JAQUELINE ZAIA DE SOUSA

InterVis: um sistema para geração e exploração interativas de visualizações de informação

São Paulo

2016

JAQUELINE ZAIA DE SOUSA

InterVis: um sistema para geração e exploração interativas de visualizações de informação

Dissertação apresentada à Escola de Artes, Ciências e Humanidades da Universidade de São Paulo para obtenção do título de Mestre em Ciências pelo Programa de Pós-graduação em Sistemas de Informação.

Área de Concentração: Sistemas de Informação

Versão corrigida contendo as alterações solicitadas pela comissão julgadora em 28 de setembro de 2016. A versão original encontra-se em acervo reservado na Biblioteca da EACH-USP e na Biblioteca Digital de Teses e Dissertações da USP (BDTD), de acordo com a Resolução CoPGr 6018, de 13 de outubro de 2011.

Orientador: Prof. Dr. João Luiz Bernardes Jr.

São Paulo

2016

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

CATALOGAÇÃO-NA-PUBLICAÇÃO
(Universidade de São Paulo. Escola de Artes, Ciências e Humanidades. Biblioteca)

Sousa, Jaqueline Zaia de
InterVis : um sistema para geração e exploração interativas de visualizações de informação / Jaqueline Zaia de Sousa ; orientador, João Luiz Bernardes Jr. – São Paulo, 2016
89 f. : il

Dissertação (Mestrado em Ciências) - Programa de Pós-Graduação em Sistemas de Informação, Escola de Artes, Ciências e Humanidades, Universidade de São Paulo
Versão corrigida

1. Inferface gráfica. 2. Desenvolvimento de software. 3. Interface homem-computador. 4. Visualização . 5. Informação. I. Bernardes Junior, João Luiz, orient. II. Título

CDD 22.ed.– 006.6

Dissertação de autoria de Jaqueline Zaia de Sousa, sob o título **"InterVis: um sistema para geração e exploração interativas de visualizações de informação"**, apresentada à Escola de Artes, Ciências e Humanidades da Universidade de São Paulo, para obtenção do título de Mestre em Ciências pelo Programa de Pós-graduação em Sistemas de Informação, na área de concentração Sistemas de Informação, aprovado em 28 de setembro de 2016 pela comissão examinadora constituída pelos doutores:

Prof. Dr. João Luiz Bernandes Júnior

Instituição: Escola de Artes, Ciências e Humanidades da Universidade de São Paulo

Presidente

Prof. Dr. Luciano Antonio Digiampietri

Instituição: Escola de Artes, Ciências e Humanidades da Universidade de São Paulo

Profa. Dra. Veronica Teichrieb

Instituição: Universidade Federal de Pernambuco

Prof. Dr. Edgard Afonso Lamounier Júnior

Instituição: Universidade Federal de Uberlândia

Resumo

SOUSA, Jaqueline Zaia de. **InterVis: um sistema para geração e exploração interativas de visualizações de informação**. 2016. 89 f. Dissertação (Mestrado em Ciências) – Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, São Paulo, 2016.

Devido ao crescimento acelerado da quantidade de dados disponíveis para análise hoje, é comum lidar com grandes conjuntos de dados, por vezes complexos demais para serem interpretados na sua forma bruta. Por isso, as técnicas de Visualização de Informação têm o objetivo de facilitar para o ser humano a tarefa de análise e interação dos dados por meio da sua abstração em forma gráfica. Justificado pela necessidade de dar autonomia ao usuário final na criação de visualizações, este trabalho pretende delinear a importância da participação do usuário na criação e suporte à abstração da informação em forma gráfica. Para isso, foi desenvolvido o InterVis – um sistema para criação interativa de visualizações de informação a partir de dados dinâmicos, que visa a permitir que o usuário final possa gerar e editar visualizações de acordo com a sua necessidade, independentemente da natureza da informação que deve ser analisada. O sistema foi testado utilizando o questionário USE a fim de verificar se a criação interativa de visualizações de informação, sem programação, aliada ao conhecimento do usuário sobre o domínio da aplicação, é mais eficiente da perspectiva da usabilidade sem perda significativa de flexibilidade. Os testes envolveram a execução de tarefas por indivíduos de um grupo de usuários. Todos os usuários conseguiram realizar todas as tarefas de criação e exploração em tempo hábil e avaliaram positivamente o sistema, além de terem sugerido diversas melhorias e novas funcionalidades. É possível concluir que o InterVis atualmente já atende às expectativas iniciais do trabalho, ainda que haja pontos a serem aprimorados em trabalhos futuros.

Palavras-chave: Visualização de Informação. GUI. Criação Interativa.

Abstract

SOUSA, Jaqueline Zaia de. **InterVis: a system for interactive criation and exploration of information visualizations**. 2016. 89 p. Dissertation (Master of Science) – School of Arts, Sciences and Humanities, University of São Paulo, São Paulo, 2016.

Because of the growing amount of data available for analysis today, it is common to deal with large data sets, often too complex to be interpreted in their brute form. That is why Information Visualization techniques exist, to facilitate the analysis and interaction with data by humans through graphical abstractions. Motivated by the need to allow end users the autonomy to generate and edit visualizations, this work aims to underscore the importance of end user participation in the creation and support of these graphical abstractions of data. For this purpose, it was developed a system for interactive creation of Information Visualizations based on dynamic data, which aims to allow the final user to generate e edit visualizations according to their need and independently of the nature of the information that should be analyzed. This system was tested using the USE questionnaire, to verify whether this interactive creation of Information Visualizations, without programming, allied to the user knowledge of each application's domain, will be more efficient from the perspective of usability without significant loss of flexibility, as expected. The tests were compound of the tasks' execution by individuals of a users' group. All the users were able to conclude all tasks of creation and exploration in due time and evaluated positively the system, besides they have been suggested diverse improvement and new functionalities. It is possible to conclude that InterVis already fulfills the initial expectations of this work, although there are still points to be refined in future work.

Palavras-chave: Information Visualization. GUI. Interative Creation.

Lista de figuras

Figura 1 - Visualização de Informação em Mapa utilizando conceito de Overview+Detail no Google Mapas 17
Figura 2 - Formas de visualização de Informação em Mapas utilizando o conceito de Focus+Context com FisheyeView 19
Figura 3 – Visualização de Informação da ZUI Pad, que divide a visualização em partes que possuem Zoom 20
Figura 4 – Visualização gerada pelo Polaris 21
Figura 5 – Gráfico de Coordenadas Paralelas com preservação de privacidade dos dados 22
Figura 6 – Técnicas de visualização de séries de dados temporais: (a) ThemeRiver; (b) TimeSerarcher; (c) Life-Lines; e (d) NameVoyager. 23
Figura 7 – Visualizações de Informação que representam hierarquias: (a) TreeMap; (b) ConeTree; (c) DocuBurst; e (d) Tree Cube.....24
Figura 8 – Visualizações de informação que representam dados relacionados: (a) FUSE; (b) ZAME; (c) H-BLOB; e (d) Ask-GraphView.....25
Figura 9 - Visualização multifocal de clusters baseada em calendário 30
Figura 10 – Exemplo uma hierarquia de classes Monolítica e Polilítica na representação de um retângulo: (a) Em tempo de compilação; e (b) Em tempo de execução 32
Figura 11 – Visualização do VisLink combinando duas visualizações bidimensionais em um layout tridimensional com relacionamento entre os dados 40
Figura 12 - Perspectiva de Geração Interativa de Visualizações de Informação.....49
Figura 13 - Painel de prévia da visualização sendo criada 50
Figura 14 - Painel de configurações das visualizações 50
Figura 15 - Painel de conjuntos de dados abstraídos em metadados 51
Figura 16 - Perspectiva de Exibição das Visualizações de Informação.....53
Figura 18 - Módulos funcionais do sistema 56
Figura 19 - Representação do Design Pattern MVC.....56
Figura 20 – Modelo de classes do módulo de visualizações deste sistema 58
Figura 21 - Modelo de classes do módulo de interação do sistema.....59
Figura 22 - Resultado do questionário USE separado pelos quesitos avaliados.....69

Lista de tabelas

Tabela 1 – Ferramentas de geração de visualizações e suas características técnicas principais 38
Tabela 2 – Ferramentas de Visualização de Informação e seus possíveis layouts 41
Tabela 3 - *Kits* de Ferramentas e suas características técnicas principais 43
Tabela 4 - *Kits* de Ferramentas e os tipos de layout aos quais eles dão suporte.....44

Lista de abreviaturas e siglas

2D	Bidimensional
3D	Tridimensional
CSV	Comma Separated Values
GUI	Graphic User Interface
JPEG	Joint Photographic Experts Group
JSON	JavaScript Object Notation
MVC	Model View Controller
SVG	Scalable Vector Graphics
XML	Extensible Markup Language
ZUI	Zoomable User Interface

Sumário

1 Introdução 11
1.1 Criação Interativa 13
1.2 Visualização Interativa 13

2 Visualização de Informação: Principais Conceitos 14
2.1 Definição 14
2.2 Conceitos 16
2.2.1 Overview+Detail 18
2.2.2 Focus+Context 18
2.2.3 Zoom 19
2.3 Tipos de Dado 21
2.3.1 Dados N-dimensionais 22
2.3.2 Dados Temporais 23
2.3.3 Dados Hierárquicos 24
2.3.4 Dados Relacionados 25
2.4 Técnicas de Visualização 26
2.4.1 Animação 27
2.5 Técnicas de Interação 28
2.5.1 Visual Data Mining 29
2.5.2 Interação Multifocal e *Information Hiding* 30
2.6 Arquitetura da Interface 31
2.6.1 Monolítica 32
2.6.2 Polilítica 33

3 Metodologia e Trabalhos Correlatos 34
3.1 Metodologia de Pesquisa 34
3.2 Metodologia de Avaliação 36
3.3 Análise de Trabalhos Relacionados 37
3.3.1 Softwares Análogos 37
3.3.2 *Kits* de Ferramentas de Suporte ao Desenvolvimento 42
3.3.2.1 The Visualization Toolkit 45
3.3.2.2 prefuse 45
3.3.2.3 The InfoVis Toolkit 46
3.4 Considerações Finais 46

4 Solução 48
4.1 Requisitos 48
4.1.1 Criação Interativa 48
4.1.1.1 Criar e salvar relatórios interativos apresentáveis como slides 49
4.1.1.2 Criar visualizações de informação contidas nos relatórios 50

4.1.1.3 Permitir a visualização de dados estáticos e dinâmicos, tanto de arquivos quanto de bases de dados relacionais 51
4.1.1.4 Omitir os dados em sua forma bruta 51
4.1.1.5 Selecionar interativamente modelos, configurações, dados e eventos na criação das visualizações 52
4.1.1.6 Pré-visualizar o relatório durante o processo de criação a visualização 52
4.1.1.7 Editar relatórios previamente criados 52
4.1.1.8 Proporcionar um processo de criação de visualizações eficiente e útil de forma fácil e satisfatória para o usuário 53
4.1.2 Visualização Interativa 53
4.1.2.1 Visualizar um relatório previamente criado 54
4.1.2.2 Possibilitar interação com as visualizações de um relatório 54
4.1.2.3 Proporcionar uma visualização de informação útil e eficiente como resultado que possibilite ao usuário identificar informações e responder perguntas sobre os dados de forma fácil e satisfatória 54
4.2 Implementação 54
4.3 Arquitetura 56
4.4 Técnicas de Visualização 57
4.5 Técnicas de Interação 58

5 Testes e Resultados 60
5.1 Metodologia Empregada 60
5.2 População 61
5.3 Procedimento 62
5.3.1 Termo de Consentimento Livre Esclarecido 62
5.3.2 Pré-teste 62
5.3.3 Execução do Teste 63
5.3.3.1 Os Dados 63
5.3.3.2 As Tarefas 63
5.3.3.3 Avaliação de Usabilidade Pelos Participantes 65
5.4 Resultados 66
5.5 Riscos 70
5.5.1 Riscos Éticos 70
5.5.2 Riscos à Validade do Teste 71

6 Conclusão 73

Referências 77

Apêndice A – Termo de Consentimento Livre Esclarecido 84

Apêndice B – Instruções para Tarefas 87

Apêndice C - Perguntas e Resultados do Questionário 88

1 Introdução

Nos últimos anos pode ser notado o crescimento acelerado da quantidade de dados disponíveis para análise, justificado não só pela automação de atividades em inúmeras áreas, mas também devido ao aumento da capacidade de armazenagem de dados sustentada pelo avanço tecnológico (ALENCAR, 2007). Dessa forma, é comum nos mais diversos contextos a necessidade de lidar com grandes volumes de dados, assim como executar tarefas analíticas e exploratórias com estes, a fim de extrair correlações e padrões interessantes dos dados (THAKUR; RHYNE, 2009).

Aplicações no mercado financeiro e séries de código genético são exemplos de que a complexidade e volume dos dados podem omitir características valiosas no contexto da aplicação, uma vez que os dados em sua forma bruta são pouco interpretáveis de forma eficiente pelo ser humano. Algumas dessas informações podem ser facilmente reconhecíveis pelo ser humano, se utilizada uma apresentação adequada. Por esse motivo, cada vez mais técnicas de visualização de informação têm sido desenvolvidas, a fim de proporcionar ao ser humano, por meio da abstração das informações em forma gráfica de maneira interativa e exploratória, a interpretação e extração de informações de uma grande massa de dados (OLIVEIRA; LEVKOWITZ, 2003).

É possível encontrar diversas soluções no mercado com essa finalidade que se propõem a dar suporte na tarefa de análise de dados e que contam com o auxílio de gráficos e dashboards, como TABLEAU (TABLEAU, 2015), Excel (EXCEL, 2016), SAS Visual Analitics (SAS, 2015) e Oracle Business Analytics (ORACLE, 2015).

Apesar das ferramentas de visualização de informação terem se mostrado indispensáveis no suporte à interpretação de dados ao longo dos últimos anos, ainda existem impeditivos a seu uso. Um dos grandes problemas no emprego de visualização de informação é o grau de dificuldade do seu desenvolvimento, que envolve a necessidade de conhecimento avançado em programação matemática e interação para a construção de estruturas complexas e gráficos dinâmicos (HEER; CARD; LANDAY, 2005), assim como entendimento dos dados do ponto de vista da aplicação. Outro motivo que limita o uso de visualizações é a dificuldade em encontrar um modelo reutilizável, visto que cada domínio de aplicação exige técnicas diferentes de representação dos dados para gerar uma visualização bem-sucedida (LEE *et. al.,* 2004).

A necessidade de customização aliada à complexidade de geração de visualizações delineia o envolvimento de duas naturezas de conhecimento específico na criação de uma visualização de informação: a área de aplicação da ferramenta e conhecimento em programação e geração de gráficos. A primeira é necessária dada a necessidade de reconhecer quais informações interessam ao usuário para serem visualizadas. A partir do tipo de informação que se deseja explorar, a segunda serve para definir e implementar as melhores estruturas de dados e abstrações visuais para representar com excelência aquele contexto.

É possível encontrar inúmeros *kits* de ferramentas e aplicações de visualização de informação ricos em recursos, porém, enquanto alguns oferecem interface gráfica e restringem o domínio da aplicação como o *ASK-Graph View* (ABELLO; VAN HAM; KRISHNAN, 2006), *VisLink* (COLLINS; CARPENDALE, 2007), *OpenedEyes* (ALMEIDA; APOLINÁRIO, 2012) e *FUSE* (SEAH; BHOWMICK; DEWEY; YU, 2012); outros proporcionam uma gama de codificações de componentes reutilizáveis e visualizações predefinidas, como são os casos dos *kits* de ferramentas *InfoVis* (FEKETE, 2004), p*refuse* (HEER; CARD; LANDAY, 2005), *JUNG* (O'MADADHAIN et. al., 2005) e *VTK* (SCHROEDER; MARTIN; LORENSEN, 2004), por exemplo, que requerem construção ou customização para atenderem os requisitos de uma visualização de informação para uma aplicação em particular. Em ambos os casos, o usuário final da aplicação tem sua autonomia em relação à exploração de dados restrita aos recursos que a visualização predefinida lhe oferece.

