, como parte da compreensão de quem é o receptor e quais as possíveis soluções para os seus problemas, foram realizados estudos e entrevistas com analistas de dados, que são os usuários foco desse processo de *design* em que atuarão como coautores. Após a reflexão sobre suas necessidades, foi pensado um discurso que atenda aos diversos perfis de usuário desse contexto. Assim, a mensagem de metacomunicação que nosso modelo recomenda que seja planejada pelo *designer* para ser transmitida via interface do sistema ao analista de dados *co-designer* pode ser parafraseada da seguinte forma:

“No meu entendimento, você é um usuário experiente em análise de dados extraídos de mídias sociais, que compreende a necessidade e o papel de aplicações computacionais para apoiar o seu processo de análise. No entanto, o seu perfil pode variar entre alguém com pouco conhecimento de ferramentas computacionais para apoiar análise de dados e alguém com muito conhecimento. Você precisa criar visualizações dos dados que coleta, com recursos de filtro, linha do tempo, zoom, anotações, ligações entre visualizações diferentes, etc. Você pode querer manipular a organização das visualizações, utilizando ordenação, realces e formas de interação. Ao mesmo tempo que você quer apresentar os seus dados com um “caminho” pré-definido para que o leitor entenda a informação, ou melhor, a história daqueles dados que você deseja apresentar, você também deseja que o leitor/usuário final interaja com as visualizações e tenha a liberdade de percorrer caminhos diferentes do que foi pensado por você e explore

os dados livremente. Você deseja ter esses recursos disponíveis para construir a sua narrativa de forma rápida, por meio da customização do ambiente, com o objetivo de melhor apresentar os seus dados e permitir uma maior exploração por parte do usuário final. Projetei, então, este sistema customizável para apoiar o designer de visualizações narrativas interativas (você), que apresenta um conjunto de recursos para criação e organização das visualizações. Eu sei que você quer criar as visualizações e apresentar a narrativa dos dados de acordo com a sua vontade. Assim, o modelo que projetei possui como plano de fundo um papel em branco para que você possa inserir e organizar os elementos que desejar. Toda a manipulação e organização de elementos são imediatamente visualizadas por você. Para criar suas visualizações, ofereço cartas contendo as opções básicas necessárias para essa criação, como escolha do gênero de narrativa, técnicas de visualização, especificação dos elementos da narrativa, os elementos estruturais e visuais das narrativas".

Já a segunda mensagem de metacomunicação, transmitida do analista de dados *co-designer* para o usuário final, foi pensada de acordo com as possibilidades de combinações de categorias do espaço de design utilizadas em modelos de visualizações narrativas que buscam equilibrar entre a narrativa pretendida pelo autor e a descoberta da história pelo leitor, como discutido na seção 2.1. Assim, a mensagem de metacomunicação a ser transmitida via visualização narrativa interativa ao usuário final pode ser parafraseada da seguinte forma:

"No meu entendimento, você é um usuário interessado em descobrir informações a partir de dados extraídos de mídias sociais por meio de representações visuais. Você precisa de visualizações que representem os dados em seu contexto, por meio de um caminho de leitura pré-definido, mas que também lhe dêem a liberdade de percorrer outros caminhos e explorar os dados, por meio de recursos de filtro, zoom, cliques, etc. Projetei, então, essa visualização narrativa interativa que traz a narrativa idealizada por mim para estes dados, como um percurso que você pode percorrer para compreendê-los ou criar outro percurso e explorar os dados da forma que você desejar".

Este capítulo está organizado em seções, para uma melhor apresentação do modelo e do seu uso. Primeiro, apresentamos a estrutura e a descrição do espaço de design de customização de visualizações narrativas considerada pelo modelo CIViS. Em seguida, na seção 4.2, apresentamos uma sugestão de como ele pode ser utilizado em decisões de *design* no contexto para o qual foi concebido. Finalmente, na seção 5.1, apresentamos um protótipo construído para realização de uma prova de conceito do modelo.

4.1 Estrutura do CIViS

Os componentes do modelo CIViS são as estruturas descritivas básicas que irão permitir ao *designer* representar o seu projeto de visualização narrativa customizável e ao *co-designer* representar o seu projeto de visualização narrativa interativa para um contexto específico. Como mencionado anteriormente, para estruturar a customização de visualizações narrativas, o CIViS baseia-se no espaço de *design* da Engenharia Semiótica e sua visão sobre EUD (visto na seção 2.2). Este espaço de *design*, por sua vez, baseia-se no modelo de comunicação de Jakobson [Jak60], que é composto pelos elementos: emissor, receptor, mensagem, código, canal e contexto (discutidos na seção 2.3). Assim como em sua teoria base, ao usar o CIViS, os *designers* devem ter consciência desses elementos e tomar decisões refletidas sobre cada um dos aspectos do espaço de *design*, contribuindo na concretização do processo de metacomunicação.

No âmbito deste trabalho, no primeiro dos dois níveis de *design*, **contexto** refere-se ao domínio da aplicação, cujo propósito está na customização de visualizações narrativas para apresentação de dados extraídos de mídias sociais. O **emissor** da mensagem de metacomunicação é o *designer* e o **receptor** é o analista de dados *co-designer*. O **canal** é o computador e a **mensagem** é a interface do sistema. O **código** são os componentes do sistema de visualização narrativa customizável, que serão descritos ainda nesta seção.

Inserido no contexto do primeiro nível está o **contexto** do segundo nível (figura 4.2), que se refere ao domínio da visualização narrativa, em que o usuário explora e interage com a visualização criada pelo analista de dados *co-designer*, que é **emissor** desta mensagem de metacomunicação. No segundo nível, o **canal** é o dispositivo de interação (computador, *tablet*, *smartphone*), a **mensagem** é a interface da visualização narrativa interativa, o **código** são os componentes da visualização narrativa interativa já customizada, e o **receptor** é o usuário final.

Na seção que se segue, vemos em detalhe os dois níveis de *design* compreendidos pelo modelo CIViS, explicando cada um de seus componentes de visualização narrativa e suas possíveis combinações.

4.1.1 Descrição do Espaço de *Design*

Vimos anteriormente que o espaço de *design* de sistemas de visualizações narrativas customizáveis é estruturado a partir de componentes, que em seu conjunto definem o sistema de codificação, ou código, desse domínio de aplicação. Nesta seção exploramos estes conceitos, componentes e atividades em maior detalhe.

Colocando em termos objetivos, o *designer* estrutura o espaço de comunicação definindo quem é o receptor *co-designer*, o que ele pode comunicar, em qual canal e usando quais códigos. Para apoiar essas decisões do *designer*, o CIViS traz os seguintes componentes de visualização

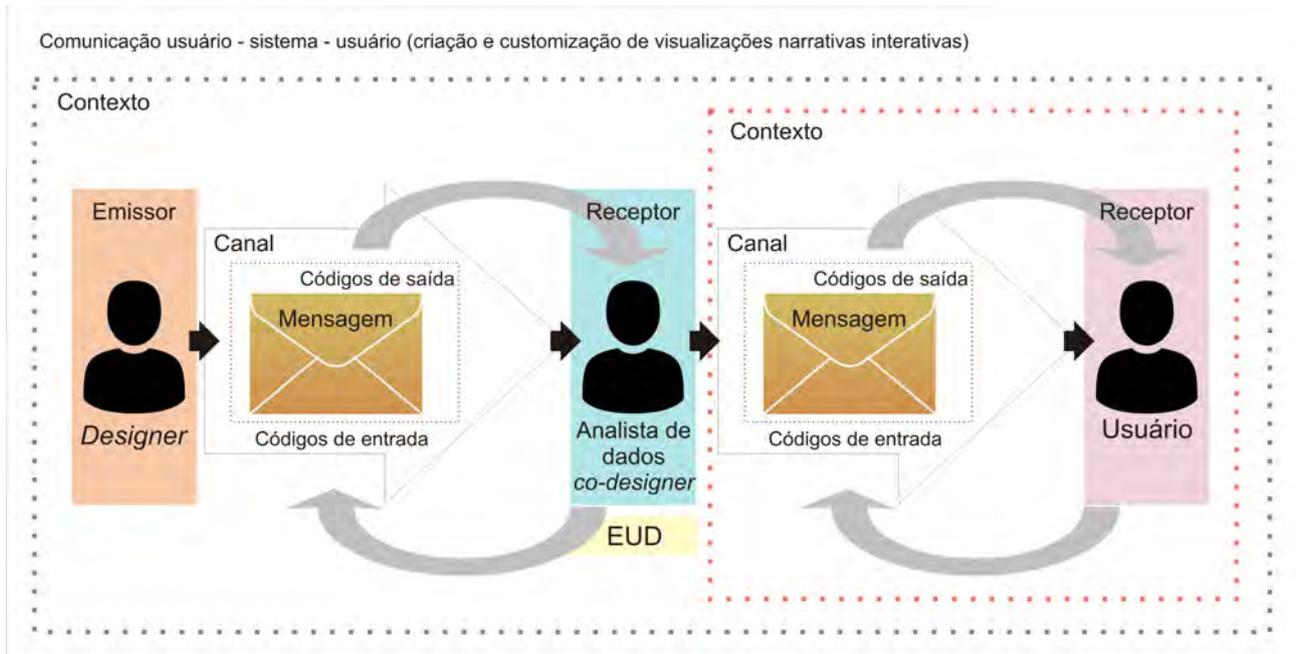


Figura 4.2 – Espaço de *Design* do CIViS.

narrativa: **Cena, Gênero, Modelo de visualização narrativa, Elementos de narrativa (evento, atores, localização, tempo), Técnicas de visualização, Elementos visuais, Elementos estruturais e Sistema de ajuda.** Tais componentes permitem ao *designer* do sistema apresentar as possibilidades de visualização narrativa ao analista de dados *co-designer* para que este compreenda quais aspectos estão presentes em seus dados e decida quais deseja utilizar na customização que irá realizar.

Assim, a atividade de customização da visualização narrativa acontece em função de um processo comunicativo mediado por computador, segundo o paradigma de computador como mídia. Nesta conversação entre *designer* e *co-designer* (motivada pela customização), a cena é a base inicial para organização dos elementos na interface. No primeiro nível, o analista de dados *co-designer* customiza a visualização narrativa para a sua audiência, o usuário final, levando em consideração o contexto de análise (que inclui os dados extraídos de mídias sociais) de um assunto ou tema específico. A estrutura deste nível é ilustrada na figura 4.3.

Em relação aos componentes de visualização narrativa previstos no CIViS, a **cena** é a estrutura base para a customização da visualização narrativa, pois consiste no espaço em que os elementos da interface serão organizados. Uma visualização narrativa pode ter uma ou mais cenas, de acordo com as decisões do analista de dados *co-designer*.

No CIViS, o **gênero** é fundamentado nos gêneros apresentados por Segel e Heer [SH10], discutidos na seção 2.1, permitindo ao analista de dados decidir sobre qual gênero utilizar e quando. São eles: *magazine style*, *annotated chart*, *partitioned poster*, *flow chart*, *comic strip*, *slide show*, e *film/vídeo/animation*. A decisão sobre o gênero está relacionada a como os elementos da visualização narrativa ficarão dispostos na interface e como o usuário final poderá interagir com eles. Esta interação será definida pelo *co-designer*, indo desde pouca, ou nenhuma interatividade, até

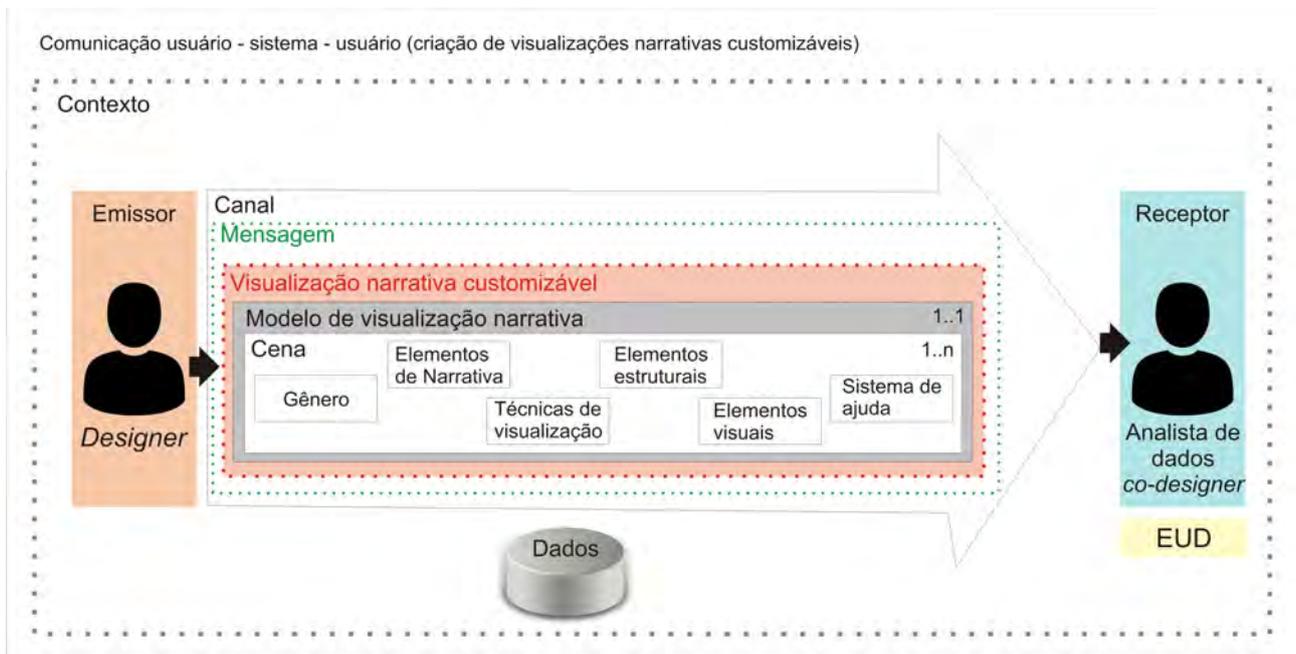


Figura 4.3 – Estrutura do 1º nível do CIViS: comunicação *designer* - *co-designer*.

muita interatividade, de acordo com os **modelos de visualizações narrativas** apresentados na Tabela 2.1: *Martini Glass Structure*, *Interactive Slideshow* e *Drill-Down Story* [SH10], que buscam o equilíbrio entre a narrativa idealizada pelo autor (pouca interatividade) e a liberdade de descoberta da história pelo leitor (muita interatividade). Neste caso, o *co-designer* pode optar pelo uso de combinações de gêneros de acordo com estes modelos híbridos.

Os **elementos de narrativa** surgirão a partir dos dados, pela identificação do conjunto de elementos proposto por Cruz et al [CM11]: evento, atores, localização e tempo. Neste contexto, o CIViS considera dados extraídos de mídias sociais que podem, ou não, apresentar todos esses elementos de narrativa e que, dos quais, podem ser extraídos: o usuário que postou, o conteúdo do *post* (texto, *link*, *hashtag* imagem, vídeo ou alguma combinação desses), a data, a hora e/ou o local. Com base nessa estrutura de dados, os elementos de narrativa podem ser mapeados da seguinte forma [Ghi17]: eventos podem ser identificados a partir de *hashtags* presentes no *post* ou pela identificação do nome de um evento no texto; atores são identificados a partir da estrutura do texto, com a localização do sujeito da frase e dos usuários autores do *post* ou mencionados nele; a localização é identificada a partir dos locais nos quais o evento ocorre, de um *post* georreferenciado, ou de locais mencionados no *post*; por fim, o elemento tempo é obtido por meio da data e hora do *post* ou data de ocorrência do evento. Como exemplo, apresentamos na figura 4.4 um *post* do Twitter com realce das informações que podem ser extraídas dele: evento, tempo, ator(es) e localização.

As **técnicas de visualização** podem ser quaisquer técnicas já consolidadas e conhecidas e que sejam adequadas ao tipo de dado que se deseja representar, podendo ser um gráfico de barras, um gráfico de linha, uma tabela, um mapa de calor, um grafo, entre outras técnicas. Os **elementos visuais** são os recursos de auxílio à visualização narrativa, tais como estruturas visuais,



Figura 4.4 – Exemplo de elementos de narrativa extraídos de um *post*.

realce e/ou orientação de transições entre técnicas de visualização ou entre cenas, por exemplo. Por fim, os **elementos estruturais** são aqueles que auxiliam e facilitam o uso da narrativa em visualizações por meio de ordenação, interatividade e mecanismos de mensagem, como já discutido na seção 2.1.

O **sistema de ajuda** é um componente baseado no método da Engenharia Semiótica para construção de sistemas de ajuda, proposto por Silveira et al. [SdSB03], que apoia *designers* em decisões de como falar diretamente aos usuários da aplicação sobre a aplicação e sobre como eles podem fazer um melhor uso dela. Estes componentes estão representados na figura 4.3.

Como dito anteriormente, uma visualização narrativa pode conter uma ou várias cenas, dependendo da organização planejada pelo analista de dados *co-designer*, e cada cena pode ter um ou vários componentes do CIViS. Vemos na figura 4.5, que uma cena possui um ou mais elementos de narrativas e também um ou mais gêneros. Cada conjunto de dados possui pelo menos um elemento de narrativa presente em sua estrutura, uma vez que esses elementos são extraídos dos dados para se determinar quais informações serão apresentadas aos usuários [Ghi17]. Os elementos estruturais e visuais estão relacionados com as técnicas de visualização e irão estruturar a narrativa, influenciando em como serão as interações do usuário final com a interface. Esses componentes de visualização narrativa estão presentes no espaço de design e irão guiar o *designer* e o *co-designer* na reflexão sobre o problema a ser solucionado.

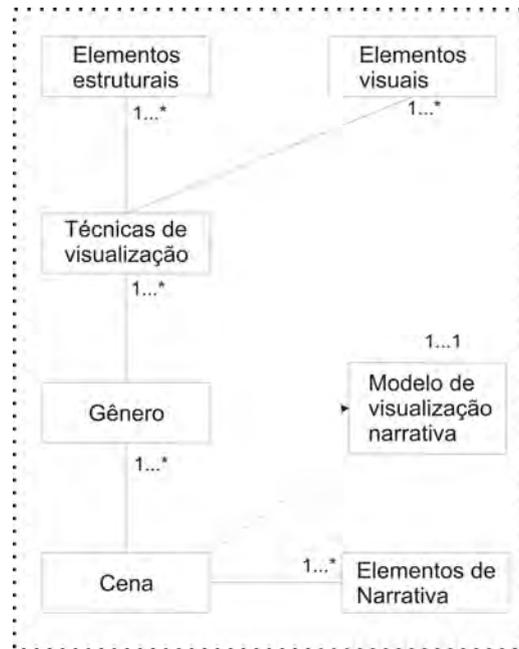


Figura 4.5 – Relacionamento entre componentes da cena de visualização narrativa no CIViS.

No segundo nível, o espaço de *design* é estruturado a partir de componentes que compõem a visualização narrativa customizada pelo analista de dados *co-designer*, como representado na figura 4.6. Esses componentes serão apresentados na ordem em que o analista os criou, ou seguindo um caminho pensado para a narrativa, mas o usuário tem a liberdade de interagir com eles e seguir outro caminho.

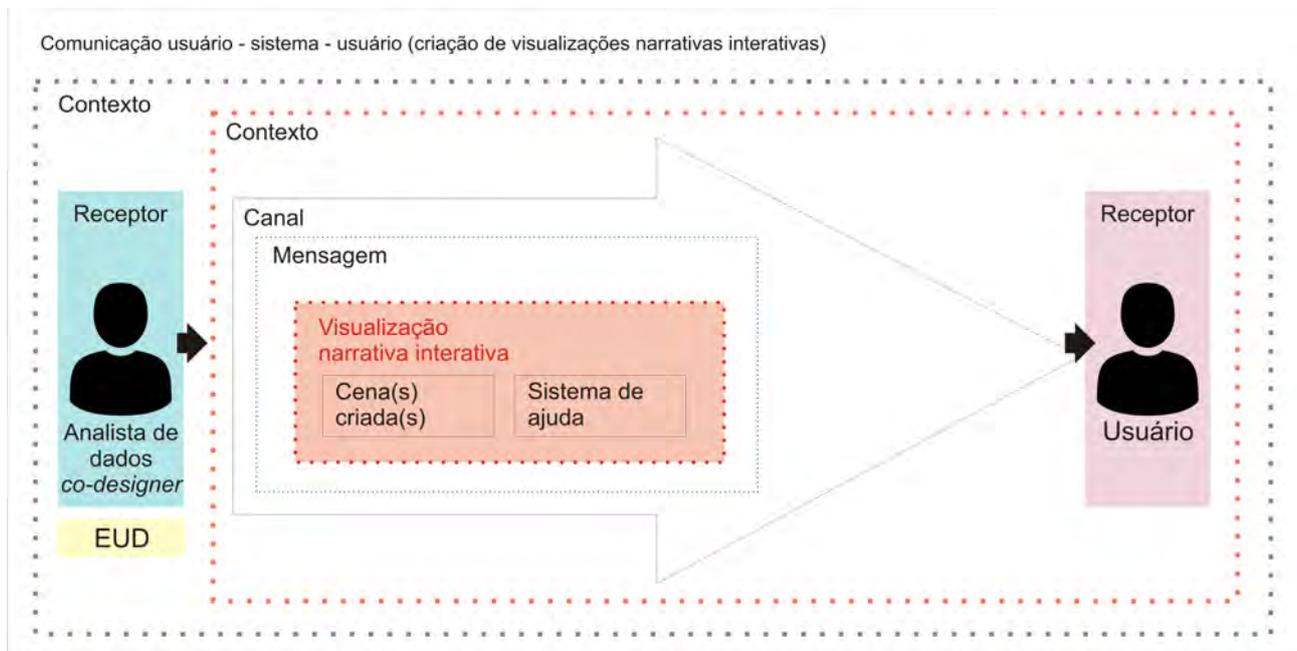


Figura 4.6 – Estrutura do 2º nível do CIViS: comunicação *co-designer* - usuário final.

O CIViS estrutura o relacionamento entre os componentes de visualização narrativa de acordo com as decisões de *design* em seus dois níveis. Para as decisões de primeiro nível, um

ponto importante desse estudo foi a identificação das personas, como descrito na seção 3.5. Aqui, as personas representam os perfis de analistas de dados que irão customizar as visualizações narrativas para os usuários. A definição das personas foi um passo importante neste estudo, pois possibilitou a identificação de características diferentes para perfis diferentes de analistas de dados, referentes às habilidades com ferramentas computacionais. O cenário 2, apresentado a seguir, ilustra este contexto.

CENÁRIO 2

João, Maria e José são pesquisadores e analistas de dados de mídias sociais, com perfis diferentes em relação ao domínio de ferramentas computacionais para este fim. No entanto, todos eles reconhecem a necessidade e a importância das ferramentas no processo de análise e gostariam de poder utilizá-las sem depender da colaboração de profissionais da área de tecnologia. Eles querem escolher e decidir como os dados serão apresentados, criar sequências de visualizações, representar mudanças ao longo do tempo, entre outras coisas. Diante disso, Carol, a *designer* de interação, continuou a sua investigação para entender melhor essas necessidades e tentar encontrar uma solução. A questão que guiou esse entendimento foi:

- Como atender os diferentes perfis de analista de dados?