Este trabalho se justifica pela necessidade de permitir que o usuário final tenha autonomia para gerar e editar visualizações de acordo com a sua necessidade independentemente da natureza da informação que deve ser analisada. Deseja-se, portanto, delinear a importância da participação do usuário na criação e suporte à abstração da informação em forma gráfica e a importância da usabilidade e viabilidade envolvidas na utilização da visualização de informação como ferramenta para as mais diversas tarefas. Para tanto, é desenvolvido e avaliado o InterVis – uma aplicação que fornece uma interface gráfica com o usuário (GUI) para criação interativa de visualizações de informação, que não demanda ao usuário conhecimento técnico na área de visualização ou necessidade de produção de código para gerar as representações gráficas interativas e não necessita de contato direto com os dados brutos.

Além disso, outro fator levado em consideração nesta solução é a proteção dos dados em sua forma bruta, motivado pela necessidade de impedir seu mal-uso escondendo

informações sensíveis no contexto da aplicação, como dados pessoais ou informações corporativas de risco (DASGUPTA; KOSARA, 2011), situação comum, por exemplo, em domínios de aplicação na área da saúde, onde é imprescindível a omissão de dados pessoais.

O objetivo deste trabalho, portanto, é desenvolver e testar este sistema para criação interativa de visualizações de informação a partir de dados dinâmicos, que visa a permitir que o usuário final possa gerar, editar e explorar visualizações de acordo com a sua necessidade independentemente da natureza da informação que deve ser analisada.

Os principais requisitos do trabalho podem ser divididos em duas perspectivas: a criação e a exploração interativas de representações visuais:

1.1 Criação Interativa

- Criação e salvamento de visualizações tanto com dados estáticos quanto com referência a um banco de dados relacional;
- Seleção interativa das coleções de dados oriundas de um banco de dados relacional de forma abstraída;
- Seleção interativa de modelos, configurações e eventos na criação das visualizações;
- Análise de usabilidade do processo de criação interativa de visualizações de informação.

1.2 Visualização Interativa

- Exibição e possibilidade de interação com uma visualização previamente criada;
- Análise de resultado da criação e utilização de visualizações de informação.

Este trabalho está organizado em sete capítulos: essa introdução, uma breve exposição dos conceitos de Visualização de Informação, a discussão das metodologias de pesquisa e avaliação estudadas, a descrição da solução proposta, os testes e seus resultados e, por fim, a conclusão do trabalho.

2 Visualização de Informação: Principais Conceitos

Nos próximos três anos serão gerados mais dados do que em toda a história pregressa do ser humano, oriundos de sistemas que, muitas vezes, armazenam informações de transações e acontecimentos do dia a dia de forma automática. O resultado disso são bases de dados multidimensionais, das quais se crê que seja possível extrair informações valiosas no contexto da aplicação; porém, dada a complexidade e volume de dados, torna-se inviável para o ser humano interpretar e analisar estas informações de forma textual (KEIM, 2002).

Uma vez que a forma como os dados serão representados depende diretamente dos problemas que o usuário deseja resolver (SHNEIDERMAN, 1996), a visualização pode ser feita de diversas formas, que variam de acordo com os tipos de dados e de relacionamentos entre eles. As visualizações que utilizam linha do tempo (PLAISANT *et al.*, 1996), por exemplo, são ideais para descrever históricos pessoais e outras séries temporais, enquanto grafos são bastante úteis na visualização de relacionamentos (GANSNER; HU; KOBOUROV, 2010).

Por esse motivo, a área de visualização de informações visa a possibilitar a exploração das habilidades naturais de percepção visual do ser humano de encontrar e interpretar informação (KENNEDY; MITCHELL; BARCLAY, 1996). A criação da visualização ideal deve levar em conta os tipos de dados a serem exibidos e como é necessário interagir com eles, a fim de obter a visualização que melhor represente os dados. Por esse motivo, uma técnica de visualização de informação pode ser classificada de acordo com três critérios: os dados que ela exibe, os tipos de visualização que ela utiliza e, por fim, as formas como é possível interagir com ela (KEIM, 2001).

2.1 Definição

"A Visualização de Informação deve fazer pela mente o que o que os automóveis fazem pelos pés" (CARD, 2008, tradução nossa). Seguindo esta ideia, Card (2008) define Visualização de Informação como uma coleção de tecnologias que utilizam processamento visual para amplificar a compreensão humana de informação abstrata; que vai ao encontro de Keim et. al. (2006) que afirmam que Visualização de Informação visa a proporcionar a comunicação de dados abstratos por meio do uso de interfaces visuais interativas. Nestas definições, utilizadas como base no presente trabalho, é possível notar propositalmente a não especificação de quais tecnologias e interfaces podem ser utilizadas e qual tipo de processamento visual. Isso por que

a intenção deste trabalho é discutir a ferramenta de visualização do ponto de vista da usabilidade em função da tarefa a ser executada pelo usuário e classificá-la de acordo com o tipo de informação a ser abstraído, a representação visual utilizada e as formas de interação disponíveis (KEIM, 2002).

Além da Visualização de Informação, outro tópico amplamente estudado é o de Visualização Científica, definida como o uso de tecnologias de processamento gráfico como ferramenta para compreensão de dados obtidos por meio de simulações ou medições físicas (HABER, 1990); ou ainda o uso de computação gráfica para criar imagens que representem números, resultados ou conceitos científicos (MCCORMICK; DEFANTI; BROWN, 1987). Enquanto a Visualização Científica tem sua evolução no fim dos anos 1980, a Visualização de Informação se desenvolve mais tarde, em meados dos anos 1990. Apesar de ambas se desenvolverem em conjunto no decorrer dos anos trazendo a discussão sobre suas reais diferenças e semelhanças (RHYNE *et. al.*, 2003), pode-se mencionar como diferença principal que enquanto a Visualização Científica tem foco na representação de dados espaciais associados a processos científicos, geralmente tridimensionais, como aplicações nas áreas de geologia, topologia ou dinâmica de fluídos; a Visualização de informação tem como objetivo a criação de metáforas visuais para dados frequentemente não espaciais, não-numéricos e multidimensionais, que auxiliam na análise de dados como documentos, redes sociais e informações corporativas (CHEN, 2005).

Dada a relevância que a visualização de informações possui hoje e o objetivo do usuário ao optar pela sua utilização, pode-se citar como essencial o princípio da transparência de Rutkowski (1982), que afirma que, quando o usuário concentra sua energia na tarefa que está a executar, a ferramenta aparenta desaparecer. Assim, pode-se afirmar que a visualização não deve ser percebida pelo usuário como ferramenta, mas como uma abstração onde a tarefa a ser executada por meio da informação apresentada será o foco da sua atenção.

Ou seja, a visualização de informação tem o objetivo de ser uma ferramenta de auxílio na execução de tarefas e tomada de decisões. Uma vez que a quantidade de dados que é possível ser exibida textualmente para que um ser humano a interprete é próxima de cem itens, fazê-lo dessa forma torna-se praticamente impossível quando se trata de coleções de dados de milhões de itens. Logo, o volume de dados disponíveis hoje com potencial utilidade acaba se resumindo a grandes bancos de dados subutilizados no seu contexto de aplicação (KEIM, 2002).

Além disso, outra questão importante na visualização de informação e que muitas vezes não recebe atenção o suficiente é a sua usabilidade, definida como: "a medida pela qual um produto pode ser usado por usuários específicos para alcançar metas específicas com efetividade, eficiência e satisfação em um contexto específico de uso" (ISO/EIC, 1998). Um dos possíveis motivos é a particularidade de cada visualização de informação, que requer estudos e testes planejados individualmente em função do tipo de tarefa a ser executada e dos tipos de dados a serem explorados (FREITAS *et al.*, 2002). Como o objetivo da Visualização de Informação é explorar a capacidade cognitiva do ser humano e dar suporte a tomada de decisões por meio de interfaces visuais interativas (KEIM *et al.*, 2006), a usabilidade tanto na criação interativa de visualizações de informação, quanto na utilização da visualização criada se fazem imprescindíveis no processo de desenvolvimento desta solução.

2.2 Conceitos

A finalidade da Visualização de Informação é proporcionar o caminho que permita ao usuário executar suas tarefas da forma mais fácil, intuitiva e proveitosa, sem se preocupar com a forma como a visualização deve ser conduzida. Pensando nisso, um dos conceitos mais relevantes em Visualização, conhecido como o Mantra de Shneiderman, afirma: "Visão geral primeiro, zoom e filtragem, e então detalhes por demanda" (SHNEIDERMAN, 1996, tradução nossa). Aprofundada por autores em diversos domínios de aplicação como, por exemplo, Chen (2004); Plaisant, Carr e Shneiderman (1995) e Herman, Melançon e Marshall (2000), esta abordagem é largamente empregada no desenvolvimento de técnicas de Visualização de Informação, visto que atende às necessidades básicas do usuário.

De acordo com Shneiderman, a primeira etapa para executar uma tarefa deve ser obter uma visão geral do cenário em que esta será desempenhada, o que permite ao usuário traçar seu objetivo e mapear o escopo em que deve trabalhar. Estendendo este conceito à visualização, é importante ter uma visão geral da mesma, pelo fato de reduzir a procura, permitindo a detecção de padrões globais e facilitando a escolha do próximo passo a ser tomado pelo usuário (CARD; MACKINLAY; SHNEIDERMAN, 1999).

Uma vez que o usuário tenha a visão geral dos dados e dos padrões contidos neles, o próximo passo é focar em uma parte específica e analisá-la mais detalhadamente (FEW, 2006). Para isso, utilizando técnicas como filtragem, o usuário pode aproximar-se dos dados e filtrar aqueles que não lhe interessam, a fim de obter uma visualização detalhada e objetiva de parte

dos dados. A busca e a filtragem de dados podem ser processos delicados, pois se trata de uma decisão do usuário de retirar ou selecionar informações, a fim de reduzir os dados visualizados e facilitar o seu manuseamento. Esse processo de refinamento e busca é progressivo e conhecido como *hierarquical decision-making* (CARD; MACKINLAY; SHNEIDERMAN, 1999). É possível que neste momento seja necessário o acréscimo de detalhes individuais ou de um grupo de dados, na medida em que o usuário os procura; como rótulos, descrições e links para outras informações. Uma forma de fazer isso sem interferir na visualização é possibilitar que o usuário selecione os itens da visualização para obter uma nova janela com estas informações complementares (SHNEIDERMAN, 1996).

Também é possível executar buscas dentre os dados, a fim de gerar uma nova visualização somente com os dados de interesse ou complementar a visualização atual. Isso equivale a dizer que o sistema deve ter a capacidade de responder a perguntas do usuário a respeito dos dados que estão em análise (ANDRIENKO; ANDRIENKO; GATALSKY, 2007). A busca pode ser feita utilizando linguagem escrita, de forma que o usuário deva descrever quais dados deseja buscar de acordo com a ferramenta de busca utilizada, ou utilizando elementos gráficos de busca visual, como menus, caixas de seleção e ícones (CATARCI *et al.*, 1997).

Figura 1 - Visualização de Informação em Mapa utilizando conceito de Overview+Detail no Google Mapas

Fonte: Jaqueline Zaia, 2015

Baseado na possibilidade de interesse em mais de um nível de granularidade na informação, existem diversas técnicas bifocais, ou seja, técnicas com mais de uma perspectiva do dado. Pode-se classificar essas técnicas em Overview+Detail e Focus+Context (PIETRIGA; APPERT, 2008). Além disso, é possível, por meio da utilização de Zoom, obter uma visualização bifocal utilizando a dimensão física do tempo, conforme descrito a seguir.

2.2.1 Overview+Detail

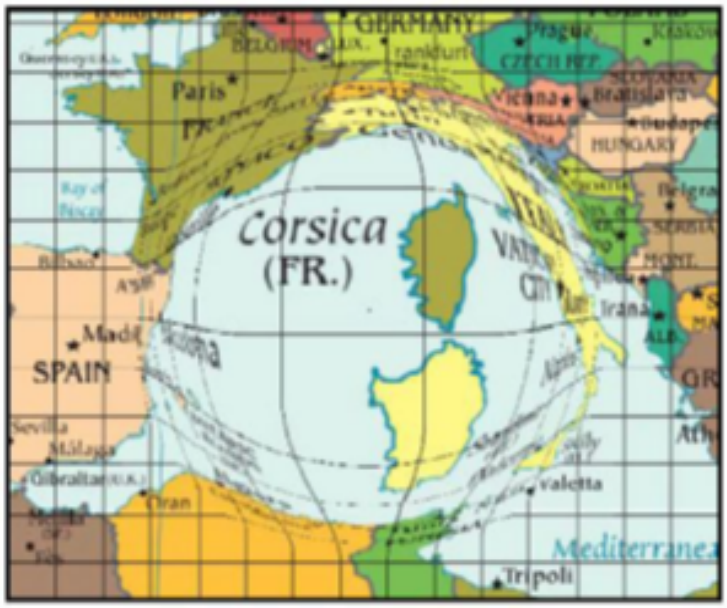
As técnicas de visualização Overview+Detail utilizam dois espaços diferentes para exibir simultaneamente o contexto geral e a visão detalhada dos dados. A interface mais comum reserva uma pequena parte da tela para a visão global, porém esta também pode ocupar a maior área, ou ainda ambas as partes, a geral e o detalhe, podem ser exibidos em espaços iguais (PLAISANT; CARR; SHNEIDERMAN, 1995). É ideal para cenários em que pode ser necessário visualizar detalhes sem perder o contexto global, além de prover um meio de interação, em que o usuário pode, por exemplo, navegar pelo contexto global a fim de encontrar os detalhes que o interessam. Existem diversas interfaces que empregam este conceito, não só em visualizações de informação, mas na navegação em geral, em que há separação de dois espaços, onde o usuário pode interagir separadamente em pelo menos um destes, mas suas ações se refletem em ambos. Um exemplo do uso de Overview+Detail é o Google Mapas, exibido na figura 1, em que pode se notar o mapa em detalhes no espaço maior e em uma parte, no espaço inferior direito da visualização, mostra-se o contexto geral da aplicação. Outras formas de Overview+Detail podem empregar o uso de miniaturas, barras de rolagem, caixas de contexto e sobreposição de conteúdo (COCKBURN; KARLSON; BEDERSON, 2007).

2.2.2 Focus+Context

Já as técnicas Focus+Context unem na mesma visualização, sem separação temporal ou espacial, ambos os níveis: o detalhado e o contextual, de forma que o contexto se mantenha e uma área exiba o detalhe por meio de zoom ou, muitas vezes, utilizando distorções. Um exemplo de visualização que utiliza a abordagem Focus+Context que é frequentemente utilizada são as *Fisheye Views* (FURNAS, 1986), como mostrado na figura 2. Nela, usa-se uma função de grau de interesse em cada ponto da visualização (Degree of Interest- DOI) para gerar uma distorção que mantém o centro de interesse aumentado enquanto a imagem é comprimida

no resto do espaço. Uma vantagem do uso de visualizações bifocais é que podem dispensar o uso do zoom e de barras de rolagem ao unir na mesma visualização ambos os níveis geral e detalhado (PLAISANT; CARR; SHNEIDERMAN, 1995). É importante notar que o uso do zoom por si só não implica a ocorrência de uma interface bifocal, uma vez que este atua como uma separação temporal do contexto global e dos detalhes, porém não fornece referência contextual momentânea entre ambos.

Figura 2 - Formas de visualização de Informação em Mapas utilizando o conceito de Focus+Context com FisheyeView



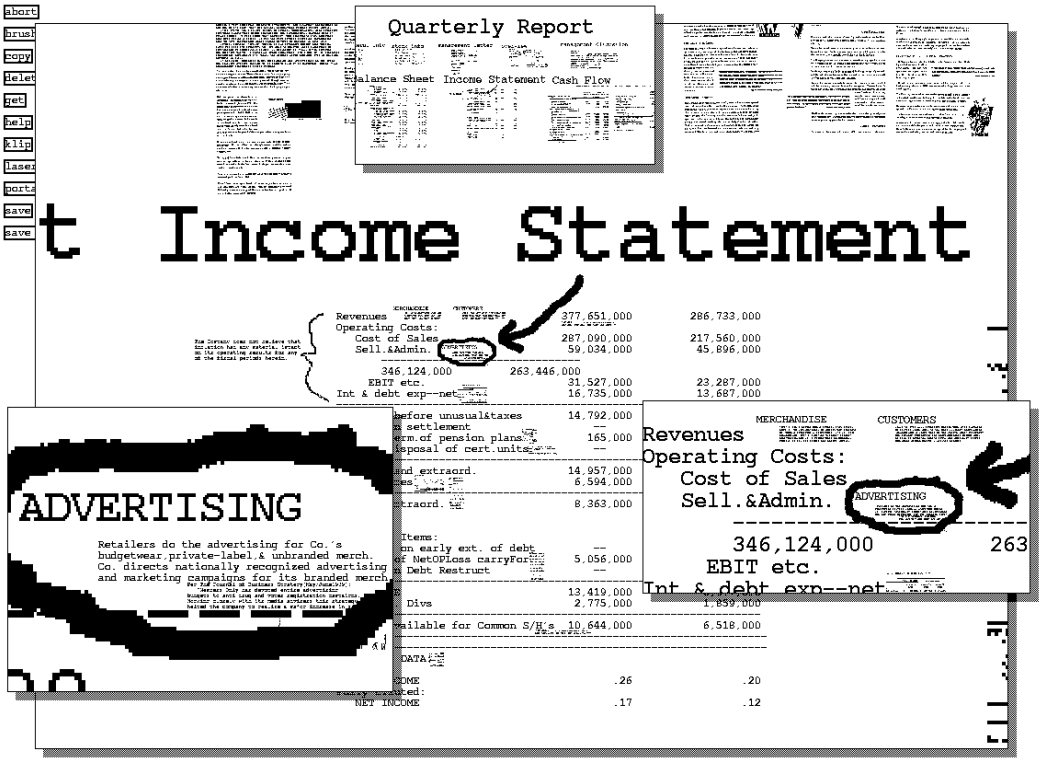
Fonte: Furnas (1986)

2.2.3 Zoom

Por fim, é possível classificar as técnicas bifocais com utilização de Zoom, que permite que ambas as interfaces contextual e detalhada sejam visualizadas, porém, diferentemente das técnicas de Overview+Detail e Focus+Context, sua exibição utiliza o mesmo espaço na tela e as interfaces são separadas pela dimensão física do tempo; ou seja, é possível ter a visão geral e detalhada na mesma visualização, mas não ao mesmo tempo (COCKBURN; KARLSON; BEDERSON, 2007). Vale ressaltar que o recurso de Zoom pode ser utilizado também juntamente com outros conceitos de técnicas, como é o caso da técnica utilizada como exemplo na figura 1, que apresenta não só o conceito de Overview+Detail, mas também possibilita a utilização de Zoom na tela de detalhe; conforme discutido por Hornbæk *et al.* (2002).

Baseado nesta técnica, é possível encontrar na bibliografia o uso do termo Zoomable User Interfaces (ZUI), que, apesar de discutidas desde o início da década, não possuem uma definição aceita com unanimidade. Uma vez que o espaço na tela e o tamanho que os objetos da visualização ocupam são fundamentais para o entendimento e organização da informação (FURNAS; BEDERSON, 1995), é possível reconhecer uma ZUI por meio de dois requisitos: a organização dos objetos na tela é feita em função do espaço e do tamanho que ocupam; o usuário pode interagir diretamente com o espaço e tamanho dos objetos por meio de *zooming* e *panning* (HORNBAEK et. al., 2002).

Figura 3 – Visualização de Informação da ZUI Pad, que divide a visualização em partes que possuem Zoom



Fonte: Perlin e Fox (1993)

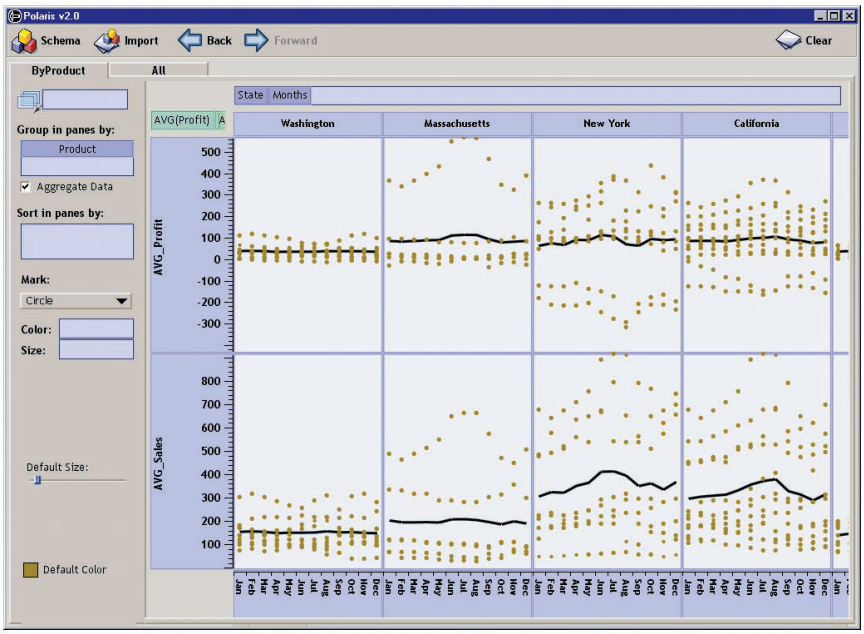
Apesar de a forma mais comum de utilização deste conceito ser o Zoom geométrico, ou seja, em que o tamanho do Zoom determina linearmente o tamanho do objeto exibido na tela, é possível que existam relações mais complexas entre o objeto e seu tamanho – este é o caso do Zoom Semântico (PERLIN; FOX, 1993), em que a visualização muda não só geometricamente quando utilizado o Zoom, mas também exibe mais ou menos detalhes e características. Como

exemplo de ZUI's, pode citar o *Pad*, ilustrado na figura 3, em que é possível ver um documento na íntegra e retângulos que o sobrepõem e mostram porções específicas do documento com zoom (PERLIN; FOX, 1993); e seus sucessores *Pad*++ (BEDERSON; HOLLAN,1994) e *Jazz* (BERDERSON et. al., 2000).

2.3 Tipos de Dado

De acordo com Shneiderman (1996), os dados podem ser organizados de acordo com os problemas que o usuário deseja resolver. A partir disso, ele propõe a taxonomia da tarefa pelo tipo de dado, em que conjuntos de dados podem ser divididos em categorias que inferem sua dimensão e seus relacionamentos: dados de uma, duas, três ou várias dimensões; dados temporais; dados que representem hierarquias (árvores) e relacionamentos (grafos). É importante mencionar que uma coleção de dados pode pertencer a mais de uma categoria.

Figura 4 – Visualização gerada pelo Polaris



Fonte: Stolte, Tang e Hanrahan (2002)

Além disso, os dados geralmente consistem em um grande número de itens, cada um correspondendo a uma observação, em que cada item possui variáveis ou dimensões (KEIM, 2002). Quando analisadas atomicamente levando em consideração a forma como serão visualizadas, as variáveis que compõem o dado podem ser classificadas como nominais, ordinais, quantitativos e intervalos (BERTIN, 1983), também utilizadas de forma resumida em

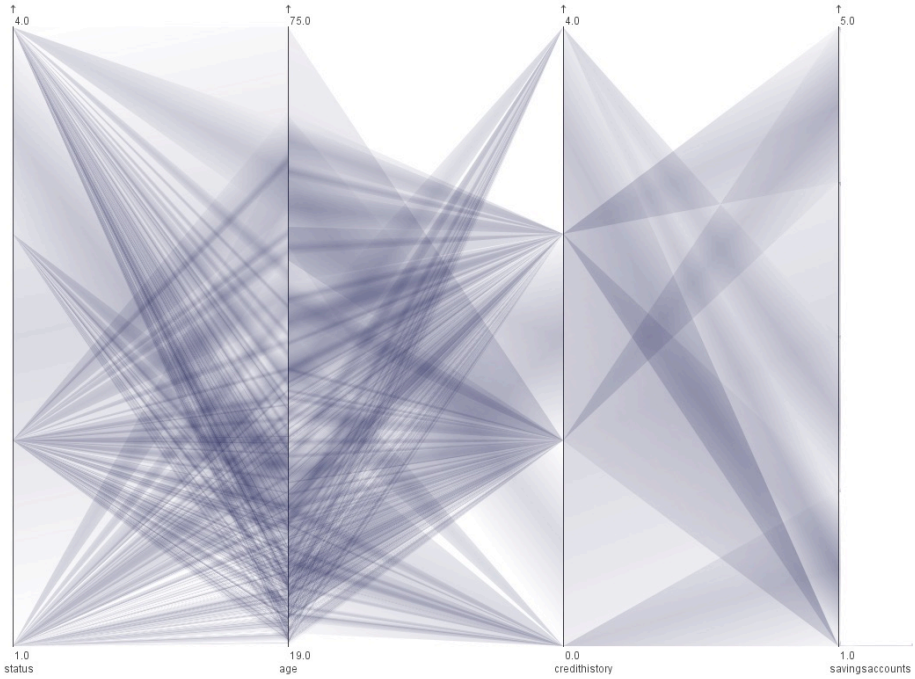
ordinais e numerais, em que variáveis nominais são, em algum nível, consideradas ordinais (STOLTE; TANG; HANRAHAN, 2002), levando em conta, por exemplo, a ordem lexicográfica do nome; e intervalos não são considerados.

Em seguida, seguem as classificações e exemplos de tipos de dados:

2.3.1 Dados N-dimensionais

Dados N-dimensionais são dados com uma, duas, três ou várias dimensões, que podem ser representados de forma tabular em tuplas e podem ser representados de acordo com o número de dimensões. Exemplos de visualização de múltiplas dimensões de dados são o *Polaris,* ferramenta que possibilita criação de diversas visualizações como gráficos de barras e dispersão, ilustrado na figura 4 (STOLTE; TANG; HANRAHAN, 2002); e a técnica de Preservação de Privacidade em Coordenadas Paralelas, ilustrada na figura 5, em que é possível notar quatro coordenadas representando variáveis e linhas entre elas representando os valores, que são desfocados a fim de preservar a privacidade dos dados individuais (DASGUPTA; KOSARA, 2011).

Figura 5 – Gráfico de Coordenadas Paralelas com preservação de privacidade dos dados

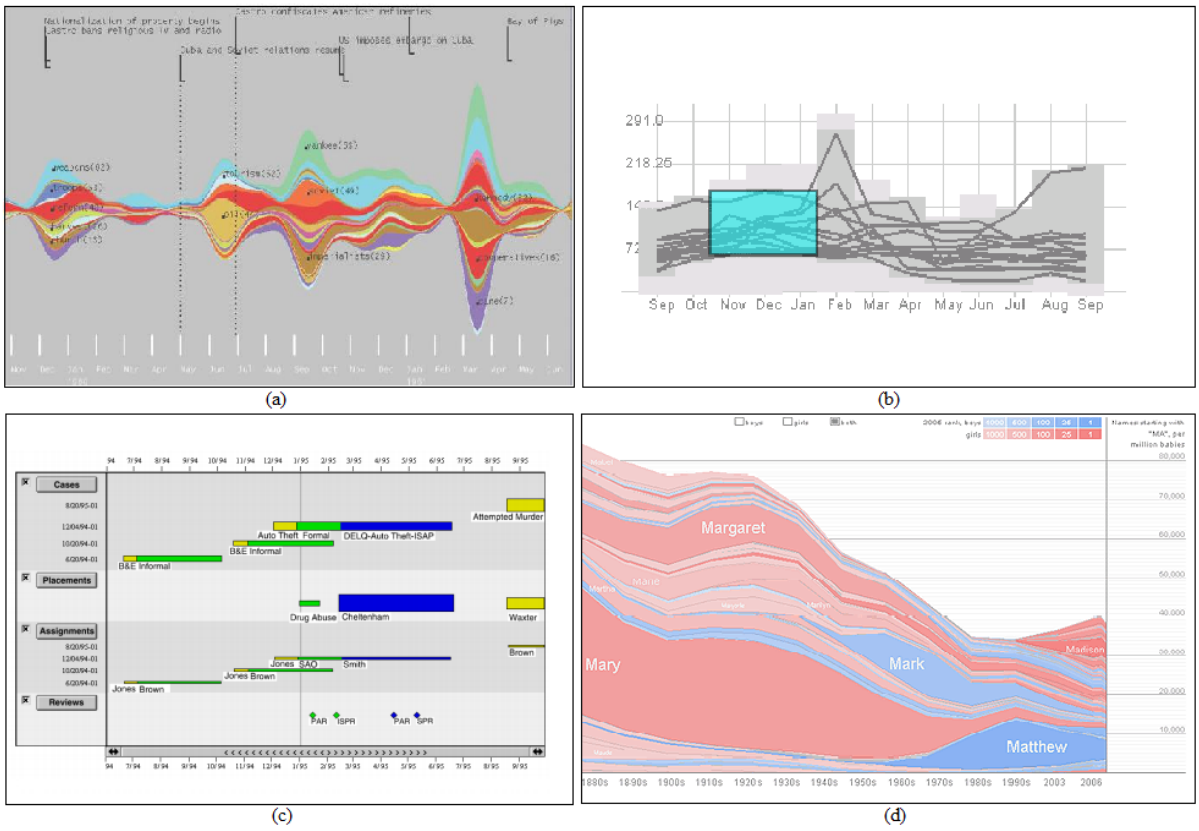


Fonte: Dasgupta e Kosara (2011)

2.3.2 Dados Temporais

De acordo com Granger e Newbold (1986), entendem-se séries de dados temporais como uma sequência de observações ordenadas por um parâmetro de tempo, que podem ser medidas de forma contínua ou em intervalos discretos de tempo e que, idealmente, consistem de observações temporalmente equidistantes umas das outras e sem valores faltantes (YAFFEE; MCGEE, 2000). Ou seja, são consideradas coleções de dados temporais aquelas que possuem pelo menos uma variável temporal, que possui alguma referência de data ou hora, de forma que torna-se possível ordená-la em função do tempo. Dados temporais são frequentemente visualizados com o auxílio de linhas do tempo, como pode-se verificar na figura 6, nas visualizações *ThemeRiver* (HAVRE *et al.*, 2002), *TimeSearcher* (HOCHHEISER; SHNEIDERMAN, 2004), *Life-Lines*, (PLAISANT et al., 1996) e *NameVoyager* (WATTENBERG, 2005).

Figura 6 – Técnicas de visualização de séries de dados temporais: (a) ThemeRiver; (b) TimeSerarcher; (c) Life-Lines; e (d) NameVoyager.