Isso impacta no quanto cada perfil irá necessitar de recursos de ajuda no processo de *design*. Assim, para todas as personas identificadas, independente do nível de domínio de tecnologias, o CIViS sugere um sistema de ajuda como auxílio à customização da visualização narrativa. Da mesma forma, no segundo nível de *design* as decisões do analista de dados *co-designer* serão baseadas no que ele deseja comunicar ao usuário final e como ele pode ajudar o usuário a utilizar melhor a visualização narrativa interativa criada.

4.2 O papel do CIViS no processo de decisões de *design* de visualizações narrativas

Como apresentado na seção anterior, o CIViS descreve o processo de metacomunicação na criação de visualizações narrativas customizáveis, segundo a teoria da Engenharia Semiótica e sua visão sobre EUD [Sou05]. Assim, o *designer* necessita tomar várias decisões no momento da criação da mensagem que irá transmitir ao seu usuário. Nesta seção, apresentaremos um exemplo de como o CIViS pode ser utilizado para apoiar as decisões do *designer* e do analista de dados *co-designer* sobre a mensagem que irão comunicar.

4.2.1 Decisões do *designer*

Inicialmente, o *designer* precisa refletir sobre como ele poderá auxiliar os diferentes perfis de analista de dados *co-designers* na customização de visualização narrativa e o que ele irá oferecer para isso. Assim, algumas perguntas podem guiá-lo nessas decisões:

1) Como deve ser a visualização narrativa customizável de dados extraídos de mídias sociais?

Com base na nossa sugestão de uso do CIViS, a visualização narrativa customizável deve ser pensada a partir de elementos e características próprios desse tipo de visualização, que permitirão a definição dos componentes que serão oferecidos ao *co-designer*. Para iniciar o *design* de visualização narrativa customizável, o *designer* precisa definir os componentes que os analistas de dados *co-designers* poderão querer utilizar. O conjunto de decisões tomadas pelo *designer*, considerando todos os componentes, irá formar o modelo conceitual de visualização narrativa referente à sua customização. O cenário 3 ilustra o primeiro contexto: mensagem do *designer* para o analista de dados *co-designer*.

CENÁRIO 3

Após entrevistar analistas de dados de mídias sociais e compreender suas práticas e necessidades quanto à visualização de dados, a *designer* de interação Carol projetou um modelo de visualização em que o analista de dados atua como *co-designer* por meio da customização de visualizações narrativas, podendo contar uma história com seus dados por meio de visualizações. Para criar esse modelo, Carol baseou-se em estudos que discutem e apontam visualizações narrativas como uma forma de instigar a exploração dos dados no usuário final. Assim, ela elencou um conjunto de componentes que julga fundamentais na criação de uma visualização narrativa e que acredita que os analistas de dados *co-designers* queiram utilizar durante a customização da visualização. Esses componentes ficarão disponíveis para que o analista os utilize e combine de acordo com a visualização narrativa desejada. Carol também pensou em um sistema de ajuda para explicar ao analista de dados *co-designer* o que são e para que servem os componentes. Ela decidiu que cada visualização narrativa poderá ter uma ou mais cenas e cada cena deverá ter ao menos um gênero, elemento de narrativa, técnica de visualização, elemento visual, elemento estrutural e um sistema de ajuda. A questão que guiou esse entendimento foi:

- Como possibilitar que o analista de dados atue como *co-designer* da visualização narrativa interativa?

Após a definição dos componentes, o *designer* irá refletir e definir o tipo de ajuda que cada perfil de analista de dados *co-designer* pode desejar. Para isso, seguindo a sugestão do CIViS, o

designer deve se basear no método para criação de sistemas de ajuda *online* [SdSB03], em que ele poderá expressar as razões pelas quais projetou aquela visualização narrativa customizável e qual a melhor forma de o analista de dados *co-designer* utilizá-la.

2) Como será o sistema de ajuda?

Como dito anteriormente, o sistema de ajuda do CIViS é baseado no método da Engenharia Semiótica para construção de sistemas de ajuda *online*, proposto por Silveira et al. [SdSB03], em que o *designer* comunica explicitamente a sua visão de *design* aos usuários. No primeiro nível do CIViS, essa comunicação será entre o *designer* e o analista de dados *co-designer*, e no segundo nível, entre o analista de dados *co-designer* e o usuário final.

Com base em nossa sugestão de uso do CIViS, o *designer* deverá acoplar a cada componente uma explicação textual sobre o que ele é e para que serve. Considerando os diferentes perfis de analistas de dados *co-designers*, ele deverá disponibilizar pequenas porções de conteúdo sobre cada componente e permitir que o analista de dados *co-designer* acesse e aprofunde o acesso ao conteúdo de acordo com as suas necessidades, até receber as informações que deseja. O *designer* deverá, também, disponibilizar uma ajuda geral com a sua visão de *design*, com informações sobre o domínio e sobre visualização narrativa, bem como exemplos e/ou cenários de uso.

Para isso, o modelo sugere que o *designer* siga os passos do método para construção de um sistema de ajuda *online* [SdSB03], que é uma parte essencial da aplicação, de acordo com a Engenharia Semiótica. O *designer* pode, então, falar diretamente aos usuários - no caso deste estudo, aos analistas de dados *co-designers*, quais são as razões de seu *design* e como eles podem fazer um uso melhor da aplicação.

A partir das reflexões do *designer* sobre as respostas às duas perguntas discutidas e das sugestões do CIViS, espera-se que ele consiga estruturar o *design* de visualização narrativa customizável para o analista de dados *co-designer* que, então, poderá iniciar suas reflexões de *design*.

4.2.2 Decisões do *co-designer*

Assim como o *designer*, o analista de dados *co-designer*, no momento em que estiver pensando sobre a customização da visualização narrativa ou, em outras palavras, criando a narrativa (história) dos seus dados extraídos de mídias sociais, também pode ser guiado por algumas perguntas. A figura 4.7 apresenta a sugestão do CIViS de fluxo de ações que poderá apoiar essas decisões. O fluxo é cíclico mas, cada etapa pode ser revisitada e repensada a qualquer momento.

Assim, as decisões do analista de dados *co-designer* correspondem a encontrar respostas para as questões apresentadas. A seguir, apresentamos estas perguntas e respostas contendo sugestões de como o CIViS pode apoiar essas decisões de *design*.

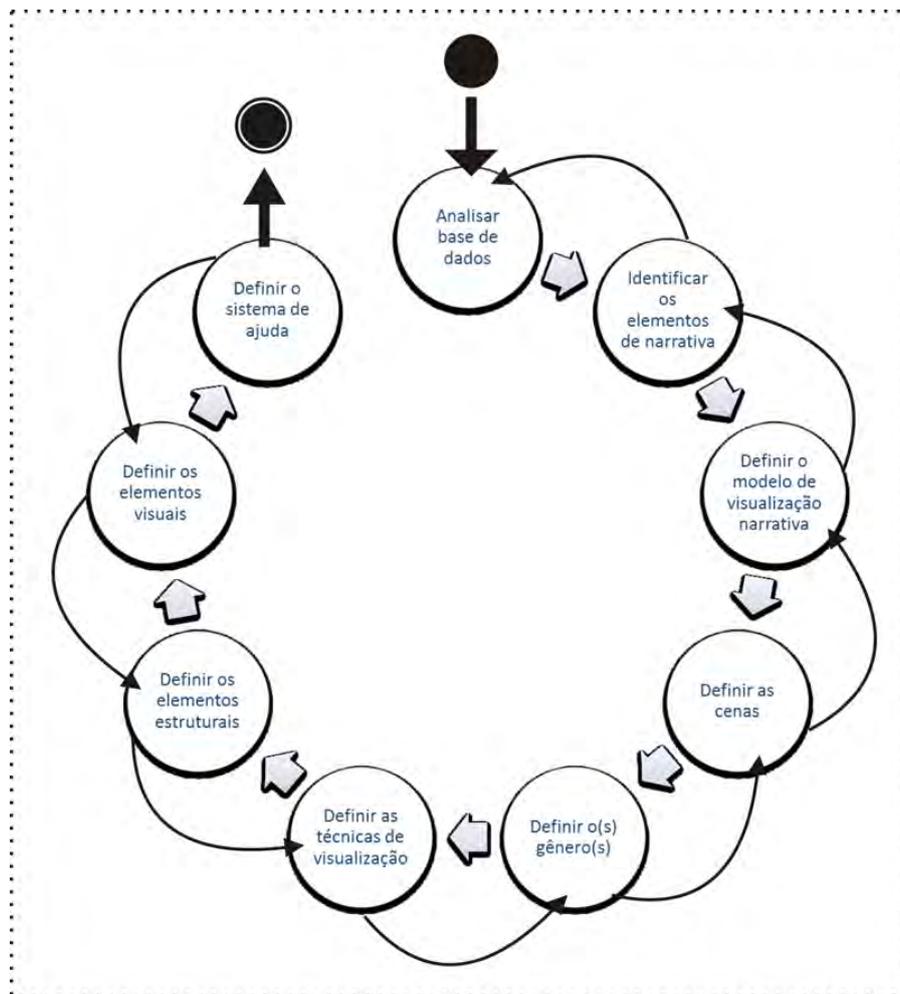


Figura 4.7 – Fluxo de ações/decisões do *co-designer* sugerido pelo CIViS.

1) Quais dados eu tenho e que história eu irei contar com eles?

A primeira decisão do analista de dados *co-designer* refere-se à que história irá contar, de acordo com os dados que possui e com o objetivo em relação àqueles dados. Para o entendimento dos dados, o CIViS se baseia no processo de transformação de dados em narrativas descrito por Lee et al. [LRIC15], em que o analista faz uma exploração inicial dos dados com o objetivo de entendê-los e criar fragmentos de uma base de dados. Estes fragmentos podem compreender, por exemplo, um subconjunto, dados sequenciados ou dados procedentes de *insights* obtidos por meio da análise, e podem ou não ser selecionados para uma representação visual na sua história.

No CIViS, consideramos que o analista já conhece a estrutura dos dados e sabe o que é possível extrair dela para ser apresentado na visualização narrativa. Suponhamos que ele tenha coletado dados durante os jogos olímpicos de 2016, no Brasil, e queira contar a história do comportamento dos brasileiros online no Twitter em relação à possível ameaça de terrorismo durante este evento. Neste caso, ele sabe quais dados ele coletou e, com isso, pode refletir sobre qual(is) história(s) irá contar com eles. Estes dados do terrorismo nas Olimpíadas, inclusive, serão o contexto para o estudo de uso do CIViS, realizado com analistas de dados *co-designers* (Apêndice E.6).

De posse dos dados, ele pode, por exemplo, identificar as datas das postagens, os usuários que postaram e os que foram mencionados em *posts*, locais, *hashtags* e demais dados que julgar necessários para a sua análise. De acordo com o objetivo da coleta, ele conseguirá, a partir dessa análise inicial da base de dados, refletir sobre a história que irá contar com esses dados.

2) Como irei estruturar essa narrativa?

Conforme nossa sugestão de uso do CIViS, a narrativa deverá ser estruturada em uma ou mais cenas e cada cena deverá conter, pelo menos, uma unidade de cada componente do modelo. Sugerimos, ainda, que as decisões de estruturação da narrativa sigam um fluxo de passos (Apêndice E.4) que permita ao analista de dados conhecer os elementos de cada componente e refletir sobre as escolhas deles. Seguindo esse fluxo, o analista pode ir aprofundando o seu conhecimento sobre as decisões tomadas e pode revisitar passos anteriores para rever algumas decisões. Em todo o processo, à medida em que precisar de auxílio ou de informação sobre como estruturar a narrativa, o analista de dados *co-designer* terá acesso ao sistema de ajuda pensado pelo *designer* para apoiá-lo na reflexão e tomada de decisão. Esta ajuda pode ser em forma de um material de apoio contendo informações sobre as possibilidades e sobre os componentes do modelo.

O caminho a ser percorrido pelo usuário/leitor na visualização narrativa será evidenciado pelo analista de dados *co-designer* no momento da criação da visualização. No entanto, neste mesmo momento, o analista de dados *co-designer* deve inserir componentes de interação que não estejam neste caminho pensado por ele mas, que permitam ao usuário explorar e percorrer outros caminhos. Isso significa disponibilizar vários outros dados relacionados ao contexto da história e que não foram evidenciados pelo analista de dados *co-designer* em sua narrativa.

3) Como poderei oferecer ajuda?

Dentro da proposta do CIViS, o analista de dados *co-designer* poderá incluir informações nos componentes de cada cena para explicar ao usuário final do que se trata e como ele pode interagir com a visualização para explorar os dados. Além disso, ele deverá incluir uma explicação sobre de que se trata aquela visualização narrativa interativa e qual foi o caminho pensado por ele para o usuário percorrer, embora a exploração de dados “fora desse caminho” seja livre.

Como sugestão de uso do CIViS, o analista de dados *co-designer* deverá exercitar a resposta às perguntas relacionadas a seguir, conforme indicado por Silveira et al. [SdSB03], para cada elemento presente na visualização narrativa interativa criada e, assim, construir o sistema de ajuda. As perguntas irão ajudá-lo a pensar nos mecanismos de ajuda que considera necessários oferecer aos usuários/leitores da sua visualização narrativa:

- O que é isto?
- Como faço isto?

- Onde está. . . ?
- Para que serve isto?
- E agora?
- A quem isto afeta?
- Por que eu devo fazer isto?
- Onde eu estava?
- De quem isto depende?
- O que aconteceu?
- Quem pode fazer isto?
- Epa!
- Por que não funciona?
- Existe outra maneira de fazer isto?
- Socorro!

Para isso, poderá fazer uso dos demais componentes do CIViS, como mecanismos de mensagem, interatividade, vídeo ou animação, entre outros, que julgar mais adequados para auxiliar o usuário final na exploração da visualização narrativa.

5. ESTUDO DO USO DO CIViS

No contexto da Engenharia Semiótica os *designers* podem contar com o apoio de ferramentas epistêmicas, compostas de linguagem e modelos conceituais, para auxiliar nas atividades de *design* de sistemas [Sou05]. Elas buscam proporcionar ao *designer* uma reflexão tanto sobre o problema que ele possui e precisa resolver, como, também, acerca de seu conhecimento sobre a solução que irá propor. Neste trabalho, desenvolvemos um modelo para guiar *designers* e analistas de dados *co-designers* na reflexão e customização de visualizações narrativas interativas, e aqui prosseguimos com sua análise.

Tendo em vista o entendimento do contexto, perfil dos analistas de dados e as suas necessidades, obtidos por meio dos estudos anteriormente descritos (capítulo 3), criamos o modelo CIViS e, em seguida, realizamos este estudo com foco na análise de seu uso pelos analistas de dados atuando como *co-designers* de uma visualização narrativa interativa. Assim, o presente capítulo apresenta e discute os resultados obtidos a partir deste estudo do uso do modelo CIViS por potenciais usuários, capazes de atuar como analistas de dados *co-designers*. Por se tratar de um estudo exploratório, nossa avaliação tem uma abordagem qualitativa, cuja metodologia é detalhada a seguir. Para realização desta atividade, os analistas de dados de mídias sociais foram convidados a utilizar o CIViS para auxiliar na criação de uma visualização narrativa interativa e, em seguida, relataram a sua opinião sobre o modelo. Tínhamos como objetivo obter relatos das possíveis compreensões do que é o modelo CIViS, dificuldades enfrentadas pelos analistas de dados *co-designers* em compreender o modelo e, finalmente, obter informações necessárias para sua melhoria.

Para isso, definimos as seguintes questões de pesquisa, a serem respondidas com base nos resultados obtidos a partir deste estudo:

- **RQ01:** Quão úteis são os componentes do CIViS para customização de uma visualização narrativa interativa?
- **RQ02:** Quais as impressões dos analistas de dados sobre o modelo CIViS em relação à reflexão que ele proporciona?

Nas seções seguintes serão apresentadas a prova de conceito do CIViS (seção 5.1), a metodologia escolhida para este estudo (seção 5.2), sua execução (seção 5.3) e os resultados obtidos (seção 5.4).

5.1 Prova de conceito do CIViS

Com o objetivo de facilitar o entendimento do CIViS e possibilitar uma visão geral das decisões do *designer* ao fazer uso do modelo, criamos um protótipo, em papel, contendo todos os

seus componentes. O objetivo do protótipo é apresentar ao analista de dados *co-designer* toda a estrutura que o CIViS oferece para auxiliá-lo na criação de uma visualização narrativa interativa. A figura 5.1 apresenta a ilustração do protótipo que contém folhas de papel, representando as cenas, cartas, *post-its* e lápis, representando os demais componentes do CIViS. A figura 5.2 apresenta a foto da mesa com os materiais organizados para a realização da tarefa pelos participantes do estudo.

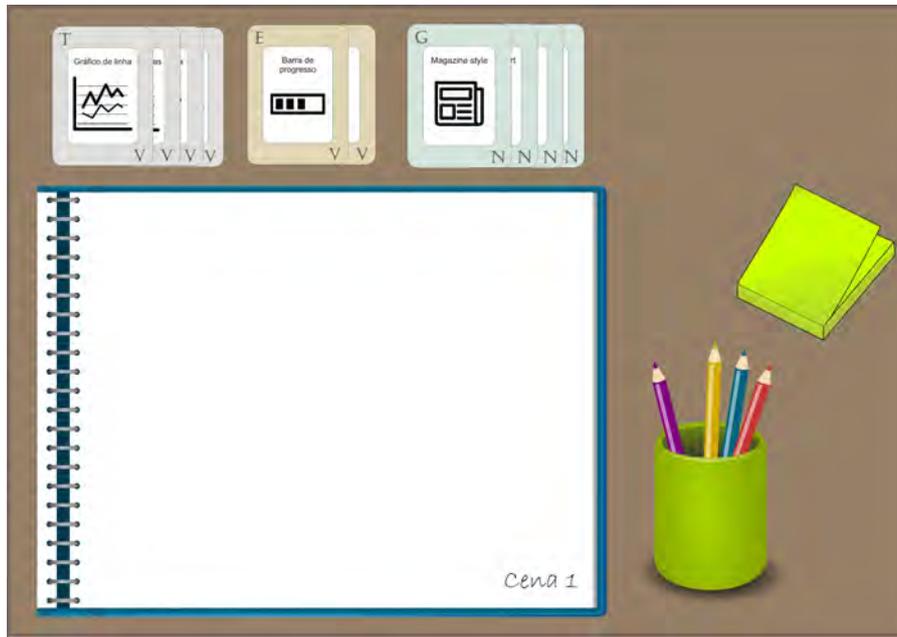


Figura 5.1 – Protótipo do espaço de *design* do 2º nível do CIViS: comunicação *co-designer* - usuário.

As cartas representam os componentes: Modelos de visualização narrativa, Elementos de narrativa (evento(s), ator(es), localização, tempo), Gênero, Técnicas de visualização, Elementos visuais, Elementos estruturais e Sistema de ajuda. Conforme apresentado na figura 5.3, a frente da carta apresenta as letras iniciais do tipo de componente, por exemplo: “T” e “V” para Técnicas de visualização, o nome do tipo do componente e o seu ícone.

5.2 Metodologia

Nesta seção são abordados o planejamento, procedimentos e decisões que, em seu conjunto, resultaram na metodologia escolhida para a condução deste estudo, que resultará na análise do modelo proposto. A abordagem inicial consistiu na elaboração de um exemplo de uso do modelo CIViS, para o qual foram propostos um cenário e uma tarefa, como descritos no apêndice E.6. Estes e demais materiais foram analisados a partir de consulta com um especialista em IHC, conhecedor da teoria Engenharia Semiótica e suas ferramentas epistêmicas, levando a uma melhor estruturação da tarefa proposta e seus respectivos materiais e métodos.

A metodologia adotada faz uso de métodos qualitativos interpretativos [DL13], que são apropriados para estudos como este, que visam explorar com maior profundidade uma (ou mais

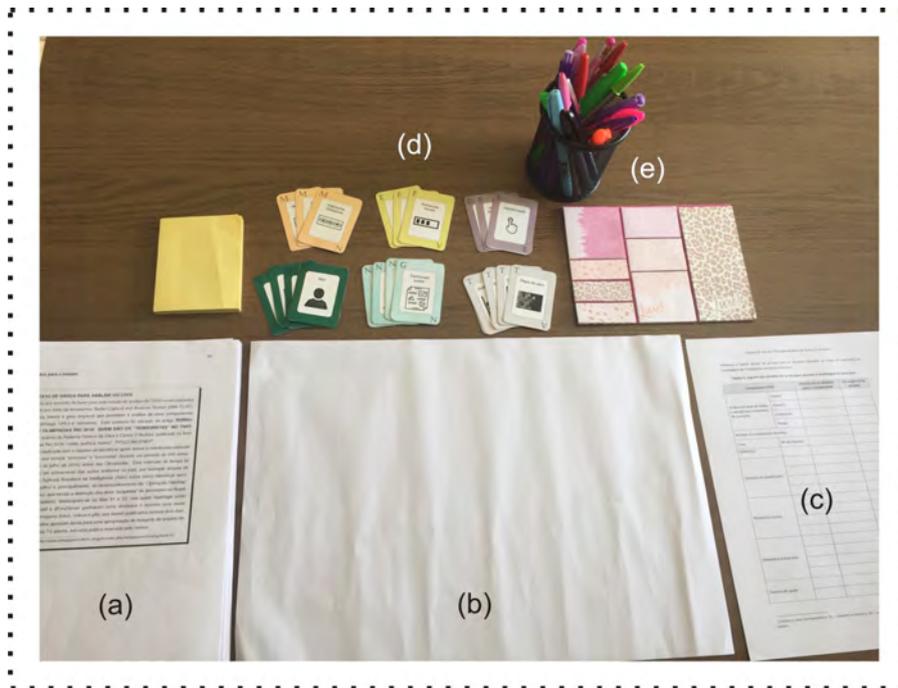


Figura 5.2 – Materiais organizados para realização da tarefa pelo analista de dados. (a) Cenário, dados, tarefa, fluxo de passos e tutorial do CIViS; (b) Cena; (c) Tabela de registro de decisões; (d) cartas que representam componentes do CIViS (havia mais de uma instância de cada carta; e (e) *post-its* e canetas coloridas.



Figura 5.3 – Modelo de carta do protótipo do CIViS em frente e verso. A frente contém o nome e ícone do componente ao qual a carta corresponde - neste exemplo a técnica de visualização gráfico de linha - e o verso contém a explicação desse componente, como parte do sistema de ajuda.

de uma) questão de pesquisa específica. De forma complementar, é importante notar que com a adoção de métodos interpretativos, o pesquisador geralmente está envolvido em uma experiência sustentada e intensiva com os participantes, o que pode moldar suas interpretações durante o estudo [Cre14].

Anteriormente à condução do estudo do uso do modelo CIViS, planejado para ser executado com a participação voluntária de analistas de dados de mídias sociais experientes, realizamos uma análise piloto para auxiliar nos ajustes finais dos materiais que compõem o modelo, assim

como possíveis ajustes no cenário proposto para a execução da tarefa. Este piloto durou 1 hora e foi realizado com uma Analista de Sistemas, com mestrado e experiência específica em análise de dados de mídias sociais. Segundo relato, as ferramentas de visualização disponíveis no mercado são, em sua maioria, de difícil aprendizado e compreensão. Após a execução do piloto, a voluntária refletiu sobre a atividade recém concluída e compartilhou suas reflexões. Em sua opinião, o modelo é útil na criação de visualizações e a inclusão da narrativa lhe parece promissora. Sentiu falta de um instrumento para registro das decisões, pois teve dificuldade em lembrar as razões das suas escolhas no decorrer deste processo. Isto levou à criação da “Tabela de registro de decisões” (ver o apêndice E.8), adicionada ao conjunto de materiais oferecidos pelo modelo CIViS. Feitos os ajustes, passamos à condução do estudo de uso do CIViS.

5.3 Perfil dos participantes

Para realizar o estudo do uso do CIViS, foram recrutados 4 participantes experientes em análise de dados de mídias sociais. Entendemos que a experiência dos analistas confere credibilidade às suas interpretações dos dados e aos usos que fizeram do CIViS, contribuindo com indícios acerca da validade do modelo. Para facilitar o entendimento sobre as falas e comportamentos dos participantes, agregamos alguns dados de perfil na tabela 5.1. O estudo com estes potenciais usuários *co-designers* do CIViS foi realizado em sessões individuais, em um laboratório de pesquisa da Universidade Federal de Minas Gerais - UFMG, com duração aproximada de 90 minutos cada sessão.

Tabela 5.1 – Perfil dos participantes do estudo do uso do CIViS.

ID	Idade	Gênero	Formação	Analisa dados extraídos de mídias sociais?	Persona a que se enquadra
P1	23	M	Comunicação Social	Sim	3
P2	34	F	Comunicação Social	Sim	2
P3	27	M	Comunicação Social	Sim	3
P4 ¹	60	F	História	Sim	1

Inicialmente, cada participante foi informado sobre o objetivo do estudo, que era perceber as possíveis compreensões sobre o que é o modelo CIViS, dificuldades enfrentadas pelos participantes em compreender o modelo e obter informações necessárias para sua melhoria. Após concordar em participar, o participante assinou o termo de consentimento livre e esclarecido (Apêndice E.2) e respondeu a um questionário para obtenção de dados sobre o seu perfil (Apêndice E.3). Em seguida, o CIViS foi apresentado ao participante que, após ouvir a explicação do modelo, foi convidado a utilizá-lo como guia para montar uma visualização narrativa interativa. Para apoiar o planejamento e a criação da visualização, o participante recebeu um cenário e um conjunto de dados (Apêndice E). Além disso, ele recebeu um material impresso com a explicação do CIViS, o

¹Participou também do Estudo 4 - Entrevistas com Analistas de Dados de Mídias Sociais, descrito no capítulo 3.

fluxo de passos do modelo (Apêndice E.4) e uma tabela para registro das decisões tomadas (Apêndice D). Assim, conhecidos o cenário, o conjunto de dados e o modelo, foi solicitado ao participante que realizasse a tarefa de criar uma visualização narrativa interativa utilizando o protótipo em papel do CIViS e seguindo o seu fluxo de passos, detalhados logo adiante.

Durante o *design* da visualização narrativa interativa, o participante deveria preencher a tabela de registro das decisões, de acordo com aquelas por ele tomadas. Ele também foi informado de que poderia utilizar o material de apoio sobre o CIViS e tirar dúvidas com a pesquisadora. Por fim, o participante recebeu orientação sobre o fluxo de passos para a utilização do modelo, explicando-se que este consiste em um ciclo iterativo, podendo ser revisitado para novas reflexões e decisões. Seguem os passos que compõem o fluxo recém descrito:

1. Analisar a base de dados que você possui e refletir sobre qual história é possível contar com esses dados.
2. Definir o modelo de visualização narrativa que irá guiar os passos seguintes.
3. Refletir sobre a(s) cena(s) que irá(ão) compor a narrativa e sua estrutura em relação à quantidade de *frames*, podendo ter 1 ou mais.
4. Identificar os elementos de narrativa presentes nos dados e definir quais serão utilizados.
5. Guiado pelo modelo de visualização narrativa escolhido, definir o(s) gênero(s), a(s) técnica(s) de visualização, o(s) elemento(s) estrutural(is) e visual(is) que irão compor a(s) cena(s).
6. Refletir e definir o sistema de ajuda para a visualização narrativa interativa criada.

Por fim, o participante foi convidado a participar de uma entrevista semi-estruturada sobre a tarefa e também sobre possíveis descobertas, problemas encontrados e suas opiniões sobre o modelo CIViS. Ao final da execução do estudo, obtivemos, como evidências empíricas: *i*) os protótipos criados pelos participantes; *ii*) os seus discursos e falas sobre as decisões tomadas nesse processo; e *iii*) as suas impressões e opiniões sobre o CIViS, coletadas por meio de entrevista semi-estruturada.

5.4 Análise da construção da narrativa

A análise do material coletado durante a realização deste estudo ocorreu em três etapas, segundo a natureza do domínio desses dados. Primeiramente, analisamos as respostas do questionário usado para obtenção do perfil dos participantes; então, prosseguimos com a análise do *design* de visualização narrativa interativa criados por eles; e, por fim, analisamos as respostas da entrevista semi-estruturada sobre a tarefa e sobre o modelo CIViS.

A partir do questionário de perfil, verificamos que todos os participantes fazem ou já fizeram análise de dados extraídos de mídias sociais e já utilizaram ferramentas de software para

ajudar nas análises, tais como Gephi, NodeXL, Tableau, TAGS, Netvizz, TCAT, entre outros. Apenas um participante não utiliza software (P4), mas possui parceiros que colaboram com essa parte nas análises que faz. Em relação à facilidade de uso ou aprendizado, de uma forma geral, os participantes disseram que algumas dessas ferramentas são fáceis e outras, como o Gephi, são muito complexas. Para P1 *“existe uma diferença entre os [programas de] software pensados para a experiência do usuário e os que se preocupam exclusivamente com os dados”*. Já no que diz respeito à forma que escolhem para apresentar os dados analisados ao usuário final, consumidor de suas análises, os participantes relataram que os apresentam em forma de texto, por meio de vídeos, ilustrados por gráficos, por meio de gráficos impressos em papel ou *print* dos mesmos inseridos em apresentações de *slides* e também apresentam os dados em relatórios qualitativos.

Quando questionados sobre a possibilidade de apresentar os dados extraídos de mídias sociais por meio de uma narrativa, utilizando técnicas de visualização de dados, todos responderam achar a ideia muito interessante. P1 disse que *“seria perfeito se eu pudesse construir uma narrativa assim! Seria muito interessante porque quando a gente fala em narrativa, já pensa em tempo. E a dimensão da temporalidade ainda é complexa”*. P2 vê a possibilidade do uso de narrativas como uma forma de contemplar o caráter qualitativo dos dados. P3 acredita que *“seria o modo ideal para apresentar os dados. Entretanto, criar uma narrativa eficaz com esse tipo de informação levaria muito tempo e seria difícil”*. Já P4 acredita que teria que ser algo dinâmico, *“que conseguisse mostrar o processo de mudança permanente das redes”*. É importante destacar que os participantes responderam a esse questionamento antes de conhecerem o significado formal de visualização narrativa, que foi enunciado no momento da apresentação do modelo CIViS. Antes disso, o conceito formal era por eles desconhecido.

Após responderem o questionário de perfil, os participantes iniciaram o *design* da visualização narrativa interativa de acordo com o fluxo de passos do CIViS (Apêndice E.4) e com base no cenário de dados disponibilizado (Apêndice E.6), conforme descrito anteriormente. Durante esse processo reflexivo, ao longo do qual os participantes compunham suas cenas, fomos registrando alguns momentos dessa criação por meio de fotos, mostradas logo adiante, seguidas das falas dos participantes sobre as decisões tomadas. A primeira dessas imagens registra a mesa de P1, mostrada na figura 5.4. Segundo P1:

“Pra contextualizar o evento, um poster é bem legal. Pra permitir a navegação eu colocaria um monte de gráfico. Trabalhar com a dimensão de tempo é fundamental, um mapa para apresentar a geolocalização dos tweets, um mapa de calor é legal para representarmos a visualização de hashtags, pra ver qual está com mais intensidade que outra, e o grafo com a distribuição de perfil. Tempo, espaço, hashtag e perfil. Isso poderia estar em um poster ou então em um flow chart, se for nessa sequencia que acabei de descrever. Mas, aí a narrativa fica fechada, né. Quando eu penso em utilizar o Drill Down, eu estou pensando que o grafo será interativo, você pode clicar em um nó, movimentá-lo de lugar, visualizar as informações a partir daquelas conexões. Se você pensar bem mesmo, com um mapa você pode fechar tudo. Você coloca um mapa na

tela e ao lado dele algumas opções: distribuição por localização, várias coisas, como quais equipes ficaram onde, linha de tempo - você dá um play e aí vai aparecendo as hashtags no mapa ao longo do tempo, ou então você vai colorir de acordo com qual hashtag repercutiu mais. Se você tem recursos de interatividade, dá pra fazer “mágica” e mostrar muita coisa. (...) Eu fico pensando no conteúdo da mensagem, isso me preocupa. Não estou desmerecendo a área da computação (que processa linguagem natural) que consegue interpretar linguagem textual, mas se você faz uma análise humana de uma hashtag, na minha opinião, essa análise será mais rica.”



Figura 5.4 – Foto do momento em que P1 finalizava as suas decisões sobre a visualização narrativa interativa a ser criada.

Percebemos aqui que P1 acha difícil que algoritmos realizem análises de forma tão rica quanto um ser humano. Note que, sua preocupação com o conteúdo de uma *hashtag* tem relação com a qualidade dessa análise de conteúdo. Além disso, observamos que P1 compreendeu os componentes do CIViS e conseguiu associá-los de forma que a narrativa pensada por ele (e guiada pelo leitor) viesse a ser concretizada. Deixou claro sobre o quanto acha importante que a narrativa seja guiada pelo leitor, para que ele tenha a liberdade de interagir com os pontos da história que mais lhe interessarem. Mesmo que o autor seja quem decide, de fato, o que será apresentado ao usuário-leitor a partir das interações, para P1 não faz sentido o autor determinar o caminho a ser seguido na visualização narrativa. Continuando as suas reflexões, ele disse:

“Um das outras opções do grafo, o nó será o perfil do usuário. Se eu clicar no nó, você terá acesso a quais mensagens esse nó deixou rastro na rede, pra poder construir alguma construção qualitativa daquilo ali... e não deixar um computador processando aquilo ali. A princípio, é muito bom andar com os dois juntos: ferramentas computacionais e análises humanas. Uma outra exibição que eu colocaria seriam das tabelas: qual

foi o perfil mais retweetado e qual foi a mensagem mais retweetada. Acho legal usar realce para chamar a atenção para as análises mais "marcantes", recursos de navegação que te oriente onde você está e te permita voltar para onde você estava antes."

Assim como P1, para criar a sua visualização narrativa, P2 escolheu o modelo *Drill Down Story*. Sobre os elementos de narrativa, conforme disponibilizados pelo modelo CIViS, comenta:

"Bom, a primeira coisa que eu consigo identificar é o tempo, que é o tempo da coleta, né? Foi a primeira coisa que me chamou a atenção. Outra coisa que a gente consegue ver são os eventos, tanto o evento das Olimpíadas quanto os eventos sociopolíticos que estão permeando também a narrativa. Depois de abrir estes dados, eu consigo identificar quem são os atores e eu acho muito difícil conseguir, nesse contexto, trabalhar com a localização. Eu queria muito ver de onde estava vindo a conversa, mas é complicado ainda. Então, eu acho a localização de dados de mídias sociais o pior índice para trabalharmos em um modelo narrativo. O contexto de ser um evento internacional acontecendo no Brasil, pode ajudar em alguns aspectos. O tempo, pra mim, é super importante, porque eu gosto de observar os picos de conversa, em que eu vou identificar o que me interessa."

Note que P2 foi capaz de reconhecer nos dados os elementos do modelo, tais como tempo, eventos, atores e localização. Durante seu processo criativo, o participante continuou compartilhando seu raciocínio, dizendo: *"eu quero colocar um frame sobre o evento e colocar um Flow Chart para apresentar os dados que quero. A questão do tempo e do ator, isso pra mim, funciona muito bem junto. Até a localização! São coisas que, no final, vão determinar, no início, uma análise contextual, e, depois, algumas situações específicas. Os gráficos ajudam muito quando a gente precisa mensurar."* A mesa de P2 é apresentada na figura 5.5.

P2, assim como alguns participantes do estudo 4 (capítulo 3), demonstrou dificuldades em manipular e entender a técnica de visualização grafo, possivelmente devido à dificuldade relatada sobre a complexidade das ferramentas que possibilitam a criação dessa técnica, como menciona na explicação de suas escolhas:

"Eu vou colocar o slideshow aqui e depois eu revejo essa minha escolha. Eu tenho muita dificuldade com grafo, porque para nós, da comunicação, ele é muito complexo. E às vezes ele evidencia muito o nó e, pra mim, aqui o nó não tem tanta importância. Não acho que ele traz tanta informação assim. Então, pensando que uma narrativa tem de ser de fácil compreensão, eu usaria a tabela, o gráfico de barras, e o de linha. Aqui, no primeiro momento contextual eu colocaria a tabela com dados da olimpíadas. No gráfico de linhas eu poderia comparar se teve mais hashtags sociopolíticas ou mais comerciais. Um gráfico de barras também para mostrar as hashtags mais usadas. Então, isso tudo é só dado contextual. Apresentaria quais eventos que aconteceram até chegar nessa história de terrorismo. Para o elemento evento, a ordenação é importante, a



Figura 5.5 – Foto do momento em que P2 finalizava as suas decisões sobre a visualização narrativa interativa a ser criada.

interatividade é importante o tempo todo, o realce e a barra de progresso também são muito úteis. E o mecanismo de mensagens vai servir pra colocar títulos e algumas poucas informações breves. E a interatividade, né, que dá liberdade para o leitor percorrer o caminho dele e interagir com os elementos que mais lhe interessem.”

P3 escolheu o modelo *Interactive Slideshow* e disse que começaria a narrativa falando sobre o evento Olimpíadas no Rio, por meio de uma animação, e em seguida falaria dos atores envolvidos, por meio de um *partitioned poster*. Segundo ele, “*é sempre bom ter um texto no meio das imagens e gráficos para introduzir e explicar o contexto*”. Para mostrar os atores, usaria um grafo com o recurso de anotações, dizendo de onde postaram, em que dia e horário, etc. E, por meio do grafo, seria possível observar a organização dos *clusters* formados na rede naquele período. Assim, “*os leitores podem percorrer e interagir com a visualização e descobrir, por exemplo, quem estava falando o quê nesse contexto de terrorismo e olimpíadas*”. Depois disso, viria um mapa com as marcações de localização dos *tweets*, porque, segundo ele, é interessante visualizar no mapa a origem das postagens. E ainda ressaltou que, “*os recursos de interatividade estariam presentes em todos os gráficos*”. É importante lembrar que P3 não escolheu o elemento de narrativa tempo para a sua visualização, no entanto, utilizou técnicas que dependem desse elemento. Ressaltamos que, apesar de ter pensado no gráfico de linhas, inicialmente, no final ele descartou, por achar desnecessário depois de toda a reflexão. A mesa de P3 é apresentada na figura 5.6.

P4 escolheu *Martini Glass* porque acredita ser o melhor para criar uma visualização narrativa interativa. Para explicar sua escolha, começou perguntando “*de onde que apareceu discussão sobre terrorismo no Brasil, que nunca teve? De onde tem essa relevância de quantos tweets se coletou sendo que pouquíssimas pessoas usam o Twitter no Brasil? Então, eu não conseguiria fazer uma visualização sobre isso se eu não explicasse o contexto em que isso aconteceu.*” Dadas as observações sobre o cenário e respeitando a lógica do modelo de narrativa escolhido (*Martini*



Figura 5.6 – Foto do momento em que P3 finalizava as suas decisões sobre a visualização narrativa interativa a ser criada.

Glass), P4 começou a sua narrativa escolhendo o gênero Vídeo, seguido de *Flow Chart* e, para finalizar, o gênero *Magazine Style*. P4 gostaria que o seu leitor tivesse oportunidade de buscar mais material na Internet, além de todos os que ele disponibilizou na visualização, debatesse com essa narrativa que ele criou, ou acrescentasse ou discordasse, mas apresentasse argumentos, ou seja, teria que ter oportunidade de o leitor adicionar dados, além de seus posicionamentos. Segundo P4, seria muito interessante conseguir buscar “*imagens, vídeos, gráficos por meio da imagem que você possui. Eu tenho uma foto de tal momento, posto e o sistema de busca me traz de volta todas as fotos daquele momento ou daquela pessoa*”. Em sua narrativa, P4 apontou que, para o contexto posto, os elementos de narrativa que usaria seriam evento, tempo e localização. A mesa de P4 é apresentada na figura 5.7.

Segundo P4, a discussão sobre terrorismo no Brasil começou porque a Copa e as Olimpíadas exigem determinadas legislações de cada país por onde elas passam, e uma delas é “*essa coisa do terrorismo*”. Para a participante, essa lei possibilita não só formas de treinamentos do aparato repressor de cada país, como a venda de armas: “*essa legislação abre uma brecha no aparato político dos Estados para que possibilite esse tipo de coisa. Óbvio que na Inglaterra, na Alemanha, nos Estados Unidos, etc., isso faz sentido mas, aqui não faz.*” Com isso, P4 quis mostrar que o contexto da discussão apresentada no cenário deste estudo é, na verdade, muito maior do que está descrito e, em certa forma, é desproporcional a preocupação com relação ao terrorismo no Brasil. Conforme disse:

“A mídia fez um estardalhaço e aí as pessoas começaram a falar e discutir sobre terroristas. Então isso tem a ver com essa parte de, claramente, você criar um contexto que justifique a lei, ao mesmo tempo que acontece uma mudança na discussão. E são nesses momentos em que se mudam os contextos e você dá uma historicidade para o terrorismo no Brasil. No fundo, essa legislação sobre o terrorismo é uma proteção dos

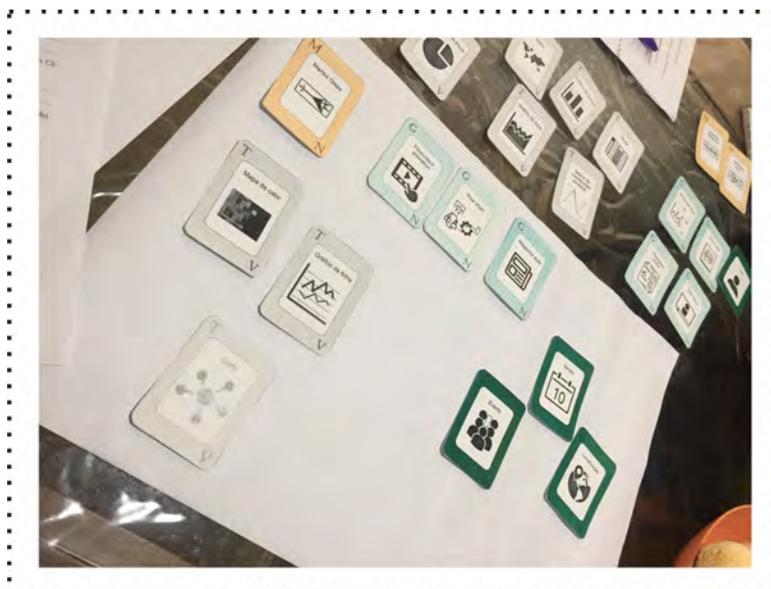


Figura 5.7 – Foto do momento em que P4 finalizava as suas decisões sobre a visualização narrativa interativa a ser criada.

*negócios estrangeiros em outros Estados nacionais. E aí, essa “operação hashtag”, que é ridícula, criou toda essa coisa de anti-terror e tal. Por isso, você tem que fazer uma contextualização para o seu leitor, criando uma lógica, considerando que a imprensa é fundamental nisso. Porque, se você não entender o tempo e essas múltiplas temporalidades, que fala de uma coisa, que vira outra, que aciona outras, se você não entender essa dinâmica, isso pode virar uma piada, entendeu? Mas, na verdade, como falei, há uma lógica maior do que isso. Nesse caso, então o elemento **evento** é importante, porque é o contexto deste evento.”*

Note que P4 enfatiza a questão dos momentos em que as mudanças de contexto ocorrem porque, segundo ela, é isso que as visualizações de dados ainda não conseguem representar. Trata-se dessa dinamicidade que transforma o contexto da discussão. Por isso, frisa a importância da contextualização na apresentação dos dados e informações aos leitores. Com relação às escolhas específicas para a criação da sua visualização narrativa, P4 não escolheu **atores** porque acha que é muito diluído, que não especificam e evidenciam quem são eles, em termos de relevância. Por fim, escolheu utilizar localização (Brasil), mas não acha tão fundamental quanto os outros dois elementos anteriormente escolhidos (evento e tempo). P4 continuou descrevendo as decisões sobre a criação de sua visualização da seguinte forma:

“Eu usaria Mapa de Calor para mostrar a intensidade dessa discussão cronologicamente, antes e depois desse período que foi recortado para essa pesquisa que você trouxe de cenário aqui. Poderia pegar a mídia mesmo, porque isso bombou mais na mídia, pra mostrar as notícias ligadas a esse tema. (...) Primeiramente, eu queria mostrar, ao longo do tempo, a quantidade de tweets nesse período, então vou usar o Gráfico de Linhas porque acho que as pessoas lêem melhor. Escolho também um Grafo, para

fazer relação do conteúdo postado nas redes formadas, como imagens, por exemplo. Acho que nessa visualização narrativa, com esses dados, eu usaria esses três gráficos. Mapa, é bobagem, pois aconteceu no Brasil. Eu usaria os mecanismos de ordenação para ordenar cronologicamente as notícias, pra mostrar a dinâmica da "coisa", mostrar que tem um momento que houve mais empenho, né. Então, na minha narrativa, eu vou dizer 'Olha, juntou isso com isso e com aquilo, nesse momento.' Então eu usaria a ordenação, usaria mecanismos de mensagem, e extrairia da mídia, inclusive, quando que alguma coisa aqui deu manchete. As pessoas discutem isso muito pouco mas, quando um jornal resolve transformar aquilo em manchete, onde a notícia foi diagramada, porque foi veiculada em tal lugar... tudo isso é importante para as pessoas entenderem melhor, debaterem, discutirem o assunto. Usaria mecanismos de realce porque quero evidenciar o que eu, como autora, estou discutindo, pra mostrar o meu caminho de narrativa. Um recurso importante que eu usaria é esse [referindo-se à Orientação de Transições] de a pessoa poder ir para outro lugar (outra visualização). Por exemplo, se eu digo que em outros países aconteceu algo diferente e disponibilizo um link para isso, meu leitor pode clicar, ir lá ver do que estou falando, e voltar para o ponto que estava na minha visualização narrativa.