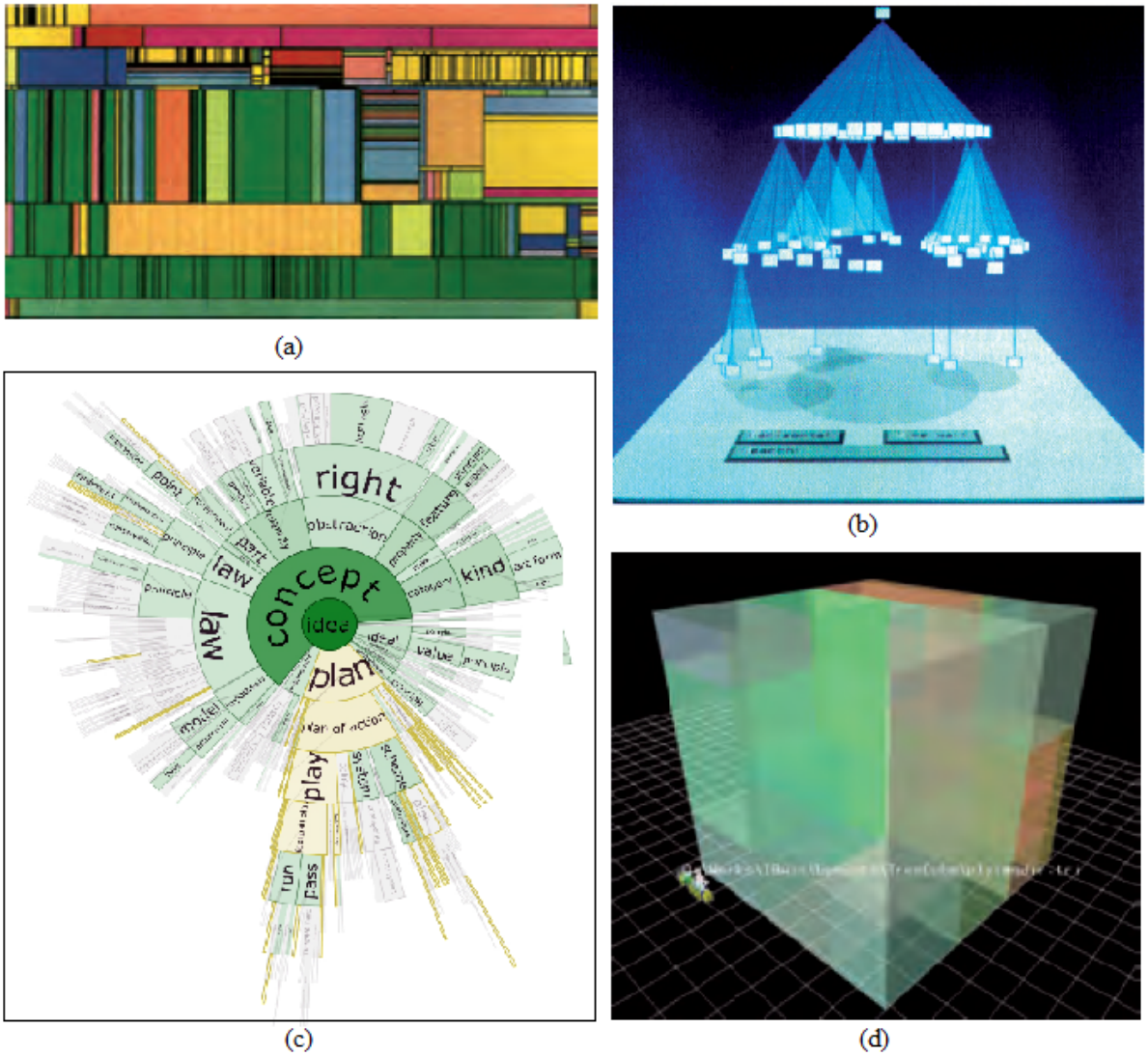


Fonte: Jaqueline Zaia, 2015

2.3.3 Dados Hierárquicos

Dados Hierárquicos são aqueles que possuem um relacionamento hierárquico, ou seja, os dados estão organizados e classificados de acordo com alguma medida de importância ou dependência entre si. Podem ser citados como exemplo de situações que se encaixam nesse conceito diagramas de árvores genealógicas, classes orientadas a objetos e organizações de arquivos e diretórios (STASKO; ZHANG, 2000). Esse tipo de conjunto de dados é comumente representado por meio de estruturas em forma de árvore.

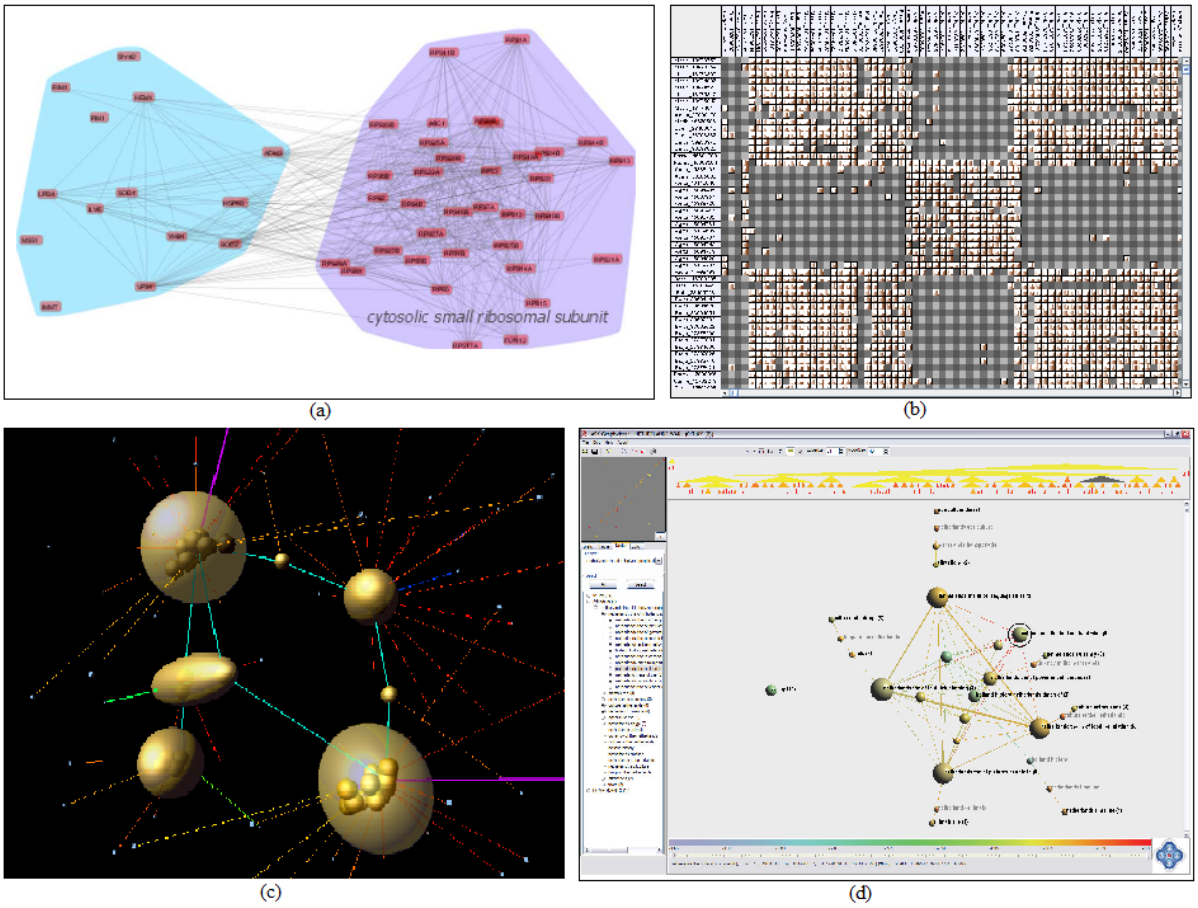
Figura 7 – Visualizações de Informação que representam hierarquias: (a) TreeMap; (b) ConeTree; (c) DocuBurst; e (d) Tree Cube.



Fonte: Jaqueline Zaia, 2015

São exemplos de visualizações de dados hierárquicos, representados graficamente na figura 7, a *Treemap* (JOHNSON; SHNEIDERMAN, 1991), *ConeTree* (ROBERTSON; MACKINLAY; CARD, 1991), *DocuBurst* (COLLINS; CARPENDALE; PENN, 2009) e *Tree Cube* (TANAKA; OKADA; NIIJIMA, 2003).

Figura 8 – Visualizações de informação que representam dados relacionados: (a) FUSE; (b) ZAME; (c) H-BLOB; e (d) Ask-GraphView.



Fonte: Jaqueline Zaia, 2015

2.3.4 Dados Relacionados

Uma coleção de dados relacionados, também conhecida como rede de dados, consiste em uma coleção de itens e relacionamentos conhecidos entre esses itens. É usual a representação de redes de dados por meio de grafos, cujos vértices representam os itens de dado e as arestas representam seus relacionamentos (O'MADADHAIN *et al.,* 2005). Mas também há outras formas como é o caso do ZAME (ELMQVIST *et al.*, 2008), ilustrado na figura 8 (b), que utiliza uma matriz de adjacência. Um exemplo usual de dados relacionados é a rede de citações em artigos, em que um artigo cita e é citado por outros artigos. A representação deste

tipo de dados pode ser observada na figura 8 em aplicações como FUSE (SEAH *et al.*, 2012), *H-BLOB* (SPRENGER; BRUNELLA; GROSS, 2000) e *Ask-GraphView* (ABELLO; VAN HAM; KRISHNAN, 2006).

2.4 Técnicas de Visualização

Além da dimensão e formato dos dados, outro conceito a ser considerado na criação de uma visualização é a sua dimensão visual, limitada às dimensões físicas mais a dimensão de tempo. Uma vez que a visualização possui essa limitação, um dos desafios da visualização de informação é utilizar da melhor forma possível essas dimensões para representar de forma abstrata os dados. O resultado final das técnicas de visualização em geral pode ser descrito como Bidimensional (2D) ou Tridimensional (3D), de acordo com a imagem gerada pela visualização. Além disso, é possível que a visualização conte com componentes visuais como legendas, menus, caixas de seleção e ícones (CATARCI *et al.*, 1997) que deem significância a componentes gráficos de diversas cores, formas, posições e tamanhos.

O posicionamento dos componentes depende do layout da visualização, que é a base na qual os itens visuais são distribuídos. Harger e Crossno (2010) classificam o layout de uma visualização de uma forma que remete à classificação de dados feita por Shneiderman (1996), dividindo a disposição dos itens em Layouts de Grafos, Layouts de Árvores, Layouts Tabulares e Layouts Georreferenciados.

Os Layouts Tabulares são aqueles ideais para a representação de dados n-dimensionais sem relacionamentos entre si. Pode-se mencionar como Tabulares o Layout de Dispersão, em que dados 2D ou 3D são representados como pontos coordenados no espaço; Layouts de Pizza, que dispõem os dados de forma circular e divide-os por setores; Layouts de Coordenadas Paralelas, que podem representar diversas dimensões em coordenadas dispostas paralelamente, em que cada espaço entre as coordenadas exibe uma dimensão dos dados; e Layouts de Barra, Linha e Área, que exibem os dados por meio de retângulos, linhas e áreas em um espaço coordenado. Para estes últimos, há também uma variação, chamada de Layout Empilhado, em que duas ou mais dimensões são empilhadas no layout.

Os Layouts de Grafos são utilizados na visualização de dados com relacionamentos e podem ser divididos em layouts circulares, onde os vértices são dispostos em função de círculo e as arestas cruzam o círculo fazendo as ligações entre os vértices; layouts radiais, em que é

escolhido um vértice concêntrico com o layout e todos os outros vértices são posicionados em função deste; layouts *force directed*, em que são utilizados algoritmos de aproximação e repulsão entre os vértices para o posicionamento; layouts hierárquicos, em que os vértices são arranjados de acordo com seus relacionamentos minimizando sobreposições e eliminando vértices menos importantes; e Layouts de Matrizes de Adjacência, que utiliza a representação gráfica de uma matriz de adjacência em grade em que cada célula da grade representa o relacionamento de dois vértices, que é ou não preenchido com cores.

Os Layouts de Árvore são tipicamente utilizados na representação de dados hierárquicos, em que a raiz da árvore são os dados de maior hierarquia e as folhas são os dados de menor. Podem ser citados os layouts *Tree Maps*, em que cada nível da árvore é representada como uma parte retangular da região ocupada pelo seu pai; Dendogramas, em que cada item da árvore é representado por um vértice posicionado juntamente com os outros itens do mesmo nível e logo após o item pai; e Árvores radiais, *Ballon Trees* e *Cosmic Tree*, similares ao Layout de Grafo Radial, que organizam os itens em linhas de círculos concêntricos, com a diferença que, por conta da natureza do dado hierárquico, não possibilita sobreposição de itens.

E, por fim, os Layouts Georreferenciados, que se tratam de layouts baseados em mapas cartográficos e frequentemente utilizam símbolos, figuras e cores para delinear coordenadas dentro do espaço. São Layouts Georreferenciados os *Map Overlays e Graduated Symbol Maps,* em que símbolos são dispostos sobre um mapa cartográfico 3D; *Choroplet Maps*, em que um mapa é colorido em função de um ou mais dados; *Dymaxion Maps*, em que mapas são desenhados em poliedros e projetados como uma imagem 2D; Cartogramas, em que projeções 2D têm o tamanho de seus itens definidos por uma variável do dado; e Globos 3D, que são nada mais que superfícies esféricas cobertas de imagem que podem ser visualizadas de qualquer ângulo.

2.4.1 Animação

Uma técnica frequentemente utilizada em visualizações é a animação, que envolve movimento de alguma das partes da visualização no qual forma e estrutura se desenvolvem no decorrer do tempo, transmitindo uma sensação de dinamismo. Essa prática, se bem utilizada, é capaz de atrair a atenção do usuário, manter sua motivação e facilitar a compreensão, aprendizagem, memória e comunicação (MOERE, 2004). Isso ocorre, pois, da mesma forma que o espaço físico na visualização é utilizado para transmitir o espaço metafórico dos dados,

mudanças reais da visualização no decorrer do tempo também podem ser eficazes para transmitir mudanças dos dados (TVERSKY; MORRISON; BETRANCOURT, 2002). Além disso, existem muitas formas de empregar a animação, que vão da simples interpolação local de estados da visualização até funções de movimentos complexos e dependentes de vários fatores da visualização.

A decisão de utilizar ou não a animação em uma visualização depende do que se deseja expressar, já que nem sempre é desejável manter a visualização presa à dimensão física do tempo. É o que ocorre nas animações de dados temporais, que são menos comuns que as representações estáticas (CRAIG; KENNEDY, 2003) pelo fato de que a exibição de dados em momentos diferentes impossibilita a tarefa de comparação e análise detalhada dos dados (TOMINSKI; SCHULZE-WOLLGAST; SCHUMANN, 2005). Em contrapartida, com o uso de técnicas simples como, por exemplo, o uso de transições entre estados da visualização, a animação pode ser eficaz quando utilizada a fim de manter o contexto e representar relacionamentos de causa e efeito (MICHOTTE, 1963).

2.5 Técnicas de Interação

As técnicas de interação são consideradas como qualquer forma que possibilite ao usuário interagir com a visualização e manipular os dados, seja por meio da interface gráfica utilizando componentes visuais como ícones e figuras, ou utilizando componentes textuais como caixas de pesquisa e filtros (CATARCI et al., 1997). Shneiderman descreve as formas de interação essenciais para qualquer ferramenta de visualização de informação baseada em sete tarefas básicas, são elas a Visão Geral, Visão Detalhada, Filtro, Detalhes por Demanda, Relacionamento, Histórico e Extração (SHNEIDERMAN, 1996).

A Visão Geral é dada quando é possível ao usuário visualizar de uma só vez toda a coleção de dados e ter uma noção do contexto que é observado. Já a Visão Detalhada é a tarefa que permite que, uma vez considerada a Visão Geral, o usuário decida interagir com uma porção específica da coleção, seja para focar em algum dado em específico ou para visualizá-lo mais detalhadamente (CARR, 1999). Ambas as tarefas são possíveis por meio da ferramenta de interação de Zoom, descrita a seguir, e estão frequentemente presentes nas técnicas que utilizam os conceitos de Focus+Context e Overview+Detail.

Uma vez que o usuário identifica a informação de seu interesse, devem vir os Detalhes por Demanda, geralmente em uma nova janela (CARR, 1999) ou possibilitando que o usuário clique sobre a visualização para que o detalhe seja exibido (AHLBERG; SHNEIDERMAN; 1994). Um Histórico das interações feitas com os dados também é outra tarefa importante, visto que proporciona ao usuário, que dificilmente encontrará as informações que precisa com uma única interação, uma forma de lembrar dos seus passos e a possibilidade de retornar e desfazer e repetir interações.