Percebemos, então, que P4 entendeu o propósito do modelo de visualização narrativa por ele escolhido, *Martini Glass*, e que a participante valoriza customizações que permitam evidenciar suas escolhas para a visualização narrativa. É importante, neste contexto, que o leitor de suas visualizações possa navegar e visitar outros pontos em uma dada visualização sem, contudo, perder o contexto, podendo retornar ao ponto de origem. Com isso, a linha da narrativa se mantém aquela por ela planejada, porém, com a flexibilidade necessária para gerar uma visualização narrativa interativa.

Observamos que todos os participantes conseguiram utilizar o modelo de forma completa, pois foram capazes de decidir sobre cada um de seus componentes no *design* da visualização narrativa interativa solicitada. É importante ressaltar que, durante a realização da tarefa, os participantes falavam sobre suas decisões e opiniões em relação aos recursos e componentes que usariam ou não no *design* e, com isso, foi possível perceber como eles iam se apropriando dos conceitos de cada componente.

As decisões tomadas pelos participantes em relação aos componentes do seu *design* são apresentadas a seguir. Para o componente **modelo de visualização narrativa**, os participantes fizeram escolhas diferentes: P1 e P2 escolheram o modelo *Drill Down Story*; P3 escolheu o *Interactive Slideshow*; e P4 optou pelo modelo *Martini Glass*. P1 argumentou que o modelo por ele escolhido é “*melhor porque dá a liberdade para o leitor já interagir desde o início com o que ele mais se interessa em saber daquela narrativa*”. Ele acha que a construção da história deve partir de quem está lendo, e usou como exemplo o recurso das abas existentes em várias aplicações, que é uma forma de dar liberdade para o leitor navegar do jeito que ele quiser. O participante P3 justificou a sua escolha dizendo: “*eu considero esse modelo melhor, porque o Martini Glass me parece um*

pouco engessado e o Drill Down Story dá uma ideia de ser hipertextual, não sei... fica uma coisa meio sem começo e nem fim, embora pareça bacana por ser bastante interativo". Já P4 escolheu o *Martini Glass* pois, segundo ele, *"é o modelo adequado para essa narrativa, pois é preciso, primeiro, contextualizar os dados para o leitor, as fontes desses dados, como se deu, como começou essa história e, aos poucos, ao longo da narrativa, ir liberando possibilidades de ele interagir com os dados"*. Assim, como podemos ver, é possível perceber, por meio das falas e questionamentos dos participantes, que eles conhecem seu público alvo e como melhor apresentar a narrativa, dado o contexto.

Para o componente **elementos de narrativa**, os participantes P1 e P2 identificaram e selecionaram os quatro elementos da base de dados que estavam relacionados ao cenário: ator(es), evento(s), localização e tempo. Eles consideraram importante inserir todos os elementos porque, mesmo sendo uma coleta durante e sobre as Olimpíadas, outros eventos poderiam ter ocorrido e ganhado destaque nesse período, como foi o caso da "Operação *Hashtag*", por exemplo. P3 identificou todos esses elementos na base mas, selecionou para o seu *design* apenas três deles: ator(es), evento(s) e localização. Ele pensou sobre as possibilidades de gráficos que poderia criar a partir desses elementos, mas não justificou a não escolha do elemento "tempo". P4 também, apesar de ter identificado todos esses elementos, preferiu utilizar no seu *design* somente: evento(s), localização e tempo. Para ele, no contexto dos dados em questão, a localização (no caso, o Brasil) ganhava destaque porque, segundo ele, se fosse em outros países, como a Inglaterra, por exemplo, as coisas não teriam acontecido da mesma forma.

Para o componente **cena**, todos os participantes foram orientados a modelar o *design* de uma cena apenas, devido à disponibilidade de tempo que eles possuíam para participar do estudo. Assim, eles decidiram sobre sua organização em termos de número de *frames*. Já para o componente **gênero de narrativa**, P1 decidiu utilizar os gêneros *Flow chart*, *Partitioned Poster* e *Annotated chart*, porque considerou que esses gêneros representariam o que, inicialmente, havia pensado a partir do modelo de narrativa que havia escolhido. P2 selecionou *Flow chart* e *Slideshow*, por motivo semelhante ao de P1. Nesse momento ainda inicial do fluxo de passos do CIViS, os participantes, todos surpreendidos com a ideia do componente de modelo de visualização narrativa, refletiram sobre os gêneros com base no modelo escolhido. P3 escolheu dois gêneros: *Partitioned Poster*, porque, segundo ele, é bom para falar dos atores; e *Film/Video/Animation* para evento. Por fim, P4 selecionou os gêneros *Film/Video/Animation*, *Flow chart* e *Magazine Style*. Segundo ela, o vídeo é uma boa forma de contextualizar a história e, como escolheu o modelo *Martini Glass*, *"se encaixa bem na questão da pouca interatividade no início da narrativa"*.

Para o componente **técnicas de visualização**, os participantes apresentaram uma certa facilidade na escolha das técnicas, possivelmente pelo fato de estarem familiarizados com esse contexto, não falaram muito sobre as suas decisões. P1 selecionou grafo, tabela, gráfico de linha, mapa e mapa de calor, enquanto P2 escolheu tabela, gráfico de linha e gráfico de barras e P4 selecionou mapa de calor, gráfico de linha e grafo. P3 foi o único que comentou extensivamente sobre suas escolhas: mapa, mapa de calor, gráfico de linha e grafo, que, segundo ele, *"no contexto*

do Twitter, funciona bem, pois você vê os clusters bem organizados”. Pontuou que, apesar de o Twitter possuir poucos *posts* geo-referenciados, acha interessante organizá-los por localização no mapa. Também explicou que, com o mapa de calor, poderia mostrar onde mais intensamente se falou sobre o assunto abordado. Selecionou, ainda, o gráfico de linha para poder mostrar o volume de *tweets* e os termos mais frequentes no decorrer do tempo, afinal, segundo ele, o gráfico de linhas é o “*pretinho básico das visualizações*”. Contudo, concluiu dizendo que “*esse último ficaria melhor usando nuvem de palavras*”. De fato, foi possível perceber que todos os participantes têm conhecimento sobre as técnicas de visualização, sendo importante ressaltar que sentiram falta de uma carta que representasse a nuvem de palavras, por considerarem uma boa técnica para ilustrar momentos importantes da narrativa.

Com relação ao componente **elementos visuais**, P1 optou por realce e orientação de transição e P2 por realce e estruturas visuais, ao passo que P3 e P4 selecionaram todos os elementos: realce, estruturas visuais e orientação de transição. Observamos que os participantes compreenderam o propósito desses elementos, que visam auxiliar o leitor na identificação dos pontos de destaque da narrativa. As escolhas foram relacionadas, principalmente, às possibilidades de destacar pontos importantes da narrativa e de permitir ao leitor a compreensão dos caminhos percorridos por ele e a relação com o que já foi visualizado e analisado por ele (leitor). Comentando sobre suas escolhas, P1 notou que o CIViS oferece elementos que possibilitam a criação de uma narrativa diversificada, lançando mão de vários recursos, tais como como texto, imagem, gráficos e interação. Contudo, reforçou a importância do uso de textos associados a todas as visualizações, ou seja, pequenos resumos, títulos ou comentários, que ajudam a contextualizar a narrativa.

Para o componente **elementos estruturais**, P1 escolheu interatividade e mecanismos de mensagem; P2 e P4 selecionaram interatividade, ordenação e mecanismos de mensagem; e P3 selecionou mecanismos de mensagem e interatividade. Com relação a este último elemento estrutural, para P1, “*se o usuário/leitor pudesse navegar e fazer suas marcações para destacar o que ele achar mais relevante, seria muito legal*”. Percebemos nesta fala o desejo de P1 em proporcionar ao leitor possibilidades de realçar os pontos que achou mais interessantes na narrativa. Por outro lado, também notamos a preocupação de P1 em disponibilizar recursos de colaboração aos seus pares, como vemos nesta fala: “*seria interessante ter um mecanismo de comentários na visualização narrativa para o caso de o usuário ser um outro analista de dados que queira opinar ou colaborar*”. Em função das escolhas dos participantes, percebemos que eles valorizam a inserção de mecanismos de mensagem e interatividade na visualização narrativa.

Finalmente, para o componente **sistema de ajuda**, os participantes foram orientados a refletir não somente sobre a forma como ofereceriam o sistema de ajuda, mas também sobre o conteúdo, a partir das perguntas que constam no Tutorial (Apêndice D e que podem ajudar o analista de dados *co-designer* a criar respostas para aquelas possíveis perguntas que o usuário final poderá fazer ao se deparar com alguma parte da visualização. P1 e P2 definiram mecanismos semelhantes, consistindo em um símbolo de ajuda associado a cada elemento da narrativa, tal que, ao passar o mouse em cima, uma caixa de texto aparece, não somente com a descrição de o quê aquele

elemento significa mas também de como o recurso pode ser utilizado. Segundo P1, associar um recurso de ajuda para cada elemento é a melhor decisão: *“Estou cansado de ver visualização em que está lá dito o que é, mas, e aí, o que você faz com isso? Não diz como utilizar. Por exemplo, se vou apresentar dois grafos de tweets, relação de hashtags e relação de usuários, é preciso explicar para o usuário que em uma opção o nó é a hashtag e na outra o nó é o usuário, e dizer para ele o que acontece quando ele clicar ou expandir uma área do grafo”*. P4 concorda com P1 e P2 quanto à forma de apresentação do sistema de ajuda, contudo, acha importante adicionar o *link* para as fontes de onde os dados foram extraídos, permitindo, assim, que os leitores possam aprofundar ainda mais a análise sobre o assunto. Por outro lado, P3 considera melhor utilizar um tutorial e uma animação com um caso de teste explicando o que é e para que serve cada recurso: *“Como já mexi muito com game, eu prefiro muito mais um tutorial, uma animação com um caso de teste, com dados falsos, de exemplo... para um software novo, que você nunca usou, acho que isso funciona melhor”*, disse ele. Complementou, ainda, que nessa animação deveria haver recursos para o usuário interagir e verificar como aquele elemento funciona, certificando-se de que entendeu.

Perguntamos a eles se sentiram falta de algum componente ou característica que o modelo não havia tratado e que poderia ser necessário em uma narrativa. Os participantes ressaltaram que nunca haviam pensado na criação de visualizações com narrativas e, por isso, não conseguiam imaginar o que poderia estar faltando, uma vez que, com o que o CIViS estava oferecendo, eles já conseguiram ampliar a visão sobre possibilidades de visualização de dados. P1 disse que, talvez, com mais tempo, pudesse analisar e verificar se um ou outro componente não se sobrepunham, podendo ser, talvez, unificados. Mas não disse quais seriam esses componentes.

Durante a realização da tarefa, percebemos que os usuários refletiam bastante sobre as decisões que iriam tomar mas, o que nos chamou a atenção foi que, no momento de refletir sobre os elementos visuais e elementos estruturais, eles revisitaram decisões anteriores para repensar e analisar, por exemplo, quais técnicas de visualização melhor se adequariam aos elementos que desejavam inserir na visualização narrativa interativa que estavam projetando. No final do seu *design*, por exemplo, P3 se questionou e refletiu sobre se teria criado uma visualização narrativa no modelo *Martini glass* ou se era mesmo um *Interactive slideshow*. Concluiu que pensou na estrutura de slides e constatou que cada slide estava funcionando como um *Martini glass*, que é o que os autores dos modelos [SH10] disseram que costumava acontecer mesmo. Assim, notamos que eles consideraram que os elementos estruturais e visuais precisam ser pensados com mais cuidado e atenção.

5.5 Análise do CIViS

De uma forma geral, os participantes não encontraram dificuldades para realização da tarefa de criar o *design* de uma visualização narrativa interativa com o CIViS. Todos consideraram

que o fluxo de passos para o *design* foi adequado para esta finalidade e não tiveram dificuldade em compreender os componentes do CIViS. Sobre o material de apoio (tutorial e fluxo de passos), os participantes relataram que não tiveram dificuldades em entendê-los e que eles ajudaram bastante durante o *design*. P3 considerou o tutorial longo mas, disse que, mesmo assim, o ajudou. Segundo o participante, o CIViS *“lembrou o modelo de mapa mental, que ajuda a esclarecer aquilo que você está pesquisando e investigando, e, no que se refere a narrativas, o modelo ajuda muito a pensar na história que você pode e quer contar”*.

Considerando que este foi o primeiro contato dos participantes com o CIViS, esperávamos que eles tivessem dúvidas e dificuldades na sua compreensão. As dúvidas que ocorreram foram em relação à correta compreensão do significado dos componentes, mas os resultados mostram que os participantes não tiveram dificuldades no uso do modelo para a criação de suas visualizações narrativas. De fato, segundo suas próprias avaliações, eles conseguiram representar com sucesso as decisões de *design* de visualização narrativa interativa para o cenário de dados extraídos do Twitter.

Os participantes relataram que, em suas atividades diárias, costumam criar uma visualização e verificar se esta é satisfatória ou não para representar os dados. Assim, essa experiência que o CIViS proporciona, priorizando a reflexão antes da ação, permitindo planejar o que é possível fazer com o dados e qual história desejam contar com esses dados, é não somente nova mas, também, agrega valor às suas atividades e, concluem, à experiência do leitor de suas visualizações. Para P1, o CIViS *“ajudou a refletir sobre o uso de cada componente de forma visual, pois a ideia de usar cartas para representar os componentes foi positiva. É um modelo bem interativo!”*. P2 disse que *“quando estamos dentro de uma narrativa, estamos imersos, né? Na minha opinião isso é o oposto do que é feito hoje. O leitor hoje que não conhece tão bem sobre visualização de dados, ele pode passar um pouco batido sobre aquela visualização. Agora, quando tem uma narrativa, vai ajudar o leitor a entender... é até um trabalho educativo. E isso o CIViS parece proporcionar.”* P3 disse que *“esse modelo me ajudou a refletir sobre como contar a história dos dados, como vai ser a narrativa, antes de você simplesmente montar a narrativa na sua mente. Ele faz pensar assim: eu tenho todos esses dados e eu posso contar a história deles. Aí você começa a analisar de onde eles vêm, para onde vão, quais são os grupos, quais são os atores... e acho que leva um tempo para você transformar esse monte de dados em uma narrativa, e o modelo ajuda a organizar esse pensamento”*.

P4 salienta que, sobre o uso do CIViS, nunca havia pensado na possibilidade de ter uma narrativa na visualização. Como disse: *“o que sempre me incomodou foi que as visualizações que as pessoas usam são assim: essa é desse jeito, essa outra é assim, e essa outra desse outro jeito.”* Para ela, na verdade, as visualizações são usadas como um instrumento de informação dentro de uma narrativa textual, e aqui, com esse modelo, é possível criar uma outra forma de narrativa, por meio de visualização, *“de uma forma que eu nunca havia pensado ainda. Eu sentia falta de uma possibilidade dessas, mas nunca tinha entendido que era possível. Porque aqui os dados entram como componentes da narrativa e isso é importante!”* (P4).

Quando questionados se acreditavam que uma visualização narrativa interativa poderia estimular a exploração de dados pelo usuário final, os participantes disseram que sim, que esse tipo de visualização tem um grande potencial para isso, principalmente pela sua característica de interatividade e navegação. P1 acredita que a visualização narrativa tem potencial para estimular a exploração de dados no usuário/leitor. Segundo ele, *“qualquer lugar que você puder facilitar a complexificação da relação e dar um empoderamento para os dois lados da relação, você melhora a narrativa”*. No entanto, a visualização narrativa interativa precisa ser planejada para atingir esse propósito. Isso pode ser ilustrado pela fala de P3:

“É muito comum você ter um monte de dados e acabar ignorando algumas potencialidades deles. Já ouvi alguém dizer que essa coisa de big data parece ser um mar de respostas esperando uma pergunta. Realmente, não é fácil pegar um grande volume de dados e transformar em informações úteis e interessantes. Assim, eu acho que se a visualização narrativa for bem feita, ela ajudará a criar a sua pergunta, a contar histórias, e, com isso, o usuário pode querer explorar mais os dados”.

P2 achou o CIViS *“fácil, interessante e que pode auxiliar bastante principalmente quem não é da área técnica, e ajuda a gente a tomar uma decisões que são bem difíceis de tomar quando se está usando uma ferramenta. E eu gosto demais da questão de usar a narrativa, de pensar em uma narrativa, justamente pra envolver melhor o leitor e criar o ambiente de imersão que eu disse.”* P2 destacou que, ao refletir sobre as decisões a serem tomadas, o CIViS facilita que o analista pense nos aspectos que envolvem contar uma história com dados.

Já P4 entende que *“ao descrever uma cena, sou eu olhando pra ela. Mas, como hoje as redes sociais promovem uma dinâmica de interação, eu gostaria que o outro também olhasse para a cena ao mesmo tempo que ele estivesse lendo a minha descrição e pudesse discutir, comentar e interagir a partir disso. E com esses componentes do CIViS, parece que eu posso ter essa possibilidade”*.

Em geral, os participantes não encontraram pontos negativos na criação de um *design* de visualização narrativa interativa usando o CIViS. Apenas P3 que considera que, quando se tem um modelo muito definido, pode acontecer de você se fechar somente àquelas possibilidades. Contudo, a responsabilidade sobre a definição do modelo recai sobre o *designer*, criador da interface customizável usada pelo analista de dados em suas customizações. Isso nos leva a refletir, portanto, que essas limitações consideradas por P3 acontecem em decorrência de uma falha na compreensão das necessidades dos usuários desta interface, os analistas de dados. Com isso, obtivemos apenas relatos positivos da experiência com o CIViS, em que os participantes citaram a possibilidade que o modelo oferece de entender os dados e ajudar a pensar na narrativa a partir deles. Para eles, isso ajuda a criar uma narrativa que responda as perguntas que os analistas de dados de mídias sociais possuem.

Por fim, após a condução deste estudo, percebemos que o CIViS proporcionou aos participantes, que foram relacionados às *personas* identificadas anteriormente, uma experiência de

refletir sobre as decisões de *design* para a criação da visualização narrativa interativa de acordo o cenário e tarefa utilizados. Com isso, podemos considerar que, independente do nível de domínio de ferramentas computacionais pelos analistas de dados de mídias sociais (que foi um ponto importante considerado na criação das *personas*), o CIViS apoiou satisfatoriamente o processo de customização de visualizações narrativas interativas pelos analistas de dados *co-designers*.

5.6 Respondendo às questões de pesquisa

A proposta deste modelo baseou-se em uma necessidade real de um grupo de pesquisadores que analisam dados de mídias sociais, por parcerias com profissionais da Ciência da Computação, para transformar os dados brutos em representações visuais dinâmicas daquilo que estava analisando. Assim, a partir de reuniões e discussões sobre mídias sociais, análise de dados extraídos dessas mídias, discussões sobre as relações de poder nelas estabelecidas, entre outros assuntos neste contexto, surgiu a ideia de criar um mecanismo de “interface” (podendo ser uma pessoa, por exemplo) que conseguisse compreender as necessidades dos analistas de dados e modelá-las para que os profissionais da computação as efetivassem. A partir dessa situação, muitas foram as reflexões e discussões de como apoiar os analistas em suas demandas. A figura 5.8 representa o que foi pensado inicialmente como o modelo e que, por meio deste trabalho, propomos como *design* de visualização com co-autoria.

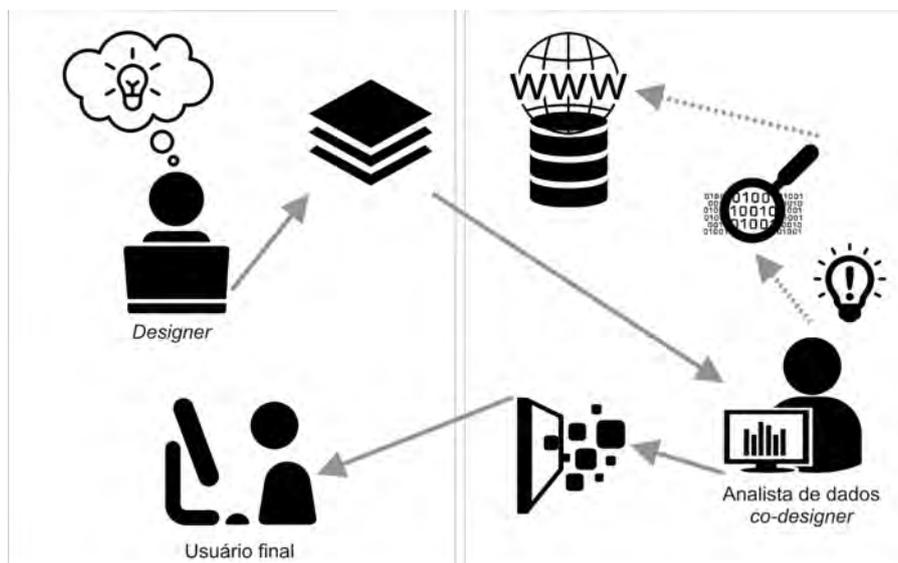


Figura 5.8 – Ideia inicial da proposta do modelo.

Seguimos, agora, com as respostas às questões de pesquisa apresentadas no início deste estudo. Com base na análise de uso do CIViS pelos analistas de dados *co-designers*, verificamos que seus componentes foram considerados úteis para customização de uma visualização narrativa interativa (**RQ01**), constituindo-se em elementos básicos para representar graficamente os dados, com recursos de texto, interação, ordenação, elementos de narrativa, entre outros, que, juntos,

possibilitam a criação de histórias com os dados. Durante a atividade, percebemos que os participantes refletiram sobre a questão de planejar a narrativa, decidir sobre como iriam começar a história e quais componentes e elementos iriam utilizar. Essa reflexão aconteceu de forma iterativa, pois antes de inserir novos elementos na narrativa, consultavam suas anotações sobre as escolhas anteriores. Eles também se preocuparam em garantir o uso de componentes de acordo com o proposto no modelo de narrativa escolhido, diferindo entre os “momentos” em que a narrativa seria guiada pelo autor e aqueles em que seria guiada pelo leitor. Talvez pela sua área de formação e de atuação, os participantes não encontraram dificuldades na compreensão dos componentes e não aconteceu de atribuírem significados que não correspondiam ao propósito do componente.

Ao final do estudo de uso do CIViS, 3 dos 4 participantes se questionaram sobre “como não pensaram sobre isso antes”, em relação a contar histórias com dados por meio de visualizações. Todos fizeram associações aos trabalhos de análises por eles já feitos e que poderiam ter usufruído desse recurso para atingir com mais propriedade os seus objetivos. Desta forma, as impressões que os analistas de dados tiveram sobre a reflexão que o CIViS proporciona (**RQ02**) estão relacionadas às possibilidades de representar aquilo que demandavam (como levantado no Capítulo 3, de estudos iniciais). Eles apontaram as dificuldades que possuem em representar determinados contextos dos dados que analisam e acreditam que o CIViS pode ajudar a criar essa representação e superar as dificuldades.