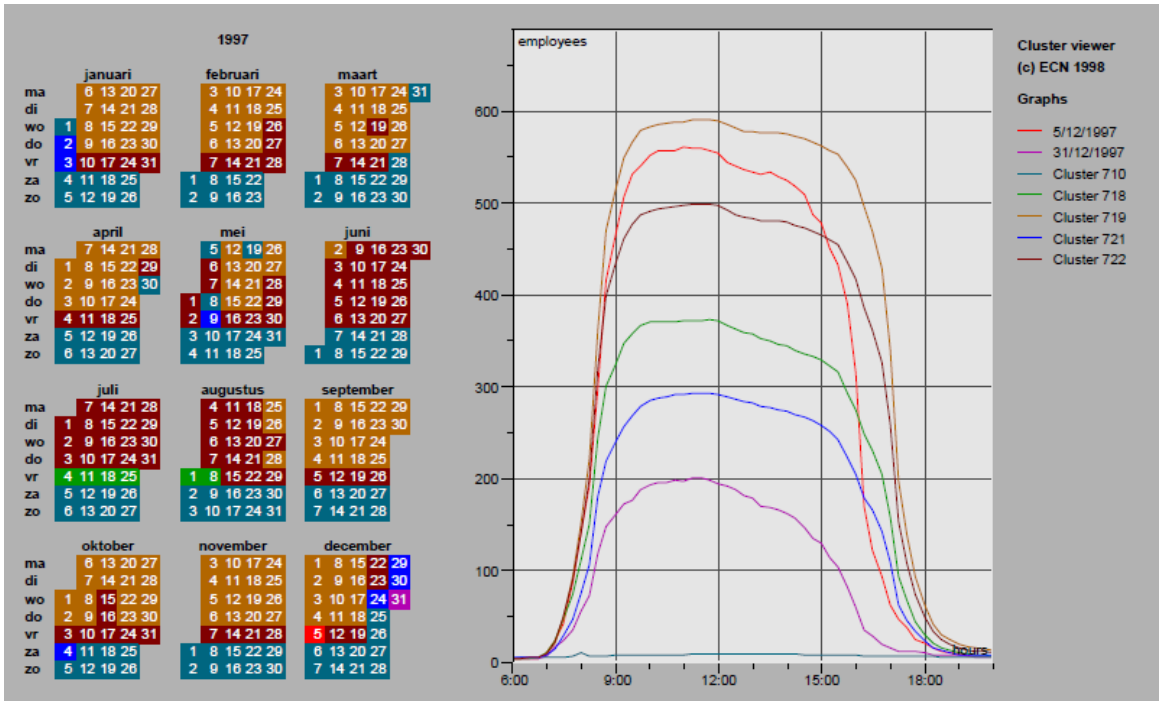
O Relacionamento entre os dados é a tarefa que possibilita ao usuário relacionar e visualizar pontos em comum entre itens buscando e conectando informações em diferentes telas. E, por fim, a Extração de dados deve possibilitar ao usuário a busca e seleção de subconjuntos de dados e seu salvamento para utilização futura.

A seguir, são descritas técnicas de interação frequentemente utilizadas e significativas para o contexto deste trabalho.

2.5.1 Visual Data Mining

Visual Data Mining ou Mineração Visual de Dados visa combinar a flexibilidade, criatividade e conhecimentos gerais do ser humano com a enorme capacidade computacional de armazenamento e processamento disponível hoje em dia, a fim de integrar o ser humano no processo de exploração de dados aplicando suas habilidades a grandes bases de dados, especialmente aquelas com ruídos e dados não-homogêneos ou sobre os quais se tem pouco conhecimento a respeito, e que é possível utilizar a intuição de forma mais proveitosa (KEIM, 2002). Visual Data Mining não é uma técnica por si só, mas um conceito que, para ser aplicado, pode utilizar um conjunto de técnicas e algoritmos que visam atender ao propósito específico de explorar informações visualmente e pode abranger diversas tarefas imprescindíveis na utilização de uma Visualização de Informação (SHNEIDERMAN, 1996).

Figura 9 - Visualização multifocal de clusters baseada em calendário



Fonte: Wijk e Selow (1999)

2.5.2 Interação Multifocal e *Information Hiding*

Outro artifício usado para lidar tipicamente com grandes séries de dados é mostrar somente a informação relevante para o usuário e esconder informação irrelevante, evitando que a visualização seja sobrecarregada e exibindo somente a informação importante (TOMINSKI; SCHULZE-WOLLGAST; SCHUMANN, 2005). Esse conceito de *Information hiding* se dá ao permitir que o usuário escolha quais informações deseja ver por meio de busca e interação com a visualização, sendo possível contextualizar, filtrar, aproximar, detalhar, relacionar e extrair os dados, além de manter histórico das operações realizadas (SHNEIDERMAN, 1996). Dessa maneira, os dados são redimensionados para atender à sua necessidade por meio da interação.

Esta capacidade de visualizar várias partes do conjunto de dados em um alto nível de detalhes ao mesmo tempo em que mantém contexto e noção de distância entre esses níveis é chamada de interação multifocal (JAVED; ELMQVIST, 2010). Esta abordagem permite interagir integrando o conceito de detalhe e contexto em várias escalas diferentes, dando suporte à tomada de decisão hierárquica do usuário (CARD; MACKINLAY; SHNEIDERMAN, 1999). Várias técnicas suportam o uso de interação multifocal, como é o caso exibido na figura 9 (WIJK; SELOW, 1999). Nela, é possível visualizar focos diferentes de informação: o

calendário mapeado com cores onde se escolhe um dia (foco no contexto anual, semanal e mensal), e um gráfico de linhas que mostra o resultado de um agrupamento dos padrões de quantidade de empregados presentes no período de um dia (foco no contexto diário e horário), combinados a fim de encontrar o melhor ponto de vista.

2.6 Arquitetura da Interface

Neste item são apresentadas as possíveis estruturas que uma ferramenta voltada para a criação de interfaces gráficas com o usuário (GUI) pode assumir, assim como as vantagens e desvantagens do uso de cada uma delas.

Quando se trata de ferramentas para auxílio à geração de uma GUI, pode-se notar que, apesar de efetivas na criação de interfaces tradicionais, estes são insuficientes quando é necessária a criação de uma interface que fuja dos componentes disponibilizados por eles; quer possua ela um componente simples como uma barra de rolagem ou objetos complexos como gráficos e telas interativas (BEDERSON; GROSJEAN; MEYER, 2004). Isso gera a necessidade de produção de grande quantidade de código para personalização e restrição da aplicação integrada, principalmente levando em conta a limitação das ferramentas, como Java Swing (SWING, 2015) e Java AWT (AWT, 2015), que possuem componentes básicos que não atendem a implementação de interfaces mais complexas; ou como Adobe Flash Builder e Macromedia Flash (ADOBE, 2015), difíceis ou impossíveis de serem estendidos e integrados com outros sistemas.

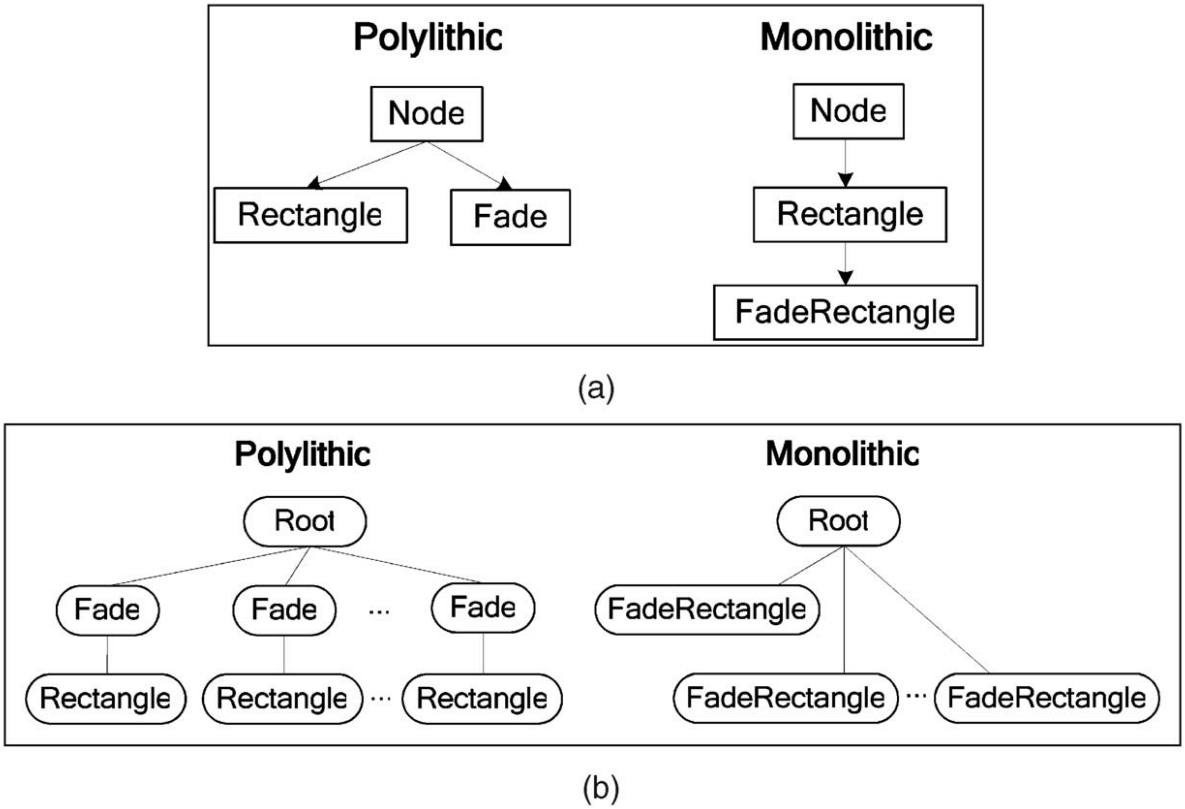
Por esses motivos, cresce cada vez mais a quantidade de *kits* de ferramentas de alto nível disponíveis para suprir essa necessidade e, junto com ela, a discussão sobre qual é a arquitetura ideal para melhorar o desenvolvimento de ferramentas com auxílio destes, evitar replicação de código e possibilitar a extensão e integração com outras ferramentas. Baseado na observação de diversos *kits* de ferramentas, Bederson, Grosjean e Meyer (2004) notaram a tendência nos *kits* de ferramentas 2D em serem implementados em função de uma estrutura mais concreta, em que os objetos tendem a ser mais parecidos com o mundo real. Em contrapartida, também foi observado por eles a ocorrência frequente em *kits* de ferramentas 3D de arquiteturas baseadas em árvores de especificações de componentes gráficos que geram uma exibição, conhecidos como *scene graphs* (BEDERSON; MEYER; GOOD, 2000).

Dada essa observação, são desenvolvidos dois *kits* de ferramentas – Jazz e Piccolo – e definidos por Bederson, Grosjean e Meyer (2004) dois estilos de arquiteturas de *kits* de ferramentas baseadas na sua composição: são eles Monolíticos e Polilíticos, descritos a seguir:

2.6.1 Monolítica

Kits de ferramentas monolíticos são definidos como aqueles que utilizam primariamente herança em tempo de compilação para estender uma funcionalidade. São caracterizados por possuírem poucas classes com as funcionalidades essenciais na construção de aplicações e tendem a ser complexos e difíceis de serem reutilizados, já que suas classes principais possuem um grande número de métodos e implementam hierarquias por herança para compor suas funcionalidades (Figura 9). São exemplos de *kits* de ferramentas monolíticos Pad++ (BEDERSON; HOLLAN, 1994), Tabula Rasa (FOX, 1998) e Jazz (BEDERSON; MEYER; GOOD, 2000).

Figura 10 – Exemplo uma hierarquia de classes Monolítica e Polilítica na representação de um retângulo: (a) Em tempo de compilação; e (b) Em tempo de execução



Fonte: Bederson, Meyer e Good (2000)

2.6.2 Polilítica

Kits de ferramentas polilíticos são definidos como aqueles que utilizam primariamente composição em tempo de execução para estender uma funcionalidade. Esta arquitetura, mais comumente aplicada na implementação de *kits* de ferramentas tridimensionais, tem como característica a criação de uma hierarquia de classes em que o componente visual é a raiz da árvore e as subclasses são componentes abstratos como linhas e formas (Figura 10 - a). Em tempo de execução, é gerada uma árvore visual que organiza os componentes na tela da forma como devem ser renderizados (Figura 10 – b). Esses *kits* de ferramentas tendem a ser mais fáceis de serem utilizados (KRASNER; POPE, 1988). Podem ser citados como exemplo Open Inventor (OPEN INVENTOR, 2015), Java3D (JAVA3D, 2015), VTK (SCHROEDER; MARTIN; LORENSEN, 2004) e Piccolo (BEDERSON; GROSJEAN; MEYER, 2004).

3 Metodologia e Trabalhos Correlatos

Neste capítulo são discutidas as metodologias de pesquisa e testes consideradas e realizadas no decorrer do trabalho, bem como os objetivos e justificativas de seus usos. Além disso, são discutidos também os trabalhos similares e suas vantagens e desvantagens.

3.1 Metodologia de Pesquisa

Este trabalho teve início com uma Revisão Exploratória sobre Visualização de Informação com evolução temporal em Redes Sociais, que inicialmente seria o foco do trabalho. Baseado nos resultados e motivado pela recorrência hoje em dia de dados que possuem referência temporal e que frequentemente são disponibilizados em grande escala, foi feita uma Revisão Sistemática (BIOLCHINI et al., 2007) das técnicas de visualização com referência temporal e técnicas de visualização de dados agrupados, ou *clusters*. A Revisão Sistemática é um método de pesquisa bibliográfica utilizado para encontrar estudos relacionados a um determinado tema, a fim de permitir uma avaliação relevante e controlada sobre toda a pesquisa (KITCHENHAM, 2004).

Como resultado da Revisão Sistemática, além das diversas técnicas para agrupamentos de dados e dados temporais, foi possível identificar os conceitos mais frequentemente utilizados, assim como assuntos pouco discutidos, como é o caso das técnicas utilizadas para validar a usabilidade das mais diversas técnicas de visualização de informação. Dentre os trabalhos, pôde-se notar a presença tanto de técnicas discutidas individualmente quanto sistemas, aplicações e *kits* de ferramentas, distribuídos em diversas aplicações, sendo as mais comuns o uso da visualização na área de genética e saúde, na busca e exploração de documentos e seus conteúdos, nas estatísticas de mercado de crédito e em históricos de relacionamentos como redes de citações em artigos e redes sociais.

Motivada pela baixa ocorrência de validações e testes significativos no desenvolvimento de ferramentas de visualização dado como resultado desta primeira revisão, foi realizada uma nova revisão exploratória sobre ferramentas e técnicas de Visualização de Informação, desta vez com maior abrangência e não mais limitado a aplicações com dados agrupados ou referenciados temporalmente. As informações foram coletadas de artigos de Visualização de Informação nos mais diversos contextos, e dados como tipos de ferramenta, tipos de dados, conceitos, técnicas de Visualização de Informação, técnicas de interação e técnicas de testes

foram mapeadas, agrupadas e comparadas. Com esta análise, pôde-se notar não só a deficiência na análise da usabilidade de técnicas de Visualização de Informação, mas também a limitação nas ferramentas existentes em proporcionar uma experiência de criação de visualizações adequada ao usuário da visualização, que deve ter conhecimento do domínio específico da informação, mas não necessariamente de programação ou visualização. Ou seja, a maioria das ferramentas de visualização analisadas considera conhecimento prévio em programação ou limita o usuário à utilização de técnicas, visualizações e dados pré-definidos; fato que, unido à pouca ocorrência de técnicas eficazes para validar Visualizações de Informação, pode dificultar o seu uso.

Ainda durante a revisão de literatura, foram desenvolvidos dois protótipos de ferramentas de visualização. O primeiro, utilizando o *kit* de ferramentas em Java VTK (SCHROEDER; MARTIN; LORENSEN, 2004), foi um protótipo com o objetivo de familiaridade com técnicas 3D, mais especificamente aquelas voltadas para redes sociais e *clusters*. Seu desenvolvimento se deu durante a primeira revisão exploratória, período em que o objetivo do trabalho ainda era diferente daquele aqui já apresentado. O segundo protótipo também foi desenvolvido em Java, porém baseado no *kit* de ferramentas prefuse (HEER; CARD; LANDAY, 2005) e já atende a alguns dos requisitos propostos na introdução. Neste protótipo, é possível criar visualizações básicas com interações limitadas como buscas textuais e filtros de forma interativa, porém funcionalidades de acesso a dados dinâmicos e interações mais complexas com grandes volumes de dados ainda não eram disponibilizadas.

No decorrer das implementações e revisões, houve uma mudança no escopo deste projeto, que deixou de abranger somente técnicas voltadas para a visualização de informações temporais e *clusters*, para passar a abranger visualizações de informação em geral. Além disso, o foco, que antes era na usabilidade e efetividade da utilização de uma Visualização de Informação, voltou-se principalmente para a usabilidade na criação interativa de uma Visualização. O objetivo deste trabalho, portanto, não é implementar uma técnica de visualização específica, mas desenvolver um sistema para criação interativa de visualizações de informação que provenha suporte à criação de técnicas diversas de acordo com a sua necessidade independentemente da natureza da informação que deve ser analisada.

3.2 Metodologia de Avaliação

Com o aumento da variedade de técnicas de visualização, torna-se importante a necessidade de aplicação de processos de avaliação mais significativos. Isso se justifica pelos possíveis problemas e vieses ao testar interfaces de visualização de informação com metodologias de teste de usabilidade não específicas para esse fim. Alguns dos problemas que podem ocorrer são: falhas na avaliação do ponto de vista da implementação de novas interfaces incompletas, o teste com usuários não apropriados para a tarefa ou com as tarefas erradas e a familiaridade dos usuários com as interfaces convencionais. A partir da discussão desses itens, são classificados os principais tipos de métodos de teste. Os primeiros são os testes formativos, em que são selecionados poucos usuários (3-5) e é observado o motivo pelo qual problemas ocorrem por meio da verbalização dos usuários ao usar a ferramenta (LANDAUER, 1988); os testes somatórios, que envolvem a medição de tarefas do ponto de vista da velocidade de duração e número de cliques ou erros por exemplo; e os testes de uso, caracterizados pela observação e registro da utilização prolongada pelo usuário (ANDREWS, 2006).

Já de acordo com Plaisant (2004), foram encontradas com mais frequência três áreas de avaliação similares: 1) experimentos controlados comparando elementos de design e ferramentas, 2) avaliação da usabilidade e 3) casos de estudo em situações reais. Assim como discutido por Andrews (2006), segundo seu trabalho, é importante considerar na avaliação a utilização de dados e tarefas o mais próximo possível da realidade para garantir a sua efetividade, uma vez que alguns dos reais benefícios da utilização de sistemas de visualização de informação são originados em tarefas mais complexas e difíceis de serem reproduzidas (ANDREWS, 2006).

Neste trabalho, são avaliadas, a princípio, apenas algumas metodologias de teste. A primeira, análise da usabilidade da técnica (MACEACHREN et al., 1998), se baseia em uma análise sequencial qualitativa para desenvolver uma compreensão geral das estratégias de exploração de dados que a ferramenta possibilita e o quanto isso facilita o usuário familiarizado com a área de aplicação. Para isso, são coletados comentários verbais durante todo o teste, assim como *logs* da ferramenta. O teste segue um protocolo de perguntas que induz o usuário a realizar tarefas pelo tempo de 45 minutos a uma hora. Knautz, Soubusta e Stock (2010) utilizam o método de avaliação de serviços SERVQUAL (PARASURAMAN; ZEITHAML; BERRY; 1988), que utiliza uma afirmação para medir as expectativas do usuário sobre o tipo de serviço

em geral a ser testado e uma para medir as percepções sobre o tipo de serviço específico utilizando uma escala. É feita a diferença entre as expectativas e as percepções para obter uma validação efetiva.

Já a avaliação feita por Chang et al. (2007) utiliza, após a aplicação de um pré-questionário e a conclusão de algumas tarefas, as perspectivas individuais de cada um dos participantes do teste. Similar ocorre na avaliação da ferramenta *AutoFocus* (FLUIT, 2005), porém a avaliação é feita por meio das opiniões recebidas via e-mail. Também é feita uma comparação quantitativa com outras ferramentas similares utilizando informações estatísticas como o tempo de realização de tarefas. Devido à baixa ocorrência de validação com o usuário, pode-se notar que há uma deficiência no desenvolvimento de novas técnicas e ferramentas de visualização.

3.3 Análise de Trabalhos Relacionados

Para este trabalho, são considerados trabalhos na área de visualização de informação, sejam eles voltados para o desenvolvimento de técnicas específicas, criações de ferramentas de visualização e também aqueles voltados para a análise de metodologias de teste e a avaliação de visualizações propriamente dita. Dada a diversidade de técnicas e ferramentas, e dado o objetivo deste trabalho, foram selecionadas as soluções consideradas mais significativas e que possuem características similares. Uma análise dessas soluções é feita em função de suas funcionalidades e detalhes técnicos, a fim de comparar as características interessantes e faltantes de cada um deles.

Além disso, são discutidas algumas bibliotecas e *kits* de ferramentas para auxílio no desenvolvimento de uma GUI para geração de visualizações interativamente, as quais foram exploradas e consideradas para serem utilizadas no desenvolvimento desta solução. Estas, portanto, não se tratam somente de trabalhos similares, mas de soluções auxiliares na construção de visualizações de informação.

3.3.1 Softwares Análogos

Dentre os trabalhos avaliados, é possível encontrar uma vasta gama de interfaces com o usuário que atendem às mais diversas áreas de aplicação com recursos próprios ou extensão de

recursos já existentes quando se trata de gerar visualizações de informação interativamente. A característica principal desses sistemas que é considerada no contexto desse trabalho é a de possibilitar que o usuário interaja com dados por meio de visualizações de informação interativas; sendo que a flexibilidade em relação aos possíveis tipos de dados, layouts e interações pode variar dependendo da tarefa que cada sistema se dispõe a dar suporte.

Tabela 1 – Ferramentas de geração de visualizações e suas características técnicas principais

Ferramenta	**Dimensão da Visualização**	**Suporte a Interações**	**Suporte a Funções**	**Formato de Entrada dos Dados**	**Contato Direto Com Dados**
InterVis	2D	Sim	Sim	CSV, XML e Banco de Dados	Não
Ask-Graph View	2D	Sim	Sim	Não fornecido	Não fornecido
Excel	2D	Não	Sim	CSV, XML, Bancos de Dados e outros	Sim
OpenedEyes	2D	Sim	Não	JSON	Não
Polaris	2D	Sim	Sim	Banco de Dados	Sim
Tableau	2D	Sim	Sim	CSV e Bancos de Dados	Sim
VisLink	2D e 3D	Sim	Não	CSV, GraphML, TreeML	Não
VisuMap	2D e 3D	Sim	Sim	ASCII, JPEG, SVG e Banco de Dados	Sim

Fonte: Jaqueline Zaia, 2016

Os softwares são analisados em função das características consideradas determinantes na escolha e utilização de ferramentas para propósitos específicos. Ou seja, a escolha por uma ferramenta de visualização de informação ideal vai depender sempre do problema que o usuário deseja resolver, que é diretamente relacionado aos tipos de dados e área de aplicação. As características observadas são a dimensão final da visualização gerada, o suporte a interações com os dados dentro da visualização e cálculos dinâmicos sobre as estruturas de dados visuais. Essas características definem o quão flexível é a visualização em relação a quais dados que se deseja visualizar. Por exemplo, caso seja necessário contar o número de ocorrências de determinada característica de uma base de registros, é necessário realizar um cálculo dinâmico no momento de interação da visualização que contabilize essa informação antes de exibi-la

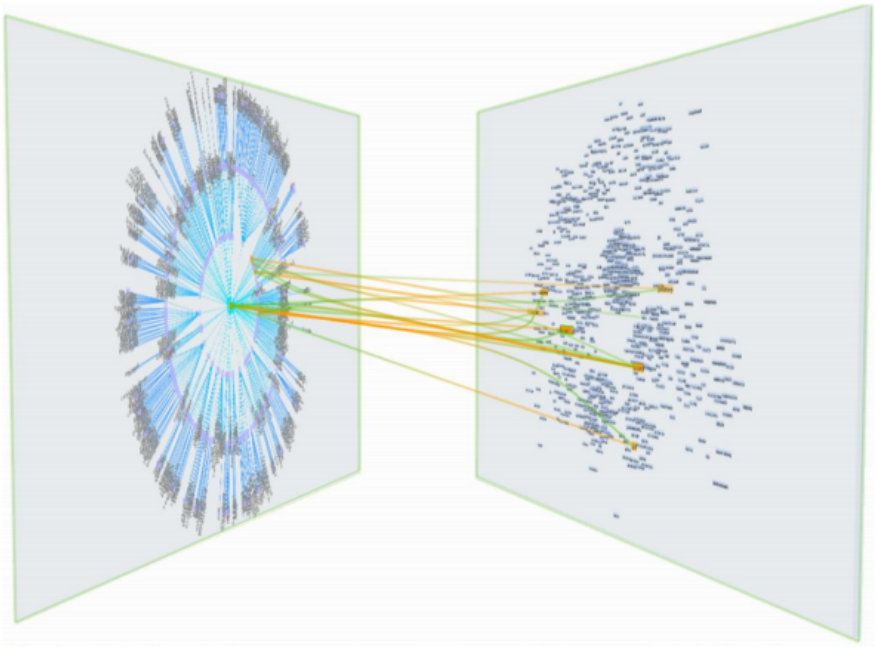
visualmente. Portanto, isso não se trata de um dado bruto disponibilizado, mas de uma informação extraída dos dados, que corresponde a uma pergunta do usuário ao sistema.

Além disso, são considerados os tipos de dados de entrada que os sistemas dão suporte; e se é necessário haver contato direto do usuário com os dados, o que implica a necessidade de exposição dos dados em sua forma bruta e possível necessidade de entendimento técnico das estruturas em que estes estão armazenados. Essas características definem a forma como o usuário precisa interagir com eles e tratá-los antes de tratar da criação e interação na visualização em si. As características, discutidas a seguir, estão consolidadas na tabela 1.

As ferramentas avaliadas são OpenedEyes (ALMEIDA; APOLINÁRIO, 2013), VisLink (COLLINS; CARPENDALE, 2007), Ask-Graph View (ABELLO; VAN HAM; KRISHNAN, 2006), Tableau (TABLEAU, 2016), Excel (EXCEL, 2016), Polaris (STOLTET; TANG; HANRAHAN, 2002) e VisuMap (VISUMAP, 2016). Além disso, uma comparação é feita com a solução proposta neste trabalho, a fim de delinear as características consideradas interessantes e em que difere esta solução das demais. Dentre as soluções avaliadas, uma das soluções de mercado mais populares é a ferramenta de geração de planilhas de dados Excel, em que uma das suas funcionalidades visa possibilitar a geração de visualizações de informação de forma interativa e não programática. Porém, apesar da sua rica interface, ela é a única que não permite interação com os dados diretamente por meio de suas visualizações de informação; enquanto todas as demais ferramentas possuem algum recurso para a edição de funcionalidades de interação, característica essa considerada importante para esta solução, de acordo com a bibliografia.

Já a dimensão final da visualização é predominantemente bidimensional dentre as ferramentas, com exceção da VisLink e VisuMap, que também dão suporte a visualizações tridimensionais. A VisLink possui um recurso diferencial – a possibilidade de comparar duas visualizações tabulares bidimensionais, formando uma visualização tridimensional que correlaciona os dados de ambas, conforme pode ser observado na figura 11. Essa característica mostra que grande parte dos problemas podem ser resolvidos satisfatoriamente com interfaces 2D.

Figura 11 – Visualização do VisLink combinando duas visualizações bidimensionais em um layout tridimensional com relacionamento entre os dados



Fonte: Collins e Carpendale (2007)

Com relação à possibilidade de realizar funções e cálculos dinâmicos sobre os dados durante a interação com a visualização, as ferramentas VisLink e OpenedEyes não possuem esse tipo de recurso, limitando as visualizações ao formato dos dados brutos. O Excel é um caso especial desta funcionalidade, pois possibilita a realização de funções complexas com os dados, porém essas não podem ser feitas interagindo diretamente com a visualização, mas atuando sobre as planilhas de dados que são fonte de dados para a mesma e que refletem dinamicamente nela. A importância desse recurso é reconhecida pelo fato de que o usuário não deve ter contato com os dados brutos, muito menos ter a necessidade de realizar tratamentos com ele antes de utilizá-lo na visualização. Logo, o InterVis deve atender a essa característica.

As entradas de dados nas ferramentas mais comuns são o CSV (Comma Separated Values) e o relacionamento com Bancos de Dados. Porém há recorrência de outros formatos populares com XML (Extensible Markup Language), JSON (JavaScript Object Notation); e menos recorrentes como GraphML, TreeML e formatos de imagem e vetor como JPEG (Joint Photographic Experts Group) e SVG (Scalable Vector Graphics). Dessa forma, entende-se que há a necessidade de diversificação de formatos de dados de entrada; logo, o InterVis fornece a interpretação de duas formas de entrada – CSV e Bancos de Dados – que podem ser expandidas na medida do necessário. A ferramenta Ask-Graph View não documenta os tipos de dados de

entrada, assim como a possível necessidade de contato direto com eles. Com exceção do VisLink e OpenedEyes, que fazem uso de metadados somente, as demais ferramentas fazem necessário o contato direto do usuário com os dados durante a criação das visualizações, excluindo a possibilidade de preservação de dados que possam ser sensíveis. Conforme discutido anteriormente, considera-se necessária a abstração dos dados, uma vez que o usuário final não deve ser exposto a possíveis dados sensíveis no contexto da aplicação.

Tabela 2 – Ferramentas de Visualização de Informação e seus possíveis layouts

Ferramenta	**Tabular**	**Grafo**	**Árvore**	**Georreferência**
InterVis	Sim	Sim	Sim	Não
Ask-Graph View	Não	Sim	Sim	Não
Excel	Sim	Não	Não	Não
OpenedEyes	Sim	Não	Não	Sim
Polaris	Sim	Não	Não	Sim
Tableau	Sim	Não	Não	Sim
VisLink	Sim	Sim	Sim	Não
VisuMap	Sim	Sim	Sim	Sim

Fonte: Jaqueline Zaia, 2016

Além das características técnicas das ferramentas, outro ponto de vista é o das funcionalidades presentes na visualização. Para delinear os tipos de visualizações presentes, uma análise é conduzida em função dos tipos de layout que cada ferramenta implementa, conforme a tabela 2. É possível observar que, com exceção do Ask-Graph View, todas as ferramentas analisadas dão suporte a algum tipo de layout tabular. Pode-se notar também a predominância de layouts de Grafo e Árvore em três ferramentas – AskGraph View, VisLink e VisuMap. Nestes, os layouts tabulares são menos explorados, enquanto os recursos para layouts de Grafos e Árvores são mais robustos e predominantes.

Isso mostra a especialização das ferramentas em função dos tipos de dados que se deseja exibir. Grafos e Árvores, por exemplo, exibem dados com estruturas muito similares, que são em função de relacionamentos entre dados – no caso a árvore representa um tipo específico de relacionamento, que é a hierarquia. Por isso, ferramentas que suportam a visualização de grafos, em geral, também suportam hierarquias e vice-e-versa. Os layouts georreferenciados, por sua vez, podem ser representados por mais da metade das ferramentas e são caracterizados por tipos de dados específicos que agregam alguma referência espacial. Das ferramentas que não dão

suporte a layouts georreferenciados, duas são voltadas para a visualização de dados relacionados e hierárquicos – VisLink e Ask-Graph View – e a terceira é o Excel, que dá suporte somente a dados tabulares. Podemos notar as especificidades de cada sistema por tipos de layouts diferentes, portanto, no InterVis, foi escolhido por implementar os três layouts mais diversos de forma básica, que pode ser estendida, assim como podem ser criados novos tipos de layout, como é o caso de layouts georreferenciados que não são dispostos no protótipo.

É possível notar uma linha evolutiva dentre as ferramentas, que estendem e complementam funcionalidades umas das outras com o passar do tempo. Esse é o caso das ferramentas Polaris e Tableau, ambas representativas na área de visualização de informação, em que a segunda é uma evolução comercial com a arquitetura e funcionalidades da primeira. Apesar de apresentarem interfaces diferentes, ambas compartilham a característica de serem ferramentas de visualização de informações multidimensionais e permitem a combinação e interação entre diversos layouts de informações relacionadas. Além disso, é observado que, de todas as ferramentas, apenas a OpenedEyes permite que suas funcionalidades sejam estendidas de maneira programática, característica interessante considerando a natureza evolutiva das ferramentas de visualização de informação e que escolheu-se possibilitar no InterVis.