Como resultado das observações e análise do uso do CIViS, percebemos que as cartas do componente de Técnicas de visualização podem ser ampliadas, contemplando uma variedade maior de técnicas, como é o caso da nuvem de palavras, que foi pensada por eles para compor o *design* das narrativas, mas, no entanto, o modelo não oferecia. Consideramos esta uma importante reflexão para a *designer* da visualização narrativa customizável oferecida nesta tese (Persona 04, de Carol Martins, aqui instanciada nesta pesquisadora), principalmente em se considerando um futuro estudo do uso do CIViS por *designers* de interação (outras instâncias da Persona 04), com foco na análise de suas decisões de *design*.

As respostas às questões de pesquisa nos levam a refletir sobre a evolução da prova de conceito e suas implicações. A partir da análise realizada, percebemos que estamos no caminho de conseguir apoiar *designers* e analistas de dados *co-designers* em suas reflexões e decisões de *design* de visualizações narrativas interativas de dados extraídos de mídias sociais. No entanto, para evoluir esse modelo, precisamos pensar em sua automatização e inserção no processo de visualização narrativa interativa customizável. Precisaremos definir um processo específico para esse contexto e as atividades que o compõem. Neste sentido, o CIViS se propõe a ser uma ferramenta para apoiar o projeto desse tipo de visualização por meio de uma ferramenta de software (como uma ferramenta CASE, por exemplo).

6. CONSIDERAÇÕES FINAIS

A análise de dados de mídias sociais vem crescendo fortemente nas últimas décadas devido à possibilidade de capturar essa grande quantidade de dados, das mais variadas fontes, de forma rápida e fácil. O olhar humano não consegue acompanhar tudo o que é dito e publicado nessas mídias e, por isso, analistas dependem cada vez mais das ferramentas computacionais para coletar, limpar, tratar e apresentar esses dados. Assim, com o advento dessa “nova” área, o processo de análise de dados de mídias sociais pode ser definido em estágios como: definir quais dados se deseja coletar, como será feita a coleta, como esses dados serão processados e, também, como serão apresentados.

Segundo Zeng et al. [ZCLL10], do ponto de vista da Ciência da Computação, a pesquisa em análise de mídias sociais ainda enfrenta vários desafios:

- as mídias sociais contém um conjunto rico de dados ou metadados, que não são tratados de forma sistemática na literatura de mineração de dados e de textos, na aprendizagem de máquina ou no processamento de linguagem natural. Especificamente na área de processamento de linguagem natural, o tratamento de textos publicados em mídias sociais, com a linguagem usada neste meio, como o “internetês”, ainda não obtém bons resultados, nesse aspecto. Por exemplo *hashtags*, opiniões, sarcasmo, ideias, avaliações e perspectivas subjetivas expressadas pelo usuário *online*, classificações, perfis de usuário, e as redes sociais explícitas e implícitas;
- as aplicações de mídias sociais são um exemplo proeminente de computação centrada no ser humano com ênfase única nas interações sociais entre os usuários. Portanto, questões como perfil de usuário dependente do contexto e necessidades de elicitación, bem como vários tipos de considerações de interações homem-computador devem ser reexaminadas;
- embora as mídias sociais prometam uma nova abordagem para o problema do ruído e da sobrecarga de informação com o processamento de informações baseado na Web, questões como inconsistência semântica, evidências conflitantes, falta de estrutura, imprecisões e dificuldades na integração de diferentes tipos de sinais existem, ainda, em abundância nas mídias sociais; e
- os dados das mídias sociais são fluxos dinâmicos, com aumento muito rápido. A natureza de tais dados e seu tamanho representam desafios significativos para a computação em geral.

Por meio dos estudos realizados nesta tese, pudemos observar que os analistas de dados de mídias sociais reconhecem algumas dessas limitações da Ciência da Computação mas, de qualquer forma, os seus desafios são relacionados a encontrar parceiros da Ciência da Computação para colaborar durante o processo de análise, na medida em que surgem demandas como a que foi muito discutida por eles nos estudos realizados de “representar a dinamicidade das redes ao

longo do tempo”, além do desafio de terem que aprender a usar ferramentas e adquirirem outras habilidades que facilitariam o trabalho envolvido nesse processo. Assim, pesquisadores de dados de mídias sociais enfrentam barreiras metodológicas e técnicas e questões sobre como a pesquisa com dados sociais online devem ser feitas, garantindo, por exemplo, validade, ética e reprodutibilidade.

De acordo com Heer e Shneiderman [HS12], o processo de análise requer um julgamento contextualizado quanto ao significado específico do domínio. Para os autores, a análise significativa consiste em explorações repetidas (iterativas) até que os usuários desenvolvam *insights* sobre relacionamentos significativos, influências contextuais específicas do domínio e padrões causais. “*Como eles organizam suas informações para análise? Quais ferramentas computacionais eles aplicam? Como eles colaboram com os outros? Quais são seus produtos de análise?*” Essas são perguntas que Chin Jr. et al. [CJKW09] apresentaram para instigar a pesquisa, o desenvolvimento e a implantação de tecnologias de informação para apoiar a análise. Nesta tese, alinhados com outros autores [DNKS10, DDCN12, Bro15, Zha17], nos concentramos nas práticas de análise de dados das mídias sociais.

De qualquer forma, o *design* de ferramentas de análise de dados é um desafio interdisciplinar que requer compreensão do domínio de análise relacionado. Com isso, nos colocamos o desafio de ingressar em uma área em que pessoas leigas em computação estão realizando análise de dados extraídos de mídias sociais e, nem sempre, possuem habilidades nas ferramentas computacionais existentes para esse fim. Realizamos estudos para compreender o universo da análise de dados de mídias sociais, principalmente as que são feitas por profissionais leigos em computação. Nosso foco esteve em entender como a visualização de dados se fazia presente no processo de análise de mídias sociais.

A questão de pesquisa que buscamos responder nesta tese foi: *Como apoiar analistas de dados na customização de visualizações narrativas de dados extraídos de mídias sociais?* Para isso, realizamos estudos para entender o contexto do problema, compreender o estado da arte da pesquisa sobre visualizações narrativas, conhecer as práticas e as demandas dos analistas de dados de mídias sociais e também entender o quanto a visualização de dados está presente nas atividades desses profissionais. Além disso, compreender o que poderia ser realizado, nesse aspecto, para apoiar o processo de análise de dados. Alinhado à nossa questão de pesquisa, o objetivo principal foi ajudar analistas de dados de mídias sociais a compreender os dados que possuem e a projetar visualizações narrativas interativas para apresentar os dados para sua audiência.

Para isso, realizamos 4 estudos para: *i)* compreender como as visualizações podem auxiliar profissionais que trabalham com mídias sociais em suas atividades diárias (estudos 1 e 2); *ii)* obter uma visão do estado da arte da pesquisa nessa área e descobrir quais técnicas, elementos e processos poderiam ser considerados na construção de visualizações narrativas, por meio de um mapeamento sistemático da literatura; e *iii)* compreender melhor as atividades e demandas dos analistas de dados de mídias sociais, bem como o processo de análise, por meio de entrevistas semi-estruturadas. Os resultados desses estudos nos permitiram identificar aspectos importantes

relacionados ao uso de visualizações, aos perfis dos analistas de dados de mídias sociais (dos quais criamos as *personas*), aos desafios e às dificuldades existentes no processo de análise de dados, e às demandas existentes em relação a visualização de dados.

Além disso, aprofundamos os nossos estudos acerca de todo o contexto envolvido nesta pesquisa e, por meio deles, observamos a potencialidade do uso de visualização narrativa de dados e a necessidade de discussão por parte da comunidade de pesquisa de visualização como um dos caminhos para desenvolver esforços para melhor compreensão e caracterização deste assunto. Isso foi uma das motivações para a solução apresentada nesta tese. Para isso, realizamos estudos para entender o contexto do problema, compreender o estado da arte da pesquisa sobre visualizações narrativas, conhecer as práticas e as demandas dos analistas de dados de mídias sociais e também entender o quanto a visualização de dados está presente nas atividades desses profissionais. Além disso, compreender o que poderia ser desenvolvido nesse aspecto para apoiar o processo de análise de dados.

Nossa proposta foi elaborar um modelo (ferramenta epistêmica) que visa fornecer ao *designer* um melhor entendimento sobre o *design* de visualizações narrativas interativas customizáveis, fundamentado na Teoria da Engenharia Semiótica e sua visão sobre EUD. A ideia é que o modelo ajude: *i)* o *designer* a elaborar seu projeto de comunicação com os usuários analistas de dados *co-designers* e nas experiências que eles terão ao utilizar o modelo; e *ii)* o analista de dados *co-designer* a refletir sobre as decisões de customização da visualização narrativa interativa que irá projetar para o seu usuário final. Com este objetivo em mente, elaboramos o CIViS - modelo de *design* de interface customizável para apoiar a construção de visualizações narrativas de dados extraídos de mídias sociais. Neste modelo, abordamos a EUD na perspectiva de analista de dados de mídias sociais atuando como *designers*.

Nossos resultados mostram que o modelo se mostrou capaz de expressar decisões de *design* referentes à customização de visualizações narrativas, em nível de *design*. Criamos uma prova de conceito do CIViS e realizamos um estudo de seu uso por analistas de dados *co-designers*, que nos mostrou que o CIViS pode ser utilizado como uma ferramenta analítica para projetar visualizações narrativas interativas. Isso demonstra a contribuição do nosso trabalho para a pesquisa em visualizações narrativas, ao propor um modelo que apoia o *design* desse tipo de visualização com foco em EUD, além de apresentar aos analistas de dados a possibilidade de se pensar visualizações de dados que contam histórias de dados. Ao longo da nossa pesquisa, trabalhamos em parceria com comunicadores sociais, jornalistas, historiadores e profissionais que fazem análise de dados de mídias sociais, dado o contexto que serviu de inspiração para este trabalho, descrito anteriormente na seção 5.6. No entanto, a partir dos resultados e análises desta pesquisa, percebemos que o modelo não possui elementos específicos para a análise de mídias sociais e, por isso, acreditamos que o CIViS pode não ser restritivo ao contexto de mídias sociais e poderia, também, ser aplicado a outros contextos de análise de grandes volumes de dados.

As limitações desta pesquisa estão relacionadas ao fato de realizarmos a análise de uso do CIViS apenas com usuários analistas de dados *co-designers* e não termos realizado com os

designers. Além disso, desejávamos ter realizado o estudo com todos os participantes das entrevistas (estudo 4), o que foi inviável realizar pessoalmente devido ao fato de que estão em cidades e Estados diferentes e isso dificultava a sincronização de tempo e o deslocamento.

O trabalho desenvolvido nesta tese trouxe contribuições em diferentes áreas. Apresentamos as contribuições na seção 6.1 e, em seguida, os trabalhos futuros, na seção 6.2.

6.1 Contribuições

A nossa principal contribuição foi a criação do modelo descritivo, baseado na Engenharia Semiótica, para apoiar *designers* e *co-designers* na criação de visualizações narrativas interativas customizáveis. Para a área de IHC, especificamente para a Engenharia Semiótica, nossa pesquisa contribui com a proposição de uma nova ferramenta epistêmica para aplicação desta teoria, envolvendo EUD para usuários analistas de dados de mídias sociais que atuarão como co-autores das visualizações. Ressaltamos que a nossa proposta contribui também com uma forma de auxiliar o usuário no aprendizado da customização por meio do tutorial elaborado como sistema de ajuda. Neste trabalho, pudemos perceber como a Engenharia Semiótica apoiou a criação de um modelo que facilitou a comunicação entre *designer* e *co-designers* por meio da compreensão de quem são os potenciais usuários EUD desse modelo, suas necessidades e suas demandas. Além disso, a compreensão desse contexto e a criação do modelo nos permitiram perceber que o modelo se aproximou de muitas necessidades apontadas pelos participantes de nossos estudos e que, por meio da análise do estudo de uso do CIViS, pudemos compreender o quanto ele foi ao encontro dessas necessidades e o quanto ele pode ser adequado para atender os analistas de dados *co-designers*.

Esta tese trouxe, também, uma contribuição para os analistas de dados de mídias sociais no que diz respeito a apresentar-lhes o conceito de visualização narrativa e a possibilidade de se contar histórias com dados, com o objetivo de instigar o engajamento da sua audiência na exploração dos dados. Além disso, alinhados com Boy et al [BRBF14], recomendamos, não somente aos analistas, mas aos usuários de mídias digitais que desenvolvam a literacia digital, ou melhor, literacia em visualização. A literacia em visualização está relacionada com as habilidades de extrair e manipular informação a partir de representações gráficas [BRBF14].

Por fim, não podemos deixar de destacar as contribuições para o fortalecimento e a importância da pesquisa qualitativa na Ciência da Computação, que nos permitiu, nesta tese, nos aproximar das áreas das Ciências Sociais e Humanas e melhor compreender seus contextos e suas demandas em relação às ferramentas computacionais.

6.2 Trabalhos Futuros

Os trabalhos futuros desta pesquisa estão relacionados com as limitações apresentadas, no sentido de refinar a ferramenta epistêmica para apoiar outros tipos de problemas, e também com novas ações que visam acompanhar o uso deste modelo, observando-o em relação ao quanto possibilita reflexões sobre decisões de criação de visualizações narrativas interativas.

Além disso, vislumbramos a melhoria e continuidade desta pesquisa, apontando algumas possibilidades como a realização de uma avaliação do uso do CIViS por *designers* de interação, bem como do uso do modelo por analistas de dados (menos experientes em análise de dados), para consolidar nossos resultados.

Alinhados com a possibilidade de o CIViS ser utilizados em outros contextos, pretendemos realizar avaliações de seu uso em contextos distintos de análise de grandes volumes de dados. Acreditamos que essas avaliações permitirão reflexões sobre possibilidades de melhoria do espaço de *design* do modelo.

Considerando a importância do sistema de ajuda em quaisquer sistemas, como trabalho futuro também pretendemos avaliar como o sistema de ajuda, da forma que está proposto, auxilia no uso do CIViS em diferentes contextos de análise de dados.

Outra possibilidade de próximos trabalhos é modelar uma arquitetura para uma ferramenta de visualização narrativa customizável com base no CIViS e, em seguida, implementá-la.

Por fim, pretendemos promover uma maior aproximação entre a Ciência da Computação e Ciências Sociais e Humanas para planejar e conduzir estudos que possam evoluir esta pesquisa, principalmente no sentido de identificar as possibilidades de interface humana e tecnológica que facilite a realização do processo de análise de dados de mídias sociais.

REFERÊNCIAS BIBLIOGRÁFICAS

- [Alb08] Albergaria, E. T. “Um modelo de interface extensível para sistemas de mineração de dados por regras de associação”, Dissertação de mestrado, Departamento de Ciência da Computação, UFMG, Belo Horizonte, Brasil, 2008, 114p.
- [ALW11] Aylett, R.; Louchart, S.; Weallans, A. “Research in interactive drama environments, role-play and story-telling”. In: Proceedings of the International Conference on Interactive Digital Storytelling (ICIDS), 2011, pp. 1–12.
- [AMPMJ08] Albergaria, E.; Mourão, F.; Prates, R.; Meira Jr, W. “Modelo de interface extensível como solução para desafios de interação em sistemas de mineração de dados”. In: Anais do Congresso da SBC – Seminário Integrado de Software e Hardware, Sociedade Brasileira de Computação (SEMISH), 2008, pp. 151–165.
- [AO05] Arksey, H.; O’Malley, L. “Scoping studies: towards a methodological framework”, *International journal of social research methodology*, vol. 8–1, 2005, pp. 19–32.
- [Bar99] Barbosa, S. D. J. “Programação via interface”, Tese de doutorado, Departamento de Informática, PUC-Rio, Rio de Janeiro, Brasil, 1999, 118p.
- [Bar06] Barbosa, C. M. d. A. “Manas-uma ferramenta epistêmica de apoio ao projeto da comunicação em sistemas colaborativos”, Tese de doutorado, Departamento de Informática, PUC-Rio, Rio de Janeiro, Brasil, 2006, 222p.
- [BBC⁺11] Bertschi, S.; Bresciani, S.; Crawford, T.; Goebel, R.; Kienreich, W.; Lindner, M.; Sabol, V.; Moere, A. “What is knowledge visualization? perspectives on an emerging discipline”. In: Proceedings of the International Conference on Information Visualisation (IV), 2011, pp. 329–336.
- [BCC15] Baxter, K.; Courage, C.; Caine, K. “Understanding your users: a practical guide to user research methods”. Morgan Kaufmann, 2015, 568p.
- [BDF15] Boy, J.; Detienne, F.; Fekete, J.-D. “Storytelling in information visualizations: Does it engage users to explore data?” In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), 2015, pp. 1449–1458.
- [BdP03] Barbosa, S. D. J.; de Paula, M. G. “Designing and evaluating interaction as conversation: A modeling language based on semiotic engineering”. In: Proceedings of the International Workshop on Design, Specification, and Verification of Interactive Systems (DSV-IS), 2003, pp. 16–33.
- [BdSP12] Baranauskas, M. C. C.; de Souza, C. S.; Pereira, R. “Prospecção de grandes desafios de pesquisa em interação humano-computador no brasil (grandihc-br)”.

In: Proceedings of the 11th Brazilian Symposium on Human Factors in Computing Systems (IHC), 2012, pp. 63–64.

- [Ben15] Benevenuto, F. “Computação social: Desafios para cooperações interdisciplinares”. In: *Análise de Redes para Mídia Social*, Editora Sulina, 2015.
- [BLHHW06] Berners-Lee, T.; Hall, W.; Hendler, J.; Weitzner, D. J. “Creating a science of the web”, *Science*, vol. 313–5788, 2006, pp. 769–771.
- [BPdS05] Barbosa, C. M. d. A.; Prates, R. O.; de Souza, C. S. “Marq-g: A semiotic engineering approach for supporting the design of multi-user applications”. In: Proceedings of the Latin American Conference on Human-computer Interaction (CLIHC), 2005, pp. 128–138.
- [BRBF14] Boy, J.; Rensink, R. A.; Bertini, E.; Fekete, J. D. “A principled way of assessing visualization literacy”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 20–12, 2014, pp. 1963–1972.
- [Bro15] Brooks, M. “Human centered tools for analyzing online social data”, Phd thesis, University of Washington, Seattle, WA, 2015, 222p.
- [CJKW09] Chin Jr., G.; Kuchar, O. A.; Wolf, K. E. “Exploring the analytical processes of intelligence analysts”. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), 2009, pp. 11–20.
- [CM11] Cruz, P.; Machado, P. “Generative storytelling for information visualization”, *IEEE Computer Graphics and Applications*, vol. 31–2, 2011, pp. 80–85.
- [Cre14] Creswell, J. W. “Research design: Qualitative, quantitative, and mixed methods approaches”. Sage, 2014, 304p.
- [DDCN12] Diakopoulos, N.; De Choudhury, M.; Naaman, M. “Finding and assessing social media information sources in the context of journalism”. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), 2012, pp. 2451–2460.
- [Dia10] Diakopoulos, N. “Game-y information graphics”. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), 2010, pp. 3595–3600.
- [DL13] Denzin, N. K.; Lincoln, Y. S. “Collecting and interpreting qualitative materials”. Sage, 2013, 656p.
- [DNKS10] Diakopoulos, N.; Naaman, M.; Kivran-Swaine, F. “Diamonds in the rough: Social media visual analytics for journalistic inquiry”. In: Proceedings of the IEEE Symposium on Visual Analytics Science and Technology (VAST), 2010, pp. 115–122.

- [dSB14a] de Sousa, T. A. F.; Barbosa, S. D. J. “Recommender system to support chart constructions with statistical data”. In: Proceedings of the International Conference on Human-Computer Interaction (HCII), 2014, pp. 631–642.
- [dSB14b] de Souza, L. G.; Barbosa, S. D. J. “Estendendo a MoLIC para apoiar o design de sistemas colaborativos”. In: Proceedings of the Brazilian Symposium on Human Factors in Computing Systems (IHC), 2014, pp. 25–28.
- [dSBdS01] de Souza, C. S.; Barbosa, S. D. J.; da Silva, S. R. P. “Semiotic engineering principles for evaluating end-user programming environments”, *Interacting with Computers*, vol. 13–4, 2001, pp. 467–495.
- [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.
- [EB11] Elias, M.; Bezerianos, A. “Exploration views: Understanding dashboard creation and customization for visualization novices”. In: Proceedings of the IFIP TC.13 International Conference on Human-Computer Interaction (INTERACT), 2011, pp. 274–291.
- [EKHW07] Eccles, R.; Kapler, T.; Harper, R.; Wright, W. “Stories in geotime”. In: Proceedings of the IEEE Symposium on Visual Analytics Science and Technology (VAST), 2007, pp. 19–26.
- [ES08] Eagan, J. R.; Stasko, J. T. “The buzz: Supporting user tailorability in awareness applications”. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), 2008, pp. 1729–1738.
- [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*, vol. 8–2, 2001, pp. 143–158.
- [FDCD12] Fisher, D.; DeLine, R.; Czerwinski, M.; Drucker, S. “Interactions with big data analytics”, *interactions*, vol. 19–3, 2012, pp. 50–59.
- [FG14] Fan, W.; Gordon, M. D. “The power of social media analytics”, *Communications of the ACM*, vol. 57–6, 2014, pp. 74–81.
- [FGY⁺04] Fischer, G.; Giaccardi, E.; Ye, Y.; Sutcliffe, A. G.; Mehandjiev, N. “Meta-design: A manifesto for end-user development”, *Communications of the ACM*, vol. 47–9, 2004, pp. 33–37.
- [Fig13] Figueiras, A. “A typology for data visualization on the web”. In: Proceedings of the International Conference on Information Visualisation (IV), 2013, pp. 351–358.