3.3.2 *Kits* de Ferramentas de Suporte ao Desenvolvimento

Os *kits* de ferramentas de suporte ao desenvolvimento são em geral bibliotecas e conjuntos de algoritmos que implementam diversas funcionalidades úteis no suporte ao desenvolvimento de GUIs em determinadas linguagens de programação. Não são, portanto, ferramentas de pronto uso, visto que demandam o conhecimento técnico e tempo de um programador para criar visualizações utilizáveis pelo usuário final. Apesar de a solução proposta nesse mestrado não ser um *kit* de ferramentas, ela possui e estende características desses *kits*; e por isso tais trabalhos são considerados parte importante da metodologia de pesquisa e desenvolvimento do InterVis. A fim de mapear as características técnicas mais significativas dentre os *kits* de ferramentas analisados, é feita uma análise em função de algumas características, como as linguagens de programação em que os *kits* de ferramenta são disponibilizados e quais dimensões finais de visualização estes implementam. Além disso, é considerado se o *kit* de ferramentas dá suporte a interações dentro da visualização e realiza funções e cálculos dinâmicos sobre as estruturas de dados visuais, como agrupamento, soma, máximo e média; possibilitando flexibilidade na análise dos dados em relação aos dados

textuais de entrada. Por fim, são observados quais os formatos de dados tratados na solução, ou seja, para quais formatos de dados o *kit* de ferramentas possui função de abstração dos dados em sua forma nativa. O levantamento de tais informações para cada um dos *kits* de ferramentas considerados é sumarizado na tabela 3.

Tabela 3 - *Kits* de Ferramentas e suas características técnicas principais

Ferramenta	**Linguagem**	**Dimensão da Visualização**	**Suporte a Interações**	**Suporte a Funções**	**Formato de Entrada dos Dados**
Flare	ActionScript	2D	Sim	Sim	CSV, GraphML e JSON
Google Vis	JavaScript	2D	Sim	Não	CSV, JSON e XML
Improvise	Java	2D	Sim	Não	CSV, XML,DBF, SHP e Bancos de Dados
JUNG	Java	2D e 3D	Sim	Sim	GraphML e NET
Prefuse	Java	2D e 3D	Sim	Sim	CSV, GraphML e TreeML
Processing	Java	2D e 3D	Sim	Sim	CSV, XML e outros
ProtoVis	JavaScript	2D	Sim	Sim	Não possui
The InfoVis Toolkit	Java	2D	Sim	Sim	CSV, TM3, DOT, TreeML e GraphML
Tulip	C++ e Python	2D e 3D	Não	Não	CSV, GML, Tulip e DOT
VTK	C++, Python, Java, e Tcl	2D e 3D	Sim	Sim	VTK, XML e outros

Fonte: Jaqueline Zaia, 2016

As ferramentas analisadas são Flare, Google Vis, Improvise, JUNG, prefuse (HEER; CARD; LANDAY, 2005), Processing, ProtoVis, The InfoVis Toolkit (FEKETE, 2004), Tulip e The Visualization Toolkit – VTK (SCHROEDER; MARTIN; LORENSEN, 2004). Dentre elas, todas disponibilizam soluções para visualização de informação em forma de bibliotecas de código, porém apenas Improvise e Tulip possuem alguma interface gráfica disponível para manusear o código. Além disso, enquanto todas as ferramentas dão suporte à criação de visualizações bidimensionais, menos da metade dá suporte a visualizações tridimensionais.

Em relação ao suporte à interação e à aplicação de funções e cálculos sobre os dados; a grande maioria das ferramentas possui algum tipo de funcionalidade de interação, e apenas três não possibilitam nenhum tipo de alteração sobre os dados. Já o suporte ao processamento da entrada de dados é fornecido por todas as ferramentas para algum tipo de arquivo, exceto o ProtoVis, que não documenta nenhum módulo de processamento dos dados de entrada. Dentre os formatos de arquivo suportados, a grande maioria dá suporte a arquivos de texto delimitados como CSV, porém há recorrência de outros formatos populares com XML e JSON; e também de formatos próprios como o VTK e o Tulip. Somente o Improvise possui módulo conexão com bancos de dados.

Tabela 4 - *Kits* de Ferramentas e os tipos de layout aos quais eles dão suporte

Ferramenta	**Tabular**	**Grafo**	**Árvore**	**Georreferência**
Flare	Sim	Sim	Sim	Não
Google Vis	Sim	Não	Não	Sim
Improvise	Sim	Sim	Sim	Não
JUNG	Não	Sim	Sim	Não
Prefuse	Sim	Sim	Sim	Não
Processing	Sim	Sim	Sim	Sim
ProtoVis	Sim	Sim	Sim	Sim
The InfoVis Toolkit	Sim	Sim	Sim	Não
Tulip	Sim	Sim	Sim	Sim
VTK	Sim	Sim	Sim	Sim

Fonte: Jaqueline Zaia, 2016

Além das características técnicas dos *kits* de ferramentas, outro ponto de vista é o das funcionalidades presentes na visualização. Para delinear os tipos de visualizações presentes, uma análise é conduzida em função dos tipos de layout que cada ferramenta implementa, conforme a tabela 4. Nela é possível perceber, assim como nos softwares análogos, um padrão em que os *kits* de ferramentas são predominantemente voltados a layouts de grafos, ou voltados para dados multidimensionais tabulares. Pode-se interpretar essa especialização como um reflexo dos tipos de dados suportados e suas respectivas estruturas de dados nativas, uma vez que grafos mostram relacionamentos entre dados enquanto dados tabulares mostram suas características somente. Enquanto a grande maioria das ferramentas dá suporte a dados tabulares, cujas estruturas de dados tendem a ser mais simples e similares a tabelas, somente metade delas atende a algum tipo de visualização georreferenciada, sendo que as ferramentas

que atendem possuem poucos recursos relacionados a elas, como poucos layouts e formas de interação. Além disso, apesar de a grande maioria das ferramentas dar suporte a grafos, pode-se notar que algumas delas possuem recursos superficiais em relação a outras, que possuem uma maior diversidade de algoritmos e layouts voltados para o tratamento de dados relacionados em grafos. Esse é o caso do JUNG, cujas visualizações implementadas são majoritariamente para grafos; enquanto o Flare possui funcionalidades para visualização de grafos menos elaboradas em relação às suas demais visualizações disponíveis.

Uma vez que esses trabalhos se tratam de *kits* de ferramentas, a ausência de algum dos recursos discutidos neles pode ser suprida pela combinação de outras bibliotecas ou até mesmo seu código pode ser estendido com novas funcionalidades, não se tratando, portanto, de uma limitação. Dentre os dez *kits* de ferramentas analisados, na etapa seguinte desse trabalho, três foram testados e avaliados detalhadamente do ponto de vista da arquitetura, módulos e classes, a fim de serem considerados para auxiliar no desenvolvimento do InterVis.

3.3.2.1 The Visualization Toolkit

O VTK, ou *The Visualization Toolkit*, é uma biblioteca de classes C++ open source para desenvolvimento de gráficos e visualizações 3D. Possui uma arquitetura polilítica que proporciona fácil utilização e reutilização nas aplicações, uma vez que segue a filosofia de que *kits* de ferramentas devem possibilitar o desenvolvimento de aplicações complexas a partir de pequenas partes bem definidas e com interfaces simples. Carrega também consigo algumas características determinantes como simplicidade de uso, uso de linguagem interpretada e portabilidade entre plataformas (SCHROEDER; MARTIN; LORENSEN, 2004).

3.3.2.2 prefuse

O prefuse é um *kit* de ferramentas em Java desenvolvido para criação de visualizações dinâmicas 2D e 3D, e propõe abstrações dos dados que possibilitam ao programador a geração de customização rápida de componentes visuais interativos, tanto para dados tabulares quanto para dados relacionados. Ao invés de proporcionar blocos visuais prontos, simplifica a criação de técnicas já conhecidas de visualização, como layouts, itens visuais e algoritmos de distorção proporcionando essas estruturas de forma que é possível integrá-las facilmente e criar novas técnicas em domínios de aplicação específicos (HEER; CARD; LANDAY, 2005).

3.3.2.3 The InfoVis Toolkit

O InfoVis Toolkit é um *kit* de ferramentas desenvolvido para dar suporte à criação, extensão e integração de componentes avançados de Visualização de Informação 2D em aplicações que utilizam Java Swing (SWING, 2015). Ele disponibiliza estruturas que proporcionam um retorno rápido em loops constantemente necessários em consultas dinâmicas. Possui componentes visuais definidos que podem ser manuseados com o auxílio de um framework que facilita a criação de interfaces de visualização. Além disso, o InfoVis também possibilita a extensão de seus componentes genéricos (FEKETE, 2004).

3.4 Considerações Finais

É possível observar a diversidade de técnicas e ferramentas já existentes dentre os trabalhos considerados neste projeto, que se vão desde algoritmos de renderização em tela até tratamento de dados para abstração visual, considerando as mais variadas adversidades na criação de visualizações de informação. Isso mostra uma área de pesquisa que, apesar de ser recente, tem muito a oferecer em seus pouco mais de 30 anos de existência; e se torna cada vez mais necessária na medida em que novas tecnologias emergem, como é o caso de tecnologias big data e outras ferramentas de processamento de grandes massas de dados voltadas principalmente para Mineração de Dados e Business Intelligence.

Em consequência do seu crescimento, pode-se observar a demanda por novas metodologias de validação de ferramentas dessa natureza, que precisam se adaptar cada vez mais na medida que aumenta a diversidade de técnicas e recursos em ferramentas de visualização de informação. Pode-se perceber, contudo, que a velocidade com que novas técnicas e ferramentas surgem não é acompanhada pela sua respectiva avaliação, principalmente do ponto de vista da usabilidade, ainda que outras características da visualização também careçam de atenção, como a questão de desempenho na renderização das imagens em visualizações de informação. Essas características são frequentemente avaliadas em escopo acadêmico, mas faltam testes mais profundos nas possíveis áreas de aplicação, a fim de tornar a visualização de informação algo que seu próprio conceito propõe, que é ser parte naturalmente necessária de uma solução.

A partir dessa análise, é válido afirmar que a área de Visualização de Informação tem uma grande diversidade de técnicas e recursos, mas ainda necessita de mais atenção no que diz respeito à usabilidade e utilidade envolvidas na sua aplicação em soluções reais.

4 Solução

Esse trabalho visa a permitir que o usuário final de visualização de informação – com conhecimento no domínio da informação, mas não necessariamente de programação ou visualização – possa gerar e editar visualizações de acordo com a sua necessidade independentemente da natureza da informação que deve ser analisada. Para tanto, baseado na análise da literatura e dos problemas discutidos anteriormente, é proposto como solução a idealização, desenvolvimento e avaliação do protótipo de um sistema para criação e exploração interativas de visualizações de informação a partir de dados dinâmicos, denominado InterVis. O protótipo é desenvolvido na linguagem de programação Java e com o apoio do *kit* de ferramentas para geração de visualizações prefuse (HEER; CARD; LANDAY, 2005), tanto no que diz respeito ao código quanto no modelo da arquitetura.

Esta solução é direcionada para a geração de visualizações de informação a serem utilizadas em problemas gerais do dia a dia de qualquer área de aplicação. No entanto, isso não implica que o InterVis deva ser suficiente para a geração de visualizações de informação mais completas e para problemas específicos, como é o caso de problemas que demandam visualizações com múltiplos layout ou que necessitam de acesso direto aos dados brutos. Para tanto, existem ferramentas mais específicas tanto em contexto de aplicação com em técnicas de visualização, que fazem um papel que esta solução não se propõe a fazer.

4.1 Requisitos

Os requisitos deste trabalho, levantados com base na análise da literatura e em discussões com profissionais da área, são divididos em duas perspectivas, sendo elas a criação e a exibição de representações visuais interativas:

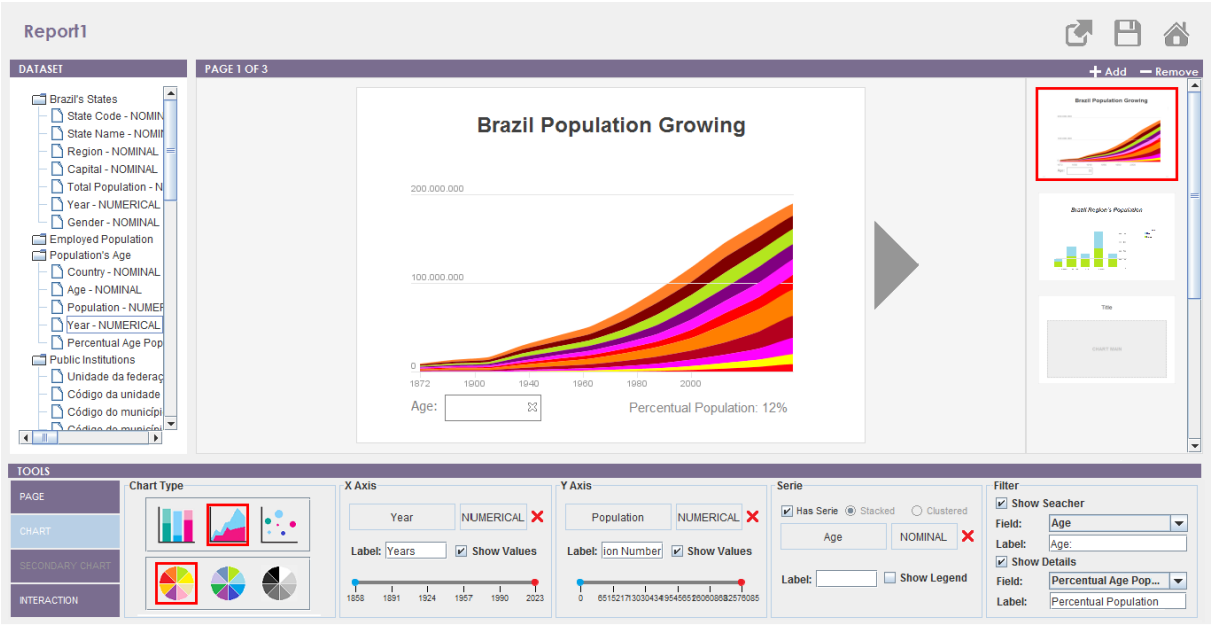
4.1.1 Criação Interativa

Esta perspectiva tem como requisito principal a geração interativa e salvamento de visualizações de informação, que pode ser feita por meio de dados estáticos ou dinâmicos com referência a um banco de dados relacional. Ela é composta primariamente de três painéis, conforme pode ser visto na figura 12: o primeiro à esquerda trás os dados abstraídos em metadados, o segundo em cima à direita trás a prévia da visualização e o último abaixo exibe

os menus para edição da visualização de acordo com os componentes selecionados na visualização. As características desta perspectiva devem atender aos seguintes requisitos:

1. Criar e salvar relatórios interativos apresentáveis como *slides*;
2. Criar visualizações de informação contidas nos relatórios;
3. Permitir a visualização de dados estáticos e dinâmicos, tanto de arquivos quanto de bases de dados relacionais;
4. Omitir os dados em sua forma bruta;
5. Selecionar interativamente modelos, configurações, dados e eventos na criação das visualizações;
6. Pré-visualizar o relatório durante o processo de criação a visualização;
7. Editar relatórios previamente criados;
8. Proporcionar um processo de criação de visualizações eficiente e útil de forma fácil e satisfatória para o usuário.

Figura 12 - Perspectiva de Geração Interativa de Visualizações de Informação



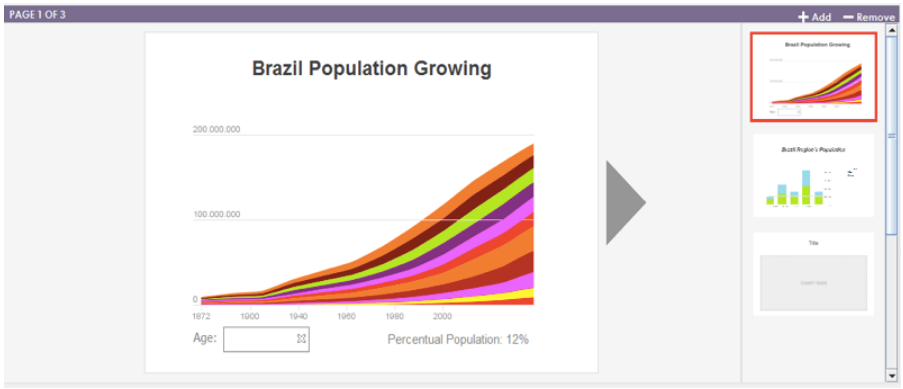
Fonte: Jaqueline Zaia, 2016

4.1.1.1 Criar e salvar relatórios interativos apresentáveis como slides

O primeiro requisito se trata da funcionalidade de permitir que o usuário crie, modifique e salve relatórios interativos, de forma que seja possível abri-los novamente no futuro. Além disso, o relatório assume um formato popularmente utilizado em diversas ferramentas, como a Microsoft PowerPoint (POWERPOINT, 2016), que utiliza páginas de slides similares a um

livro, onde as visualizações são dispostas nas páginas, de acordo com o que o usuário preferir montar. O usuário também pode navegar pelas páginas utilizando as duas setas nas laterais da página ou a subseção de visão geral de todo o relatório clicando sobre as miniaturas, conforme o que pode ser observado na figura 13. Os botões para salvar e visualizar, assim como o caminho para o menu inicial, estão no canto superior direito da perspectiva.