- [Fig14a] Figueiras, A. “How to tell stories using visualization”. In: Proceedings of the International Conference on Information Visualisation (IV), 2014, pp. 18–26.
- [Fig14b] Figueiras, A. “Narrative visualization: A case study of how to incorporate narrative elements in existing visualizations”. In: Proceedings of the International Conference on Information Visualisation (IV), 2014, pp. 46–52.
- [Fis07] Fischer, G. “Meta-design: Expanding boundaries and redistributing control in design”. In: Proceedings of the IFIP TC.13 International Conference on Human-Computer Interaction (INTERACT), 2007, pp. 193–206.
- [Gan04] Gancho, C. V. “Como analisar narrativas”. Editora Ática, 2004, 79p.
- [Ghi17] Ghidini, E. “Representação de narrativas interativas por meio de visualizações com dados extraídos de redes sociais”, Dissertação de mestrado, Faculdade de Informática, PUCRS, Porto Alegre, Brasil, 2017, 104p.
- [GP01] Gershon, N.; Page, W. “What storytelling can do for information visualization”, *Communications of the ACM*, vol. 44–8, 2001, pp. 31–37.
- [GSMS17] Ghidini, E.; Santos, C. Q.; Manssour, I.; Silveira, M. S. “Analyzing design strategies for narrative visualization”. In: Proceedings of the Brazilian Symposium on Human Factors in Computing Systems (IHC), 2017, pp. 11:1–11:10.
- [HBO10] Heer, J.; Bostock, M.; Ogievetsky, V. “A tour through the visualization zoo”, *Queue*, vol. 8–5, 2010, pp. 20:20–20:30.
- [HS12] Heer, J.; Shneiderman, B. “Interactive dynamics for visual analysis”, *Queue*, vol. 10–2, 2012, pp. 30:30–30:55.
- [HSH⁺08] Hendler, J.; Shadbolt, N.; Hall, W.; Berners-Lee, T.; Weitzner, D. “Web science: An interdisciplinary approach to understanding the web”, *Communications of the ACM*, vol. 51–7, 2008, pp. 60–69.
- [HVW07] Heer, J.; Viégas, F. B.; Wattenberg, M. “Voyagers and voyeurs: Supporting asynchronous collaborative information visualization”. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), 2007, pp. 1029–1038.
- [Jak60] Jakobson, R. “Linguistics and poetics”. In: *Style in language*, MIT Press, 1960.
- [KC07] Kitchenham, B.; Charters, S. “Guidelines for performing systematic literature reviews in software engineering”, Technical report, School of Computer Science and Mathematics, Keele University, 2007, 65p.
- [Kit04] Kitchenham, B. “Procedures for performing systematic reviews”, Technical report, Department of Computer Science, Keele University, 2004, 26p.

- [KM13] Kosara, R.; Mackinlay, J. “Storytelling: The next step for visualization”, *Computer*, vol. 46–5, 2013, pp. 44–50.
- [LFH10] Lazar, J.; Feng, J. H.; Hochheiser, H. “Research methods in human-computer interaction”. John Wiley & Sons, 2010, 560p.
- [LPKW06] Lieberman, H.; Paternò, F.; Klann, M.; Wulf, V. “End-user development: An emerging paradigm”. Springer, 2006, cap. 1, pp. 1–8.
- [LRIC15] Lee, B.; Riche, N.; Isenberg, P.; Carpendale, S. “More than telling a story: Transforming data into visually shared stories”, *Computer Graphics and Applications*, vol. 35–5, 2015, pp. 84–90.
- [Mon15] Monteiro, I. T. “Autoexpressão e engenharia semiótica do usuário-designer”, Tese de doutorado, Departamento de Informática, PUC-Rio, Rio de Janeiro, Brasil, 2015, 312p.
- [Mør97] Mørch, A. “Computers and design in context”. In: *Three Levels of End-user Tailoring: Customization, Integration, and Extension*, MIT Press, 1997.
- [Mur13] Murray, S. “Interactive data visualization for the Web”. O’Reilly Media, Inc, 2013, 474p.
- [Nie12] Nielsen, L. “Personas-user focused design”. Springer Science & Business Media, 2012, 158p.
- [PD12] Patil, T. H.; Davenport, D. J. “Data scientist: The sexiest job of the 21st century”, *Harvard business review*, vol. 90–10, 2012, pp. 70–76.
- [Pei77] Peirce, C. S. “Semiótica”. *Perspectiva*, 1977, 336p.
- [Pei92] Peirce, C. S. “The essential Peirce: selected philosophical writings”. Indiana University Press, 1992, vol. 1, 448p.
- [Pei98] Peirce, C. S. “The essential Peirce: selected philosophical writings”. Indiana University Press, 1998, vol. 2, 624p.
- [Pin06] Pinna, D. M. d. S. “Animadas personagens brasileiras: a linguagem visual das personagens do cinema de animação contemporâneo brasileiro”, Dissertação de mestrado, Departamento de Artes Gráficas, PUC-Rio, Rio de Janeiro, Brasil, 2006, 452p.
- [PR08] Petticrew, M.; Roberts, H. “Systematic reviews in the social sciences: A practical guide”. John Wiley & Sons, 2008, 352p.
- [Rus11] Russell, M. A. “Mineração de dados da web social”. O’Reilly, 2011, 360p.
- [San83] Santaella, L. “O que é semiótica”. Brasiliense, 1983, 80p.

- [San00] Santaella, L. “A teoria geral dos signos: como as linguagens significam as coisas”. Pioneira, 2000, 153p.
- [SB13] Sousa, T. A. F.; Barbosa, S. D. J. “Sistema de recomendação para apoiar a construção de gráficos com dados estatísticos”, Dissertação de mestrado, Departamento de Informática, PUC-Rio, Rio de Janeiro, Brasil, 2013, 180p.
- [SCT⁺17] Santos, C. Q.; Cunha, H.; Teixeira, C.; Souza, D. R. d.; Tietzmann, R.; Manssour, I.; Silveira, M.; Träsel, M.; Ruiz, D. D. A.; Barros, R. “Media professionals’ opinions about interactive visualizations of political polarization during brazilian presidential campaigns on twitter”. In: Proceedings of the Hawaii International Conference on System Sciences (HICSS), 2017, pp. 1939–1948.
- [SdSB03] Silveira, M. S.; de Souza, C. S.; Barbosa, S. D. J. “Um método da engenharia semiótica para a construção de sistemas de ajuda online”. In: Proceedings of the Latin American Conference on Human-computer Interaction (CLIHIC), 2003, pp. 167–177.
- [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.
- [Shn96] Shneiderman, B. “The eyes have it: a task by data type taxonomy for information visualizations”. In: Proceedings of the IEEE Symposium on Visual Languages (VAST), 1996, pp. 336–343.
- [Shn07] Shneiderman, B. “Web science: A provocative invitation to computer science”, *Communications of the ACM*, vol. 50–6, 2007, pp. 25–27.
- [Sou05] de Souza, C. S. “The semiotic engineering of human-computer interaction”. MIT press, 2005, 312p.
- [Sou13] de Souza, C. S. “Semiotics”. The Interaction Design Foundation, 2013, cap. 2.
- [Sou17] de Souza, C. S. “Semiotic engineering: a cohering theory to connect eud with hci, cmc and more”. In: *New perspectives in end-user development*, Springer, 2017.
- [SPC⁺16] Shneiderman, B.; Plaisant, C.; Cohen, M.; Jacobs, S.; Elmqvist, N.; Diakopoulos, N. “Grand challenges for HCI researchers”, *interactions*, vol. 23–5, 2016, pp. 24–25.
- [SSM16] Santos, C. Q.; Silveira, M. S.; Manssour, I. H. “Visualization and social media data analysis: Preliminary studies about data analysts’ perception”. In: Proceedings of the Brazilian Symposium on Human Factors in Computer Systems (IHC), 2016, pp. 35:1–35:4.

- [STT⁺16] Santos, C. Q.; Tietzmann, R.; Träsel, M.; Moraes, S. M. W.; Manssour, I. H.; Silveira, M. S. “Can visualization techniques help journalists to deepen analysis of twitter data? exploring the “germany 7 x 1 brazil” case”. In: Proceedings of the Hawaii International Conference on System Sciences (HICSS), 2016, pp. 1939–1948.
- [VBN⁺04] Viegas, F.; Boyd, D.; Nguyen, D.; Potter, J.; Donath, J. “Digital artifacts for remembering and storytelling: posthistory and social network fragments”. In: Proceedings of the Annual Hawaii International Conference on System Science (HICSS), 2004, pp. 1–10.
- [Vil16] Villela, M. L. B. “Um modelo de design de privacidade para o compartilhamento de informações pessoais em redes sociais online”, Tese de doutorado, Departamento de Ciência da Computação, UFMG, Belo Horizonte, Brasil, 2016, 262p.
- [War12] Ware, C. “Information visualization: perception for design”. Elsevier, 2012, 512p.
- [WGK15] Ward, M. O.; Grinstein, G.; Keim, D. “Interactive data visualization: foundations, techniques, and applications”. CRC Press, 2015, 578p.
- [WH07] Wong, J.; Hong, J. I. “Making mashups with marmite: Towards end-user programming for the web”. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), 2007, pp. 1435–1444.
- [YHYO01] Yoshioka, T.; Herman, G.; Yates, J.; Orlikowski, W. “Genre taxonomy: A knowledge repository of communicative actions”, *ACM Transactions on Information Systems (TOIS)*, vol. 19–4, 2001, pp. 431–456.
- [ZCLL10] Zeng, D.; Chen, H.; Lusch, R.; Li, S. H. “Social media analytics and intelligence”, *IEEE Intelligent Systems*, vol. 25–6, 2010, pp. 13–16.
- [Zha17] Zhang, J. “Applying visualization in digital journalism: An empirical study on the new york times”, Phd thesis, University of Alberta, Edmonton, Canada, 2017, 102p.

APÊNDICE A – ROTEIRO DO GRUPO FOCAL

Para guiar a condução do grupo focal pelo moderador, definimos o seguinte roteiro:

- Explicar o objetivo do grupo focal, o que se pretende com ele.
- Explicar o contexto da pesquisa e dos dados coletados no Twitter.
- Mostrar as visualizações criadas e, para cada visualização, perguntar:
 - que informação se consegue extrair;
 - se é possível extrair a informação de outra forma;
 - a opinião sobre a visualização;
 - sugestões de modificações e/ou melhorias.
- Das visualizações apresentadas, discutir sobre quais os participantes acham que funcionam melhor ou para quais situações cada uma delas é mais adequada.

APÊNDICE B – LISTA DOS ARTIGOS SELECIONADOS NO MAPEAMENTO SISTEMÁTICO

- S1** GERSHON N., PAGE W.: What storytelling can do for information visualization. *Commun. ACM* 44, 8 (Aug. 2001), 31–37.
- S2** KUCHINSKY A., GRAHAM K., MOH D., ADLER A., BABARIA K., CREECH M. L.: Biological storytelling: A software tool for biological information organization based upon narrative structure. *SIGGROUP Bull.* 23, 2 (Aug. 2002), 4–5.
- S3** ZELLWEGER P. T., MANGEN A., NEWMAN P.: Reading and writing fluid hypertext narratives. In *Proceedings of the Thirteenth ACM Conference on Hypertext and Hypermedia* (New York, NY, USA, 2002), *HYPERTEXT'02*, ACM, pp. 45–54.
- S4** VIEGAS F., BOYD D., NGUYEN D., POTTER J., DONATH J.: Digital artifacts for remembering and storytelling: posthistory and social network fragments. In *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on* (Jan 2004), pp. 10 pp.
- S5** APPAN P., SUNDARAM H., BIRCHFIELD D.: Communicating everyday experiences. In *Proceedings of the 1st ACM Workshop on Story Representation, Mechanism and Context* (New York, NY, USA, 2004), *SRMC'04*, ACM, pp. 17–24.
- S6** MALA T., GEETHA T.: Visualization of biological patterns in event detection and tracking based on srp algorithm. In *Information Visualisation (IV), 2005. Ninth International Conference on* (July 2005), pp. 587–592.
- S7** AKAISHI M., HORIK., SATOH K.: Topic tracer: a visualization tool for quick reference of stories embedded in document set. In *Information Visualization, 2006. IV'06. Tenth International Conference on* (July 2006), pp. 101–106.
- S8** BLUMENKRANTS M., STAROVISKY H., SHAMIR A.: Narrative algorithm visualization. In *Proceedings of the 2006 ACM Symposium on Software Visualization* (New York, NY, USA, 2006), *SoftVis'06*, ACM, pp. 17–26.
- S9** AKAISHI M., SATOH K., KATO Y., HORI K.: Narrative based topic visualization for chronological data. In *Information Visualization, 2007. IV'07. 11th International Conference* (July 2007), pp. 139–144.
- S10** ECCLES R., KAPLER T., HARPER R., WRIGHT W.: Stories in geotime. In *Visual Analytics Science and Technology, 2007. VAST 2007. IEEE Symposium on* (Oct 2007), pp. 19–26.
- S11** ZHU W., CHEN C.: Visual analytics: Storylines: Visual exploration and analysis in latent semantic spaces. *Comput. Graph.* 31, 3 (June 2007), 338–349.

- S12** SUBASIC I., BERENDT B.: Web mining for understanding stories through graph visualisation. In *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on (Dec2008)*, pp. 570–579.
- S13** FARRIMOND B., PRESLAND S., BONAR-LAW J., POGSON F.: Making history happen: Spatiotemporal data visualization for historians. In *Computer Modeling and Simulation, 2008. EMS'08. Second UKSIM European Symposium on (Sept2008)*, pp. 424–429.
- S14** FISHER D., HOFF A., ROBERTSON G., HURST M.: Narratives: A visualization to track narrative events as they develop. In *Visual Analytics Science and Technology, 2008. VAST 2008. IEEE Symposium on (Oct 2008)*, pp. 115–122.
- S15** ROSE S., BUTNER S., COWLEY W., GREGORY M., WALKER J.: Describing story evolution from dynamic information streams. In *Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on (Oct 2009)*, pp. 99–106.
- S16** MITCHELL A., MCGEE K.: Designing hypertext tools to facilitate authoring multiple points-of-view stories. In *Proceedings of the 20th ACM Conference on Hypertext and Hyper-media (New York, NY, USA, 2009), HT'09, ACM*, pp. 309–316.
- S17** SEGEL E., HEER J.: Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics* 16, 6 (Nov. 2010), 1139–1148.
- S18** DIAKOPOULOS N.: Game-y information graphics. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems (New York, NY, USA, 2010), CHI EA'10, ACM*, pp. 3595–3600.
- S19** YU L., LU A., RIBARSKY W., CHEN W.: Automatic animation for time-varying data visualization. *Computer Graphics Forum* 29, 7 (2010), 2271–2280.
- S20** NAIR V., KADUSKAR M., BHASKARAN P., BHAUMIK S., LEE H.: Preserving narratives in electronic health records. In *Bioinformatics and Biomedicine (BIBM), 2011 IEEE International Conference on (Nov 2011)*, pp. 418–421.
- S21** CRUZ P., MACHADO P.: Generative storytelling for information visualization. *Computer Graphics and Applications, IEEE* 31, 2 (March 2011), 80–85.
- S22** HULLMAN J., DIAKOPOULOS N.: Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (Dec.2011), 2231–2240.
- S23** AP CENYDD L., WALKER R., POP S., MILES H., HUGHES C., TEAHAN W., ROBERTS J.: epspread - storyboarding for visual analytics. In *Visual Analytics Science and Technology, 2011. VAST, 2011. IEEE Conference on (Oct 2011)*, pp. 311–312.

- S24** BERTSCHI S., BRESCIANI S., CRAWFORD T., GOEBEL R., KIENREICH W., LINDNER M., SABOL V., MOERE A.: What is knowledge visualization? perspectives on an emerging discipline. In *Information Visualisation (IV)*, 2011. 15th International Conference on (July 2011), pp. 329–336.
- S25** ALLEN R. B.: Visualization, causation, and history. In *Proceedings of the 2011 iConference (New York, NY, USA, 2011)*, iConference'11, ACM, pp. 538–545.
- S26** WEBER W., RALL H.: Data visualization in online journalism and its implications for the production process. In *Information Visualisation (IV)*, 2012. 16th International Conference on (July, 2012), pp. 349–356.
- S27** RIBARSKY W., SAUDA E., WARTELL Z., BALMER J.: The whole story: Building the complete history of a place. In *System Science (HICSS)*, 2012. 45th Hawaii International Conference on (Jan 2012), pp. 1864–1873.
- S28** KUHN A., STOCKER M.: Codetimeline: Storytelling with versioning data. In *Software Engineering (ICSE)*, 2012. 34th International Conference on (June 2012), pp. 1333–1336.
- S29** LUO D., YANG J., KRSTAJIC M., RIBARSKY W., KEIM D.: Eventriver: Visually exploring text collections with temporal references. *Visualization and Computer Graphics*, IEEE Transactions on 18, 1 (Jan 2012), 93–105.
- S30** MA K.-L., LIAO I., FRAZIER J., HAUSER H., KOSTIS H.-N.: Scientific storytelling using visualization. *Computer Graphics and Applications*, IEEE 32, 1 (Jan 2012), 12–19.
- S31** TANAHASHI Y., MA K.-L.: Design considerations for optimizing storyline visualizations. *Visualization and Computer Graphics*, IEEE Transactions on 18, 12 (Dec 2012), 2679–2688.
- S32** WANG X., DOU W., RIBARSKY W., SKAU D., ZHOU M. X.: Leadline: Interactive visual analysis of text data through event identification and exploration. In *Visual Analytics Science and Technology (VAST)*, 2012. IEEE Computer Society, pp. 93–102.
- S33** ITOH M., TOYODA M., KAMIJO T., KITSUREGAWA M.: Visualizing flows of images in social media. In *Visual Analytics Science and Technology (VAST)*, 2012 IEEE Conference on (Oct 2012), pp. 229–230.
- S34** LUPO L., MALIZIA A., DIAZ P., AEDO I.: Socialstory: A social storyboard system for sharing experiences in emergencies. In *Proceedings of the International Working Conference on Advanced Visual Interfaces (New York, NY, USA, 2012)*, AVI'12, ACM, pp. 786–787.
- S35** LOMBARDO V., DAMIANO R., LIETO A., PIZZO A.: Visualization of character's intentions in dramatic media. In *Complex, Intelligent, and Software Intensive Systems (CISIS)*, 2013. Seventh International Conference on (July 2013), pp. 582–587.

- S36** BOGL M., AIGNER W., FILZMOSE P., LAMMARSCH T., MIKSCH S., RIND A.: Visual analytics for model selection in time series analysis. *Visualization and Computer Graphics, IEEE Transactions on* 19, 12 (Dec 2013), 2237–2246.
- S37** HULLMAN J., DRUCKER S., RICHE N., LEE B., FISHER D., ADAR E.: A deeper understanding of sequence in narrative visualization. *Visualization and Computer Graphics, IEEE Transactions on* 19, 12 (Dec 2013), 2406–2415.
- S38** KOSARA R., MACKINLAY J.: Storytelling: The next step for visualization. *Computer* 46, 5 (May 2013), 44–50.
- S39** LEE B., KAZI R. H., SMITH G.: Sketchstory: Telling more engaging stories with data through freeform sketching. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (Dec. 2013), 2416–2425.
- S40** LIU S., WU Y., WEI E., LIU M., LIU Y.: Storyflow: Tracking the evolution of stories. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (Dec. 2013), 2436–2445.
- S41** WALKER R., SLINGSBY A., DYKES J., XU K., WOOD J., NGUYEN P., STEPHENS D., WONG B., ZHENG Y.: An extensible framework for provenance in human terrain visual analytics. *Visualization and Computer Graphics, IEEE Transactions on* 19, 12 (Dec 2013), 2139–2148.
- S42** ZHANG Z., WANG B., AHMED F., RAMAKRISHNAN I., ZHAO R., VICCELLIO A., MUELLER K.: The five ws for information visualization with application to healthcare informatics. *Visualization and Computer Graphics, IEEE Transactions on* 19, 11 (Nov 2013), 1895–1910.
- S43** LUNDBLAD P., JERN M.: Geovisual analytics and storytelling using html5. In *Information Visualisation (IV), 2013 17th International Conference (July 2013)*, pp. 263–271.
- S44** TRONCY R., MILICIC V., RIZZO G., GARCÍA J. L. R.: Mediafinder: Collect, enrich and visualize media memes shared by the crowd. In *Proceedings of the 22Nd International Conference on World Wide Web (New York, NY, USA, 2013), WWW'13 Companion, ACM*, pp. 789–790.
- S45** THIRY E., LINDLEY S., BANKS R., REGAN T.: Authoring personal histories: Exploring the timeline as a framework for meaning making. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2013), CHI'13, ACM*, pp. 1619–1628.
- S46** HULLMAN J., DIAKOPOULOS N., ADAR E.: Contextifier: Automatic generation of annotated stock visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2013), CHI'13, ACM*, pp. 2707–2716.
- S47** PANGER G., REA B., WEBER S.: Visualizing ambivalence: Showing what mixed feelings look like. In *Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2013), CHI EA'13, ACM*, pp. 1029–1034.

- S48** ELIAS M., AUFAURE M.-A., BEZERIANOS A.: *Storytelling in Visual Analytics Tools for Business Intelligence*. In *Proceedings of Human-Computer Interaction – INTERACT 2013*, Cape Town, South Africa, September 2-6, 2013. Springer Berlin Heidelberg, pp. 280–297.
- S49** YOUSUF B., CONLAN O.: *Enhancing learner engagement through personalized visual narratives*. In *Advanced Learning Technologies (ICALT), 2014 IEEE 14th International Conference on* (July 2014), pp. 89–93.
- S50** MINELLI R., BARACCHI L., MOCCI A., LANZA M.: *Visual storytelling of development sessions*. In *Software Maintenance and Evolution (ICSME), 2014 IEEE International Conference on* (Sept 2014), pp. 416–420.
- S51** WALDNER M., LE MUZIC M., BERNHARD M., PURGATHOFER W., VIOLA I.: *Attractive flicker: guiding attention in dynamic narrative visualizations*. *Visualization and Computer Graphics, IEEE Transactions on Visualisation and Computer Graphics*, v. 20, 12 (Dec 2014), 2456–2465.
- S52** FIGUEIRAS A.: *How to tell stories using visualization*. In *Information Visualisation (IV), 2014. 18th International Conference on* (July 2014), pp. 18–18.
- S53** FIGUEIRAS A.: *Narrative visualization: A case study of how to incorporate narrative elements in existing visualizations*. In *Information Visualisation (IV), 2014. 18th International Conference on* (July 2014), pp. 46–52.
- S54** AKYIGIT E., CENGIZ T., YILDIRIM O., BALCISOY S.: *Visual exploratory tool for storyline generation*. In *Visual Analytics Science and Technology (VAST), 2014. IEEE Conference on* (Oct 2014), pp. 215–216.
- S55** BELMONTE N.: *Extracting and visualizing insights from real-time conversations around public presentations*. In *Visual Analytics Science and Technology (VAST), 2014. IEEE Conference on* (Oct 2014), pp. 225–226.
- S56** CHEN N.-C., FELDMAN L., KROLL J., ARAGON C.: *Emoticons and linguistic alignment: How visual analytics can elicit storytelling*. In *Visual Analytics Science and Technology (VAST), 2014. IEEE Conference on* (Oct 2014), pp. 237–238.
- S57** MAIN C. S., PERRONE L. F., SCHROCK G. L.: *Data visualization for network simulations*. In *Proceedings of the 2014 Winter Simulation Conference (Piscataway, NJ, USA, 2014), WSC'14*, IEEE Press, pp. 3119–3130.
- S58** BALDUINI M., DELLA VALLE E., CIUCCARELLI P., AZZI M., LARCHER R., ANTONELLI F.: *Citysensing: Fusing city data for visual storytelling*. *MultiMedia, IEEE* 22, 3 (July 2015), 44–53.
- S59** ANDREWS D., BABER C.: *Visualizing interactive narratives: Employing a branching comic to tell a story and show its readings*. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2014), CHI'14, ACM*, pp. 1895–1904.

- S60** REGATTIERI L., CHARTIER R., WINDSOR J., ROCKWELL G.: Tweetviz: Following twitter hashtags to support storytelling. In HT (Doctoral Consortium/ Late-breaking Results/ Workshops) (2014).
- S61** GAO T., HULLMAN J. R., ADAR E., HECHT B., DIAKOPOULOS N.: Newsviews: An automated pipeline for creating custom geovisualizations for news. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2014), CHI'14, ACM, pp. 3005–3014.
- S62** ENDERT A., HOSSAIN M. S., RAMAKRISHNAN N., NORTH C., FIAUX P., ANDREWS C.: The human is the loop: new directions for visual analytics. *Journal of Intelligent Information Systems* 43, 3 (2014), 411–435.
- S63** SATYANARAYAN A., HEER J.: Authoring narrative visualizations with ellipsis. *Comput. Graph. Forum* 33, 3 (June 2014), 361–370.
- S64** LE MUZIC M., PARULEK J., STAVRUM A., VIOLA I.: Illustrative visualization of molecular reactions using omniscient intelligence and passive agents. *Comput. Graph. Forum* 33, 3 (June 2014), 141–150.
- S65** LEE B., RICHE N. H., ISENBERG P., CARPENDALE S.: More than telling a story: Transforming data into visually shared stories. *IEEE computer graphics and applications*, 5 (2015), 84–90.
- S66** MU M., SIMPSON S., RACE N., NIAMUT O., KOOT G., KAPTEIN R., TAAL J., MORI L.: "Let's share a story": Socially enhanced multimedia storytelling. *MultiMedia*, IEEE 22, 3 (July 2015), 54–6.
- S67** WOOD J.: Visualizing personal progress in participatory sports cycling events. *Computer Graphics and Applications*, IEEE 35, 4 (July 2015), 73–81.
- S68** LE MUZIC M., WALDNER M., PARULEK J., VIOLA I.: Illustrative timelapse: A technique for illustrative visualization of particle-based simulations. In Visualization Symposium (PacificVis), 2015 IEEE Pacific (April 2015), pp. 247–254.
- S69** SADEGHI F., TENA J., FARHADI A., SIGAL L.: Learning to select and order vacation photographs. In Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on (Jan 2015), pp. 510–517.
- S70** BOY J., DETIENNE F., FEKETE J.-D.: Storytelling in information visualizations: Does it engage users to explore data? In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2015), CHI'15, ACM, pp. 1449–1458.
- S71** AMINI F., HENRY RICHE N., LEE B., HURTER C., IRANI P.: Understanding data videos: Looking at narrative visualization through the cinematography lens. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA, 2015), CHI'15, ACM, pp. 1459–1468.

- S72** RODRÍGUEZ M. T., NUNES S., DEVEZAS T.: Telling stories with data visualization. In Proceedings of the 2015 Workshop on Narrative & Hypertext (New York, NY, USA, 2015), NHT'15, ACM, pp. 7–11.
- S73** KIM G., SIGAL L.: Discovering collective narratives of theme parks from large collections of visitors' photo streams. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (New York, NY, USA, 2015), KDD'15, ACM, pp. 1899–1908.

APÊNDICE C – ROTEIRO DE ENTREVISTA COM ANALISTAS DE DADOS

Nome: _____

Idade:

Gênero: () F () M

UF em que reside e trabalha

Profissão e formação.

Quais redes sociais utiliza? Utiliza todas com a mesma frequência?

Com qual objetivo você utiliza as redes sociais? Qual tipo de informação você acha possível obter a partir dos *posts* em mídias sociais?

Você se considera um analista de dados? Por quê?

Quais das tecnologias a seguir você conhece?

- Gephi
- Many Eyes
- Tableau
- D3.js
- Timeline.js
- Profuser
- Outros:

Você utiliza infográficos e/ou visualizações de dados para obter informações nas suas atividades de trabalho?

Você utiliza algum software/ ferramenta de visualização para criar infográficos e/ou visualizações? Qual(is)?

O quanto você considera importante visualizações de dados na análise de dados? Por quê?

Você já vivenciou alguma situação em que precisou analisar dados sem o apoio de recursos computacionais? Pode descrever? O quanto você considera que visualizações teriam ajudado nessa situação?

Em uma situação hipotética, imagine que você possua uma ferramenta em que você poderá informar quais dados deseja visualizar e como você deseja visualizar esses dados. Você conseguiria imaginar uma situação em que isso seria útil para você (pode ilustrar com situações reais como, por exemplo: dados relacionados à Copa do Mundo, às eleições, jogos olímpicos etc., ou com dados fictícios).

Você já passou por uma situação de necessidade de alguma visualização ou alguma “combinação de dados” que poderia te ajudar em um processo de análise?

APÊNDICE D – TUTORIAL DO CIVIS

Este tutorial consiste na explicação sobre os componentes do CIViS para auxiliar *designers* e analistas de dados *co-designers* no projeto de visualizações narrativas.

D.1 Análise da base de dados

Para começar, tenha a sua base de dados em mãos e identifique nela:

- a data do primeiro e a do último *post*;
- os usuários que postaram;
- os usuários que foram mencionados nas postagens;
- as *hashtags* contidas nos *posts*;
- os *links* contidos nos *posts*;
- os locais mencionados nos *posts*;
- os locais georeferenciados nos *posts*;
- os usuários correspondentes a outras mídias como jornais, programas de tv etc.;
- outros dados que julgar necessários para a sua análise.

De acordo com o objetivo da sua coleta, você consegue, então, a partir dessa análise inicial da base de dados (representada pela carta sistema de ajuda na figura D.1, pensar sobre a história que irá contar com esses dados.



Figura D.1 – Carta que representa o processo de análise dos dados.

Para projetar a sua visualização narrativa, siga os passos nas próximas seções.

D.2 Seleção dos elementos de narrativa

Após a escolha do modelo, você deverá escolher uma ou mais cartas que representam os elementos de narrativa, apresentadas na figura D.2. As descrições dos elementos estão a seguir:

- **(a) Evento:** é um acontecimento. Pode ser definido como a transição de um estado para outro, causado por um ou mais atores. Ele pode ser identificado a partir de *hashtags* presentes no *post* ou pela identificação do nome de um evento no texto.
- **(b) Ator:** ou atores, são os personagens envolvidos e que, a partir das suas ações, geram as mudanças na história. Atores são identificados a partir da estrutura do texto do *post*, com a identificação do sujeito da frase e/ou dos usuários autores do *post* ou mencionados nele.
- **(c) Localização:** diz respeito ao local no qual o evento ocorre. A localização é identificada a partir desse local, de um *post* georreferenciado, ou de locais mencionados no *post*.
- **(d) Tempo:** é o elemento que ordena e serializa a história. O elemento tempo é obtido por meio da data e hora do *post* ou data de ocorrência do evento.

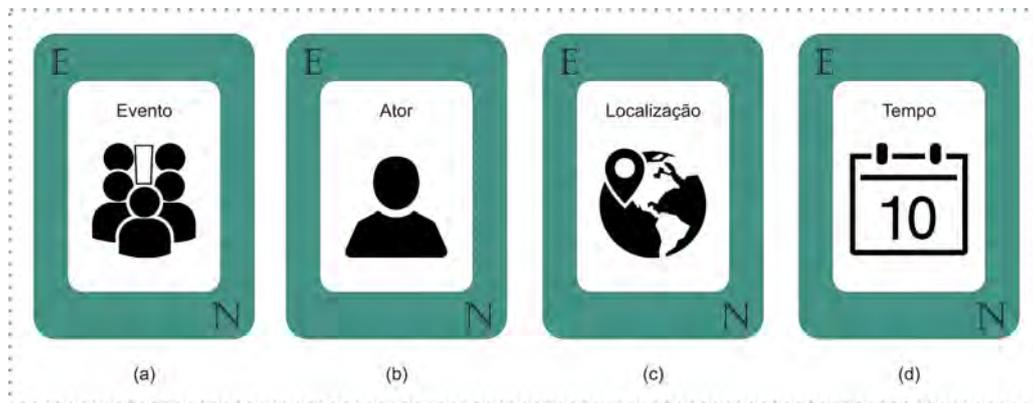


Figura D.2 – Cartas que representam os elementos de narrativa: (a) evento, (b) ator, (c) localização e (d) tempo.

D.3 Seleção do modelo de visualização narrativa

Os modelos de visualização narrativa englobam um espectro de abordagens orientadas pelos autores (que tendem a ser pouco interativas e com muito uso de mensagens) e abordagens orientadas pelo leitor (que tendem a ser mais interativas e menos ordenadas). Para a sua visualização narrativa você deverá escolher uma carta referente a um dos três modelos híbridos apresentados na figura D.3, que contêm as combinações mais usadas: *Martini Glass Structure*, *Interactive Slideshow* e *Drill-Down Story*. A seguir estão as descrições dos modelos:

- **(a) Martini Glass Structure:** inicia com a abordagem guiada pelo autor com o uso de perguntas, observações ou textos para introduzir a visualização. Assim que a visualização narrativa guiada pelo autor estiver completa, ela se abre para um estágio guiado pelo leitor, em que o usuário fica livre para explorar os dados de forma interativa. Na estrutura em forma de copo de Martini, a haste representa a narrativa guiada pelo autor, e a taça representa os caminhos disponíveis e possíveis de interação guiada pelo leitor.
- **(b) Interactive Slideshow:** segue um formato típico de apresentação de slides, mas incorpora interação do leitor com a narrativa, dentro dos limites de cada slide. Esta estrutura permite ao usuário explorar mais detalhadamente pontos específicos da visualização antes de avançar para a próxima etapa da história. No entanto, os slides, individualmente observados, geralmente funcionam como o *Martini Glass*, com a narrativa guiada pelo autor antes de permitir ao usuário interagir.
- **(c) Drill-Down Story:** apresenta um tema geral e, em seguida, permite ao usuário escolher entre instâncias particulares desse tema para revelar mais detalhes. Esta estrutura possui mais ênfase na abordagem guiada pelo leitor, possibilitando ao usuário escolher quais histórias são contadas e quando. No entanto, ela ainda requer uma quantidade significativa de autoria do *designer* para determinar os possíveis tipos de interação do usuário, quais histórias serão incluídas, e os detalhes a serem incluídos para cada história.

Assim, o CIViS oferece ao *designer* cartas que simbolizam esses modelos híbridos.

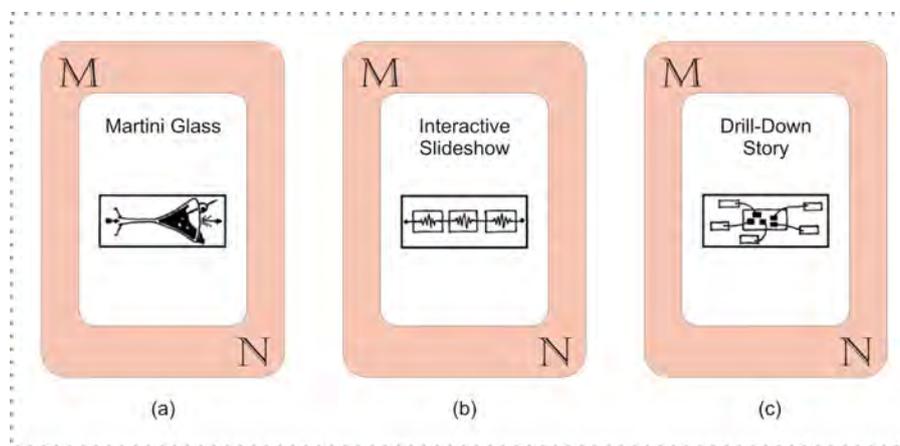


Figura D.3 – Cartas que representam os modelos híbridos de visualizações narrativas do CIViS: (a) *Martini Glass Structure*, (b) *Interactive Slideshow* e (c) *Drill-Down Story*.

D.4 Definição das cenas

A cena é o ambiente em que você irá organizar a sua história. Cada cena pode conter várias partes (*frames*) e, cada parte, um conjunto de elementos diferente. Na figura D.4 são apresentados alguns modelos de divisão de uma cena em *frames*. É importante ressaltar que sua

história pode conter uma ou várias cenas. Assim, você deve definir quantas cenas comporão a sua visualização narrativa interativa. No entanto, como essas decisões podem ser revisitadas no ciclo do fluxo do CIViS, a decisão final não precisa ser agora.

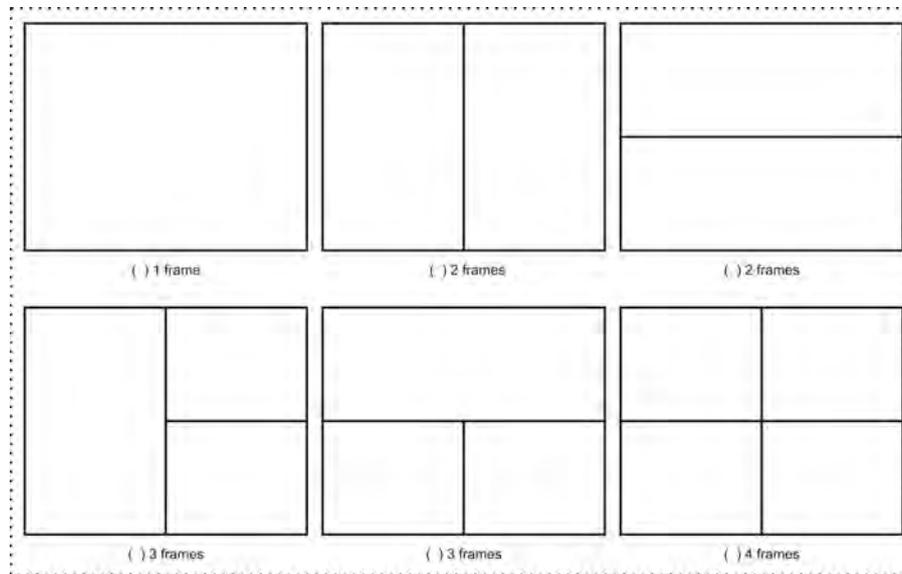


Figura D.4 – Opções de estruturas de *frames* de uma cena.

D.5 Seleção dos gêneros

Gêneros de narrativas são formas básicas de apresentar a história. Assim, você deverá selecionar uma carta que represente o gênero de narrativa que será usado na cena atual da sua história. Na figura D.5 são apresentadas as cartas que representam esses gêneros e suas descrições estão logo a seguir:

- **(a) *magazine style***: é uma forma de representação de histórias em que os dados são apresentados em um único quadro, no estilo de página de revista ou jornal.
- **(b) *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.
- **(c) *partitioned poster***: apresenta um pôster ou imagem particionada, cujas partes apresentam múltiplas visões da história.
- **(d) *flow chart***: apresenta fluxos em uma forma visual, possuindo um ponto inicial e um ponto final. Neste tipo de visualização, é possível que haja diferentes “caminhos” a serem seguidos.
- **(e) *comic strip***: possui fluxo contínuo, direcionado e com múltiplos quadros e cenas, que possuem relação causal direta entres as cenas.

- **(f) *slide show***: é uma forma de apresentação de dados de forma sequencial. A história, neste gênero, possui múltiplos quadros/cenas, representados por slides.
- **(g) *film/vídeo/animation***: narra a história de forma totalmente direcionada e pré-estabelecida. Este gênero é muito utilizado em televisão e comerciais.

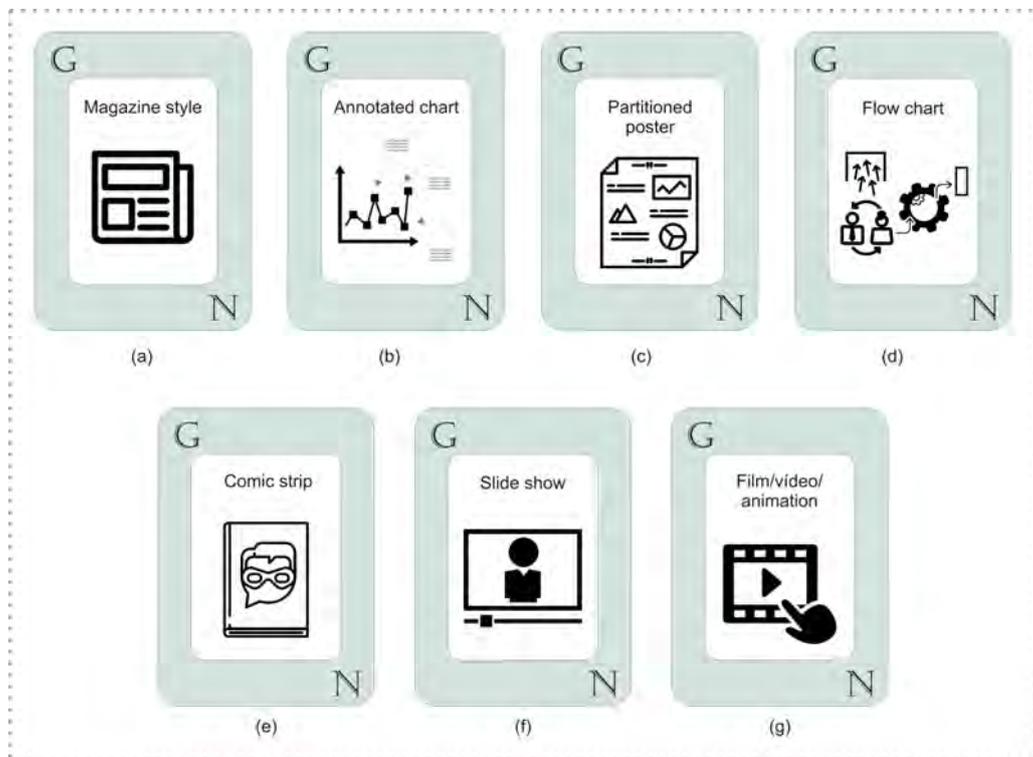


Figura D.5 – Cartas que representam os gêneros de visualizações narrativa: (a) *magazine style*, (b) *annotated chart*, (c) *partitioned poster*, (d) *flow chart*, (e) *comic strip*, (f) *slide show* e (g) *film/vídeo/animation*.

D.6 Seleção dos técnicas de visualização

Na sequência, você deverá selecionar uma ou mais técnicas de visualização que irão compor as cenas da sua história:

- **(a) gráfico de linha**: é um tipo de gráfico que exibe informações com uma série de pontos de dados, chamados de marcadores, ligados por segmentos de linha reta. É semelhante a um diagrama de dispersão, exceto que os pontos de medição são ordenados (tipicamente pelo seu valor do eixo-x) e ligados por segmentos de reta. Um gráfico de linhas é muitas vezes usado para visualizar uma tendência nos dados em intervalos de tempo - uma série de tempo. Ele é composto por dois eixos, um vertical e outro horizontal, e por uma linha que mostra a evolução de um fenômeno ou processo

- **(b) gráfico de barras:** é um gráfico com barras retangulares e comprimento proporcional aos valores que ele representa. As barras podem ser desenhadas verticalmente ou horizontalmente.
- **(c) gráfico de pizza:** é um diagrama circular em que os valores de cada categoria estatística representada são proporcionais às respectivas medidas dos ângulos.
- **(d) grafo:** é uma estrutura $G(V,E)$, onde V é um conjunto não vazio de objetos denominados vértices (ou nós) e E é um subconjunto de pares não ordenados de V , chamados arestas.
- **(e) gráfico de área:** é um tipo de gráfico que possui valores nos eixos x e y em que, geralmente, o eixo x representa dados temporais e o eixo y apresenta alguma variável categórica.
- **(f) mapa de calor:** é um tipo de representação gráfica que apresenta os valores individuais de um conjunto de dados dispostos em uma matriz e utiliza cores para representar as grandezas dos valores. É comum que os valores mais altos possuam cores mais escuras e os valores mais baixos, cores mais claras. Com o mapa de calor é possível contrastar informações e comparar rapidamente uma grande quantidade de itens.
- **(g) gráfico de coordenadas paralelas:** permite visualizar dados multidimensionais, cujas dimensões referem-se aos atributos que o gráfico possui. Os atributos são representados por 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.
- **(h) gráfico de bolhas:** permite a visualização de três variáveis, pois é composto por um eixo x , um eixo y e, além disso, é possível visualizar a proporção de um valor no gráfico por meio do tamanho da bolha.
- **(i) tabela:** é uma representação matricial, em linhas e colunas, tantas quantas a aplicação que se queira dar.
- **(j) mapa:** é uma representação do mundo físico que possibilita estabelecer relações, destacar regiões, visualizar proporções de diversos domínios de dados.

D.7 Seleção dos elementos visuais

Os elementos visuais são os recursos visuais que auxiliam no uso da narrativa e envolvem três sub-categorias, apresentadas na figura D.7 e descritas a seguir. Você deverá indicar quais elementos visuais estarão presentes em cada parte da sua visualização, sendo que um mesmo elemento pode estar presente em várias partes (frames, cenas, técnicas de visualizações).

- **(a) estruturas visuais:** mecanismos presentes na visualização que fazem com que o leitor compreenda toda a narrativa, além de identificar em que parte da história ele se encontra.

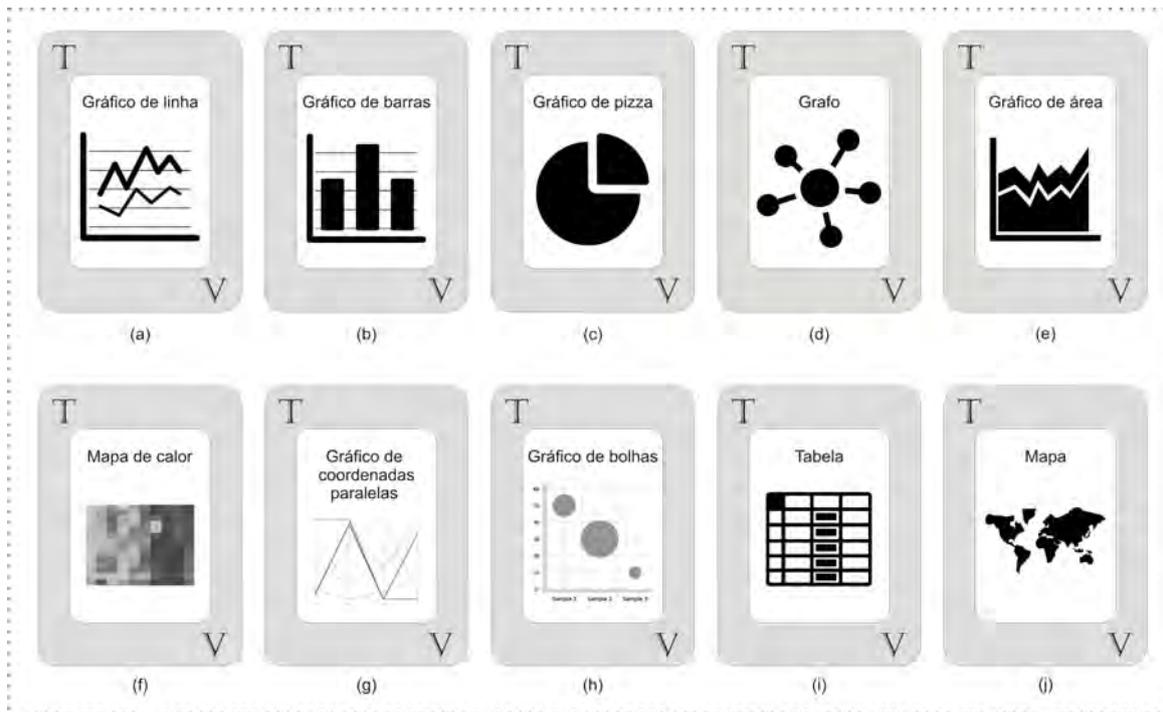


Figura D.6 – Cartas que representam as técnicas de visualização de dados: (a) gráfico de linha, (b) gráfico de barras, (c) gráfico de pizza, (d) grafo, (e) gráfico de área, (f) mapa de calor, (g) gráfico de coordenadas paralelas, (h) gráfico de bolhas, (i) tabela e (j) mapa.

No CIViS sugerimos os recursos de *splash screen*, padrões visuais consistentes, barras de progresso ou histórico das mudanças.

- **(b) 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* ou sugestão (quando o sistema sugere dados, como, por exemplo, as dez *hashtags* mais citadas).
- **(c) orientação de transições:** técnicas para movimentações dentro ou entre cenas, podendo ser, por exemplo, interação, cor, animações, linhas ou setas, botões de próximo/anterior e barras de *scroll*.

D.8 Seleção dos elementos estruturais

Os elementos estruturais, quando presentes na narrativa, poderão auxiliar e facilitar o seu uso. Também estão divididos em três sub-categorias, apresentadas na figura D.8 e descritas a seguir. Você deverá indicar quais elementos estruturais estarão presentes em cada parte da sua visualização, sendo que um mesmo elemento pode estar presente em várias partes (frames, cenas, técnicas de visualizações).

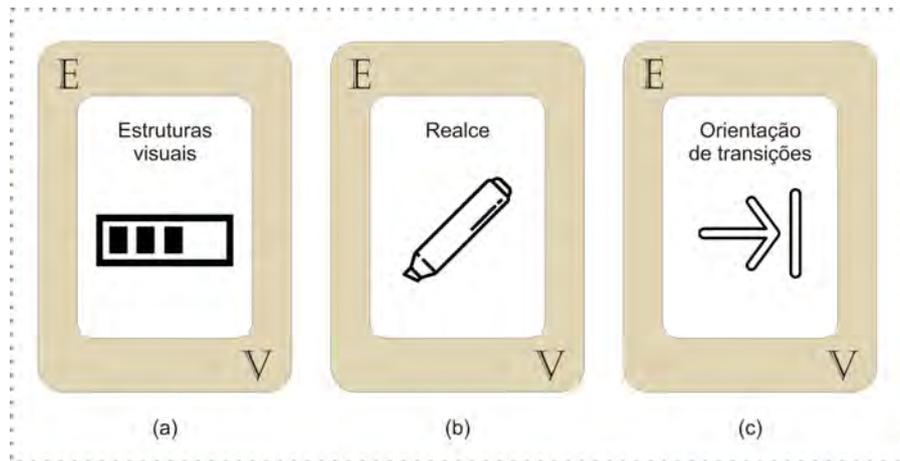


Figura D.7 – Cartas que representam os elementos visuais: (a) estruturas visuais, (b) realce e (c) orientação de transições.

- **(a) ordenação:** é a forma pela qual as cenas são ordenadas, podendo ser de forma randômica, direcionada de acordo com o leitor ou de forma linear, como em um slide, por exemplo.
- **(b) interatividade:** refere-se às diferentes estratégias de manipulação do conteúdo como filtro, seleção, navegação, *brushing*, zoom.
- **(c) mecanismos de mensagem:** são as formas que as narrativas comunicam observações e comentários ao leitor, que podem ser pelo uso de títulos, anotações, artigos, comentários, textos introdutórios ou resumos.

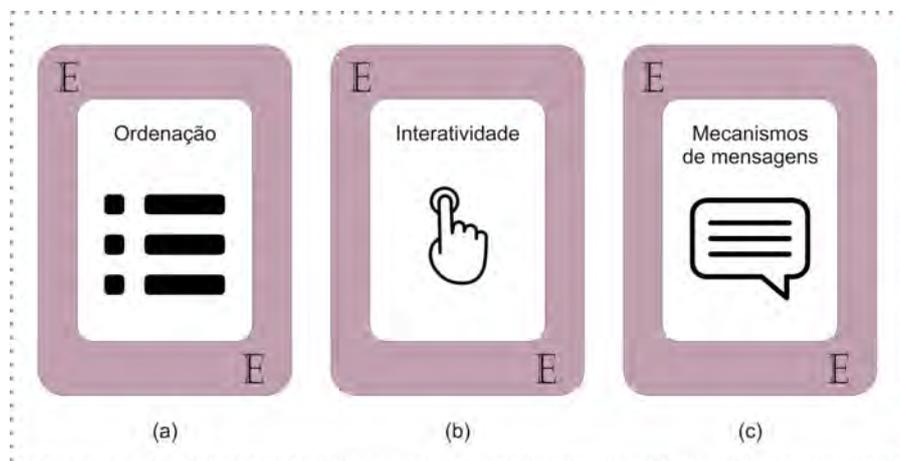


Figura D.8 – Cartas que representam os elementos estruturais: (a) ordenação, (b) interatividade e (c) mecanismos de mensagem.

D.9 Sistema de ajuda

O sistema de ajuda consiste em falar diretamente aos usuários da visualização sobre ela e sobre como eles podem fazer um uso melhor dela. Para criar o sistema de ajuda, exercite

responder as perguntas abaixo para cada elemento presente na visualização narrativa interativa criada, conforme indicado por Silveira et al. [SdSB03]:



Figura D.9 – Carta que representa o sistema de ajuda.

- O que é isto?
- Como faço isto?
- Onde está. . . ?
- Para que serve isto?
- E agora?
- A quem isto afeta?
- Por que eu devo fazer isto?
- Onde eu estava?
- De quem isto depende?
- O que aconteceu?
- Quem pode fazer isto?
- Epa!
- Por que não funciona?
- Existe outra maneira de fazer isto?
- Socorro!

Essas perguntas irão te ajudar a pensar nos mecanismos de ajuda que você considera necessário oferecer aos usuários/leitores da sua visualização narrativa. Para criar esses mecanismos, você poderá fazer uso dos demais componentes do CIViS, como mecanismos de texto, interatividade, vídeo ou animação, entre outros.

APÊNDICE E – MATERIAL UTILIZADO NO ESTUDO DE USO DO CIViS

Neste apêndice apresentaremos os materiais utilizados durante o estudo de uso do CIViS pelos analistas de dados *co-designers*.

E.1 Apresentação do CIViS aos participantes

ROTEIRO

1. Apresentação da pesquisadora
2. Apresentação da pesquisa e seus objetivos
3. Explicação oral de o que é o CIViS, apresentando o material impresso que foi utilizado
4. Explicação do objetivo do estudo de uso do CIViS
5. Explicação do Termo de Consentimento Livre e Esclarecido
6. Explicação de como o estudo de uso do CIViS foi estruturado e como aconteceria

E.2 Termo de Consentimento Livre e Esclarecido

Modelo de interface customizável para apoiar a construção de visualizações narrativas de dados extraídos de mídias sociais - CIViS

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

Termo de Consentimento Livre e Esclarecido

Eu, Caroline Queiroz Santos, orientada pela Profa. Dra. Milene Selbach Silveira, desenvolvi como parte da minha pesquisa de doutorado o modelo de interface customizável para apoiar a construção de visualizações narrativas de dados extraídos de mídias sociais, nomeado de CIViS. O CIViS consiste em uma ferramenta epistêmica, fundamentada na teoria da Engenharia Semiótica, para apoiar *designers* e analista de dados *co-designers* na modelagem de visualizações narrativas interativas. Nós agradecemos a sua participação e contribuição para o avanço da pesquisa em Interação Humano-Computador.

O objetivo desta pesquisa é analisar o uso do CIViS no apoio à customização de visualizações narrativas interativas. Para isto, os participantes deste estudo são convidados a usar os componentes do CIViS, enquanto são observados por um pesquisador. Esta observação será registrada em papel, em áudio e, eventualmente, em vídeo. As informações nos trarão dados importantíssimos para verificar a qualidade do modelo em questão.

Lembramos que o objetivo deste estudo **não é** avaliar o participante, **mas sim** avaliar o modelo que o participante estará usando. O uso que se faz dos registros efetuados durante o estudo 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 do estudo pode interrompê-lo e estará nos fazendo um favor se registrar por escrito as razões ou sensações que o levaram a esta atitude. A pesquisadora fica obrigada a descartar o estudo 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 estudo, qualquer restrição ou condição adicional que lhe pareça aplicar-se aos itens acima enumerados (1, 2, 3 e 4). A pesquisadora se compromete a observá-las com rigor e entende que, na ausência de tal manifestação, o participante concorda que rejeita o comportamento ético da pesquisadora somente as condições impressas no presente documento.
6. A pesquisadora tem direito de utilizar os dados do estudo, 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 estudo.

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: _____ – Escola Politécnica - PUCRS

E.3 Questionário Perfil

Modelo de interface customizável para apoiar a construção de visualizações narrativas de dados extraídos de mídias sociais - CIViS

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

Questionário Perfil

1) Dados pessoais:

Nome: _____

Idade: _____

Gênero: () M – () F – () Não declarar

Formação: _____

Profissão: _____

2) Análise de mídias sociais:

2.1) Você já fez análise de dados extraídos de mídias sociais? () Sim – () Não

2.2) Você utiliza (ou utilizou) algum software para ajudar nessa análise? Qual(is)?

2.3) Como você costuma apresentar a análise desses dados para outras pessoas?

2.4) Você utiliza (ou já utilizou) alguma ferramenta de visualização de dados? Qual(is)?

2.5) Qual a sua opinião sobre as ferramentas de visualização de dados que você conhece ou que já utilizou (em relação à facilidade de uso ou de aprender a usar)?

2.6) Na sua opinião, como seria apresentar os dados extraídos de mídias sociais por meio de uma narrativa utilizando técnicas de visualização de dados?

E.4 Fluxo de passos para customização de visualização narrativa interativa

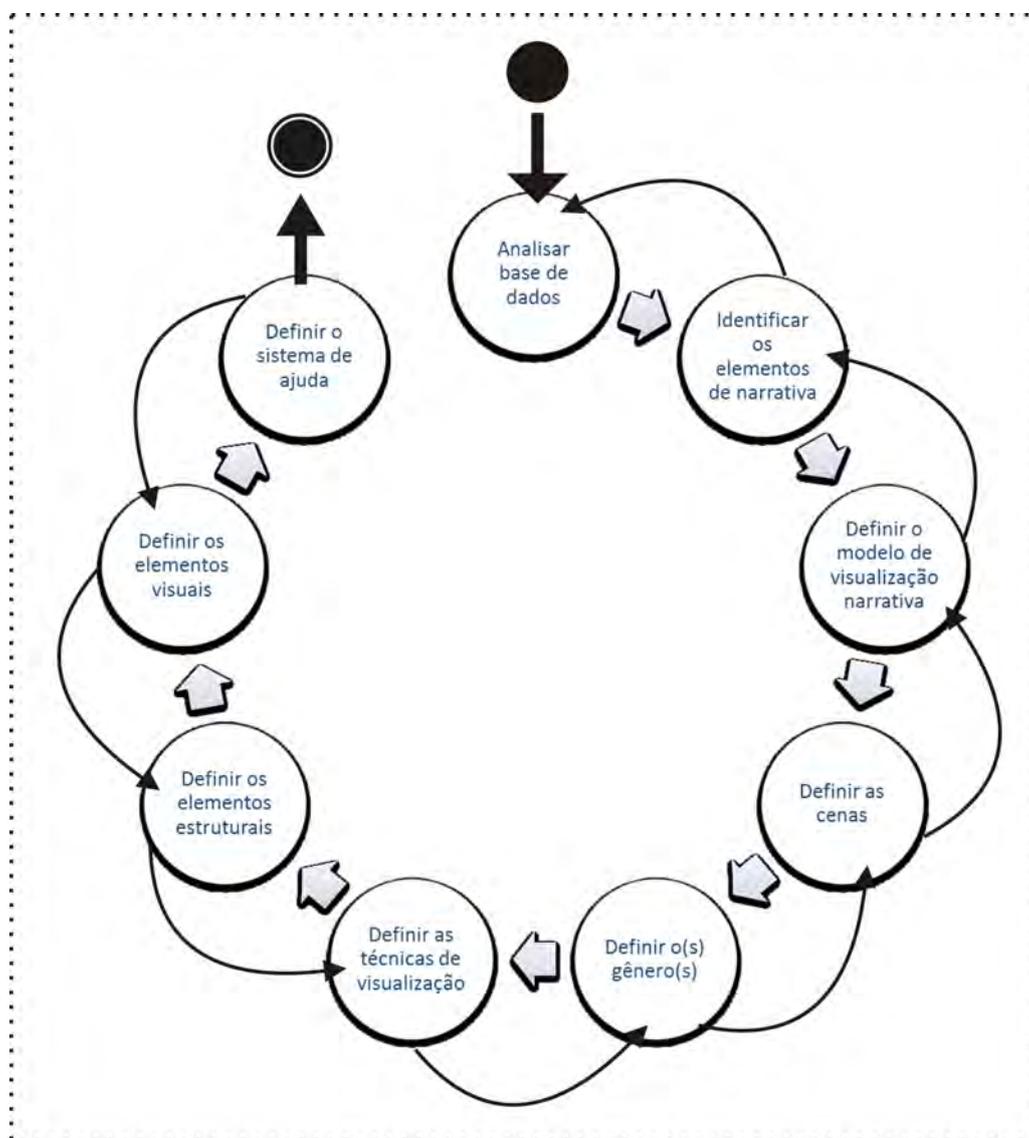


Figura E.1 – Fluxo de passos para customização de visualização narrativa interativa.

E.5 Fluxo de passos para customização de visualização narrativa interativa refinado

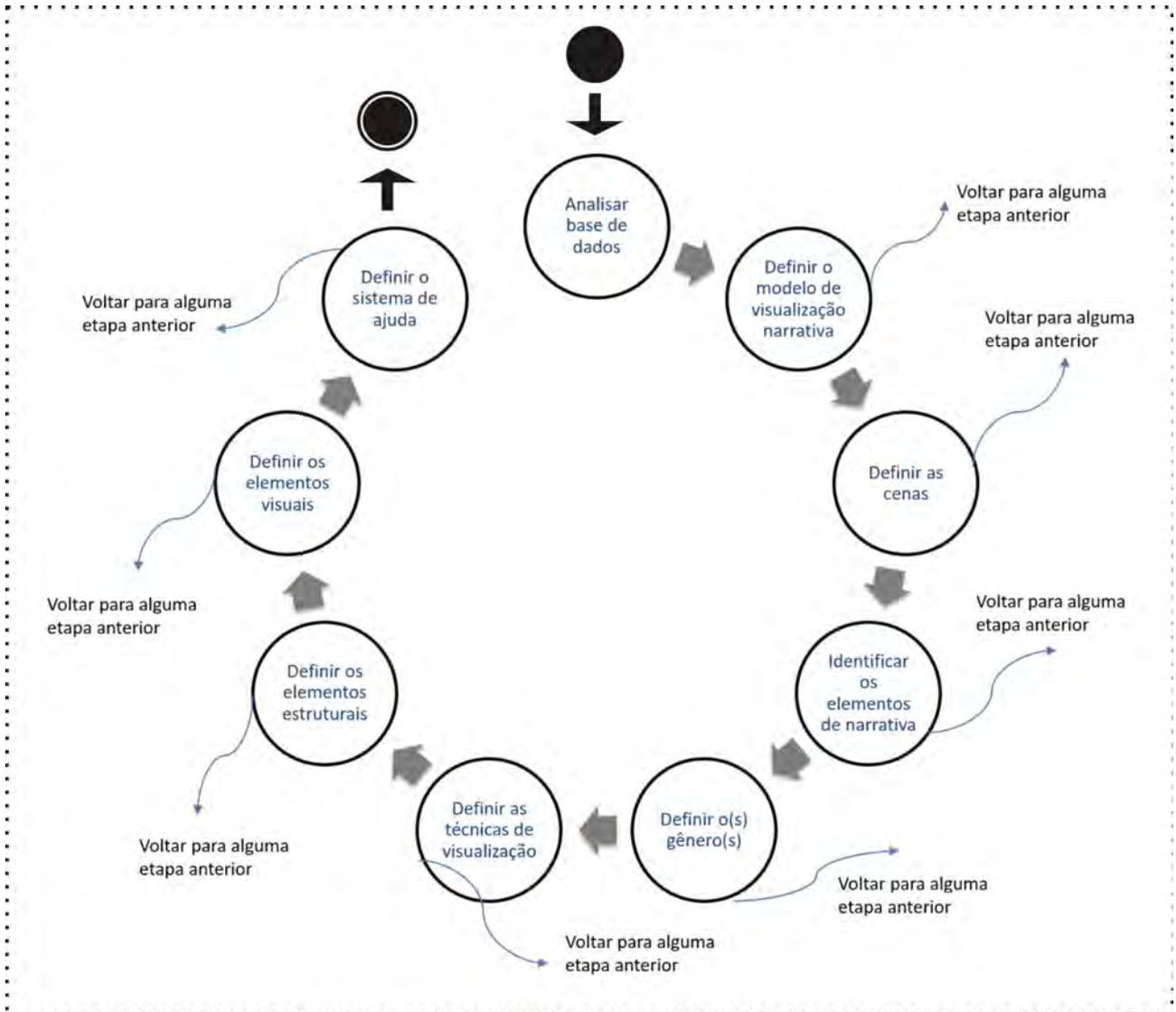


Figura E.2 – Fluxo de passos para customização de visualização narrativa interativa após análise de representação visual das possibilidades de retornar para qualquer etapa anterior.

E.6 Dados para o estudo

CONTEXTO DE DADOS PARA ANÁLISE DO CIViS

Os dados que servirão de base para este estudo de análise do CIViS foram coletados no Twitter por meio da ferramenta *Twitter Capture and Analysis Toolset* (DMI–TCAT), que coleta *tweets* e gera arquivos que permitem a análise de seus componentes (como *hashtags*, URLs e *mentions*). Este contexto foi extraído do artigo **TERRO-RISMO E OLIMPÍADAS RIO 2016: QUEM SÃO OS “TERRORISTAS” NO TWIT-TER?**, de autoria de Roberta Firmino da Silva e Carlos D’Andréa, publicado no livro “Olimpíadas Rio 2016: mídia, política, humor”, PPGCOM/UFMG^a.

A coleta foi realizada com o objetivo de identificar quais atores e referências estavam associados aos termos “terrorista” e “terroristas” durante um período de três semanas (7 a 28 de julho de 2016) antes das Olimpíadas. Este intervalo de tempo foi marcado por um acirramento das ações antiterror no país, por exemplo através de um alerta da Agência Brasileira de Inteligência (Abin) sobre como identificar terroristas (08 de julho) e, principalmente, do desencadeamento da “Operação *Hashtag*” (em 21 de julho), que iniciou a detenção dos doze “suspeitos” de terrorismo no Brasil. No cenário brasileiro, destacaram-se os dias 21 e 22, nos quais *hashtags* como #Rio2016, #Brasil e #ForaTemer ganharam certo destaque e ocorreu uma maior associação de imagens (fotos, vídeos e gifs) aos *tweets* publicados nesses dois dias. Os dados coletados apontam ainda para uma apropriação de imagens de arquivo de personalidades da TV aberta, em uma prática marcada pelo humor.

^aDisponível em <http://www.seloppgcom.fafich.ufmg.br/index.php/seloppgcom/catalog/book/41>

E.7 Execução

Nome do participante: _____

Para guiá-lo na utilização do CIViS, considere o fluxo de passos apresentados na seção E.4, o contexto de dados para análise do CIViS e o seu entendimento sobre os componentes do modelo.

TAREFA

Você está trabalhando em uma equipe que deverá analisar os dados extraídos de mídias sociais sobre terrorismo e Olimpíadas 2016. Você ficou encarregado de fazer a modelagem de uma visualização narrativa interativa, de acordo com o CIViS, a fim de contar a história desse assunto por meio de visualizações de dados. No primeiro momento, você deverá refletir sobre os dados que possui, refletir sobre as possibilidades que o modelo te oferece e definir como irá iniciar a sua narrativa interativa. A narrativa poderá ser apresentada em uma ou em várias cenas. Sua tarefa, agora, é fazer essa modelagem livremente, com base nas informações e recursos que possui. Durante a modelagem, você deverá preencher a tabela de decisões, de acordo com as suas decisões tomadas. Você pode utilizar o material de apoio sobre o CIViS e também pode tirar dúvidas comigo. Você terá aproximadamente 35 minutos para realizar esta tarefa.

E.8 Tabela de registro das decisões do *co-designer*

Análise de uso do CIVis pelo analista de dados *co-designer*

Nome: _____

Preencha a tabela abaixo de acordo com as decisões tomadas ao longo da realização da modelagem de visualização narrativa interativa.

Tabela 1: registro das decisões do *co-designer* durante a modelagem de uma cena

Componente CIVIS		Decisão do <i>co-designer</i> para o componente			Abrangência da decisão ¹	
Análise da base de dados e seleção dos elementos de narrativa	Evento(s)					
	Ator(es)					
	Localização					
	Tempo					
Modelo de visualização narrativa						
Cena	<i>Frame 1</i>	<i>Frame 2</i>	<i>Frame 3</i>	<i>Frame 4</i>	<i>Frame 5</i>	
Gênero(s)						
Técnicas de visualização						
Elementos visuais						
Elementos estruturais						
Sistema de ajuda						

¹ Coloque o valor correspondente: 01 – somente no frame x, 02 – em toda a cena y, ou 03 – em toda a história.

E.9 Roteiro de entrevista semiestruturada sobre a tarefa

Nome: _____

1) Você considera que a sequência de passos para criação da visualização narrativa interativa do CIViS representa um fluxo adequado para este fim?

Não representa 0 - 1 - 2 - 3 - 4 - 5 Totalmente

Explique: _____

2) Você teve dificuldades em compreender os componentes do CIViS?

Nenhuma 0 - 1 - 2 - 3 - 4 - 5 Muita

Explique: _____

3) O material de apoio te ajudou?

Muito pouco 0 - 1 - 2 - 3 - 4 - 5 Muito

Explique: _____

4) Quais os pontos positivos de criar uma visualização narrativa interativa usando o CIViS?

5) Quais os pontos negativos de criar uma visualização narrativa interativa usando o CIViS?

6) O uso do CIViS provocou em você alguma reflexão sobre aspectos de visualização narrativa interativa sobre os quais você não havia pensando antes? Se sim, quais?

7) A proposta do CIViS é apoiar o *designer* e o analista de dados *co-designer* a tomar decisões sobre visualização narrativa enquanto está projetando o sistema. Neste cenário, você acha que o CIViS seria uma ferramenta útil?

Explique: _____

8) Caso queira, escreva sobre suas impressões a respeito do CIViS.