Figura 13 - Painel de prévia da visualização sendo criada

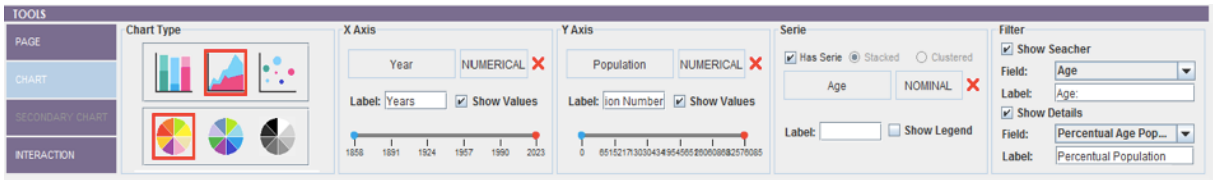


Fonte: Jaqueline Zaia, 2016

4.1.1.2 Criar visualizações de informação contidas nos relatórios

Uma vez criado um relatório no modelo de livro, o InterVis permite que o usuário crie até duas visualizações, relacionadas ou não, em cada página do relatório. Cada página do relatório trata os dados vinculados às suas visualizações de forma individual, ou seja, é possível vincular interações entre visualizações somente quando essas estão na mesma página do relatório. Isso pode ser feito na aba “Page” localizada no painél, no lado esquerdo, conforme a figura 14. Nela é possível definir qual é o layout de cada página. Já a relação de causa e efeito na interação entre os gráficos pode ser atribuída, ou não, na aba “Interaction”, localizada no mesmo lugar.

Figura 14 - Painel de configurações das visualizações

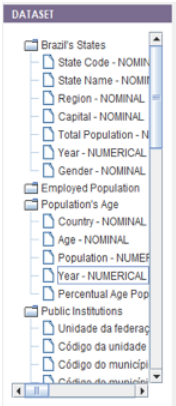


Fonte: Jaqueline Zaia, 2016

4.1.1.3 Permitir a visualização de dados estáticos e dinâmicos, tanto de arquivos quanto de bases de dados relacionais

Ao criar um relatório, o usuário pode utilizar bases de dados tanto estáticas, em que os dados são salvos junto com o relatório, quanto dinâmicas, em que somente uma referência dos dados é salva. E possível utilizar dados de duas fontes: no formato de arquivo CSV, que deve ser carregado no InterVis para que seus dados sejam exibidos na criação da visualização; e por meio de uma conexão com banco de dados relacional, em que é necessária a configuração de um banco de dados para a disponibilização das informações. Essa configuração dos dados é feita fora das perspectivas de criação em uma tela inicial, de forma que o usuário não tenha contato com os dados em sua forma bruta.

Figura 15 - Painel de conjuntos de dados abstraídos em metadados



Fonte: Jaqueline Zaia, 2016

4.1.1.4 Omitir os dados em sua forma bruta

Dada a forma como os dados são exibidos, o InterVis garante que o usuário não terá acesso direto aos dados brutos. Isso ocorre, pois, uma vez que uma fonte de dados é inserida, ela é exibida por meio dos seus metadados no formato de um menu hierárquico vertical, como pode ser observado no menu lateral esquerdo “Dataset” na figura 15. Logo, no caso de uma base de dados relacional ser carregada, por exemplo, serão exibidas as suas tabelas no primeiro nível e suas colunas com seus respectivos formatos no segundo nível do menu, sendo que os dados em si são omitidos.

4.1.1.5 Selecionar interativamente modelos, configurações, dados e eventos na criação das visualizações

Ao criar o relatório de visualizações, é possível editar suas propriedades visuais no menu inferior, conforme o que pode ser observado na figura 14. Modelos de layout e combinações de cores são disponibilizados para que o usuário escolha na subseção “Chart Type”. Além disso, há subseções para cada um dos eixos, chamadas “X Axis” e “Y Axis”, que são mandatórias de preenchimento, e para as séries de dados há uma subseção chama “Serie”, que é opcional para o caso de ser necessário exibir dados com algum tipo de classificação. Nessas subseções há características como o intervalo de dados, rótulos e valores exibidos que podem ser escolhidas, uma vez que é atribuído um dado a ele, o que se dá quando o usuário arrasta um item de dado do menu “Dataset” até uma das subseções. Por fim, há as subsseções que caracterizam cada uma das funcionalidades de interação que se encontram nas abas “Chart” e “Interaction”, são elas as configurações para utilização de filtros textual e gráfico – “Filter” e “Graphical Filter”, zoom – “Zoom” e “Semantic Zoom” e detalhes textuais sob demanda “Details”. Todas as seções das abas “Chart” e “Interaction” são exibidas e funcionam somente quando há pelo menos uma visualização válida criada, caso contrário somente as propriedades básicas da página, que se encontram na aba “Page”, como nome, disposição dos itens e aparência da página.

4.1.1.6 Pré-visualizar o relatório durante o processo de criação da visualização

Durante todo o processo de criação das visualizações, é possível visualizar a prévia do relatório no painel central, composto de uma visualização Overview+Detail do relatório, onde é possível visualizar uma página do relatório por vez no painel maior da esquerda (detalhe), e todas as páginas do relatório (visão geral), conforme já descrito no item 4.1.1.1 e exibido na figura 13.

4.1.1.7 Editar relatórios previamente criados

Uma vez criado e salvo um relatório, o usuário pode modificá-lo no futuro. Para isso, é necessário abrir o InterVis e, por meio dele, abrir para edição o relatório salvo. No caso de um relatório utilizar uma base de dados de referência, é necessário abrir também essa base de dados, seja ela um arquivo ou uma conexão com banco de dados, caso contrário, as configurações se mantém, mas não haverá dados para serem exibidos.

4.1.1.8 Proporcionar um processo de criação de visualizações eficiente e útil de forma fácil e satisfatória para o usuário

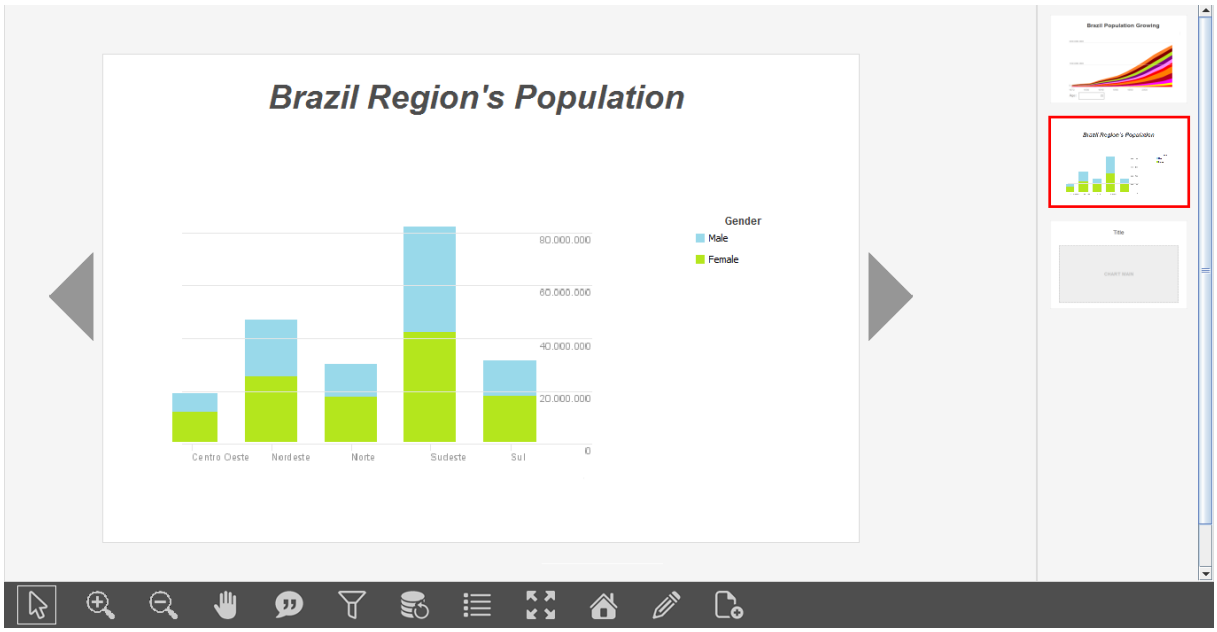
Este requisito é referente à validação do InterVis e visa a garantir que o sistema não só seja concluído baseado nos demais requisitos, mas que também atenda às necessidades do usuário de forma fácil, satisfatória e eficiente. A forma como a avaliação é conduzida para atender a esse requisito é descrita detalhadamente na Metodologia Empregada do capítulo de Testes.

4.1.2 Visualização Interativa

A perspectiva de visualização a visa proporcionar ao usuário a experiência de interagir com uma Visualização de Informação previamente criada na perspectiva de Criação Interativa e apresentada no formato de relatório, conforme os requisitos a seguir:

1. Visualizar um relatório previamente criado;
2. Possibilitar interação com as visualizações de um relatório;
3. Proporcionar uma visualização de informação útil e eficiente como resultado que possibilite ao usuário identificar informações e responder perguntas sobre os dados de forma fácil e satisfatória.

Figura 16 - Perspectiva de Exibição das Visualizações de Informação



Fonte: Jaqueline Zaia, 2016

4.1.2.1 Visualizar um relatório previamente criado

Uma vez criado e salvo um relatório, é possível abri-lo na perspectiva de exibição. Essa perspectiva é composta de um painel principal composto de uma visualização Overview+Detail do relatório, onde é possível visualizar uma página do relatório por vez no painel maior da esquerda (detalhe), e todas as páginas do relatório (visão geral), conforme pode ser observado na figura 16.

4.1.2.2 Possibilitar interação com as visualizações de um relatório

Nessa perspectiva, é possível que o usuário interaja com a visualização previamente criada, desde que ela possua funcionalidades de interatividade. Essas funcionalidades são atribuídas à visualização durante a criação da visualização, na perspectiva de edição. Se nenhuma funcionalidade é atribuída, os botões exibidos na barra inferior de funcionalidades estarão desativados. Nessa barra também se encontram os botões para mudança de perspectiva, edição, criação de nova visualização, exibição em tela cheia, atualização dos dados dinâmicos e anexo de comentários, conforme a figura 16.

4.1.2.3 Proporcionar uma visualização de informação útil e eficiente como resultado que possibilite ao usuário identificar informações e responder perguntas sobre os dados de forma fácil e satisfatória

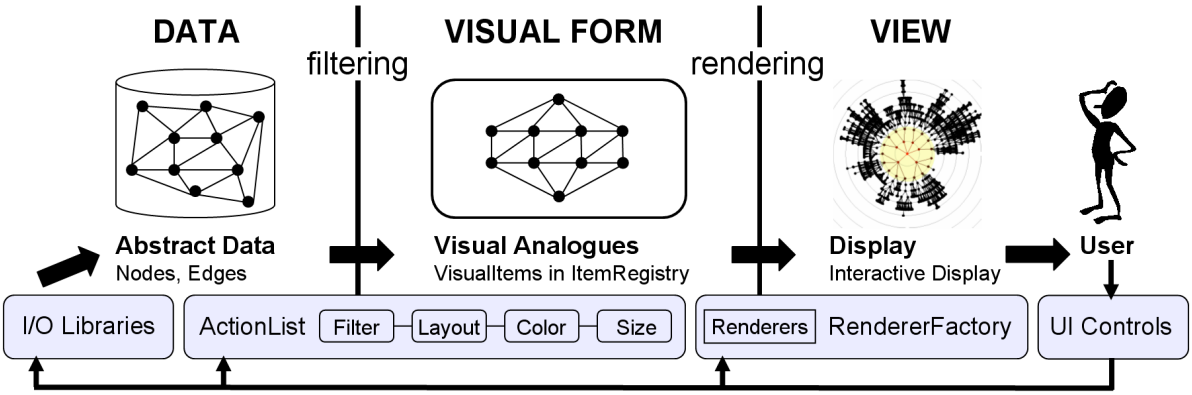
Esse requisito se refere à usabilidade da visualização de informação gerada pelo InterVis, assim como a forma que é possível abri-la, atualiza-la e interagir com ela. Para garantir seu cumprimento, a Metodologia Empregada e todo o processo que é conduzido para avaliação do InterVis se encontram no capítulo de Testes.

4.2 Implementação

Esta solução é desenvolvida na linguagem de programação Java. O funcionamento da aplicação segue o modelo proposto pelo prefuse na figura 17 (HEER; CARD; LANDAY, 2005) em que os dados são estruturados do banco de dados para estruturas abstratas. Neles são aplicados os filtros e ações, como agrupamentos e cálculos sobre os dados; em seguida são criadas as estruturas visuais, que atribuem aos dados propriedades visuais como posição, cor e

tamanho. Por fim, estas estruturas são renderizadas e exibidas na forma de interfaces interativas, que podem ser alteradas a qualquer momento, em função das ações do usuário.

Figura 17 - Modelo de interações entre os módulos do prefuse



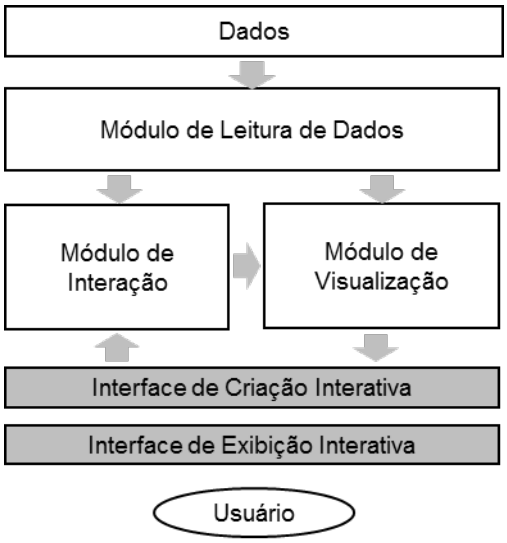
Fonte: Heer, Card e Landay (2005)

A implementação desta solução mantém o mesmo modelo e estende o funcionamento de cada um dos módulos ao nível gráfico, ou seja, cria interfaces que dão suporte à geração de visualizações neste modelo, proporcionando na perspectiva de Criação Interativa desde a seleção do banco de dados a ser utilizado até a escolha dos componentes da forma visual da Visualização e das interações. A perspectiva de Visualização Interativa proporciona o resultado do processo de criação, ou seja, a exibição da Visualização de Informação e os itens auxiliares da interação, conforme implementado interativamente na perspectiva de Criação Interativa.

Na solução, uma vez que o usuário tenha acesso aos dados em uma conexão com o banco de dados, é possível escolher entre a Exibição ou a Criação em uma tela principal. Dessa forma o módulo de leitura de dados não possui parte gráfica, uma vez que não há contato do usuário diretamente com os dados. Já os módulos de interação e visualização possuem interface gráfica, porém, enquanto que a interface de Criação Interativa proporciona acesso às estruturas de dados abstratas, assim como as estruturas de dados visuais e tipos de interação, a perspectiva de Exibição disponibiliza somente a visualização final. O módulo de visualização tem o objetivo de receber os dados e transformá-los em estruturas visuais que são exibidas na interface de acordo com as interações do usuário, ou seja, é composto de uma herança de classes que compõem a visualização, juntamente com seus recursos principais como layout e dados abstraidos. Já o módulo de interação aplica mudanças sobre os dados, lidos por meio do módulo de leitura de dados, ou seja, este módulo realiza o tratamento necessário com os dados

abstraídos e os disponibiliza para que o módulo de visualização gere novamente a visualização a cada interação (Figura 18).

Figura 18 - Módulos funcionais do sistema

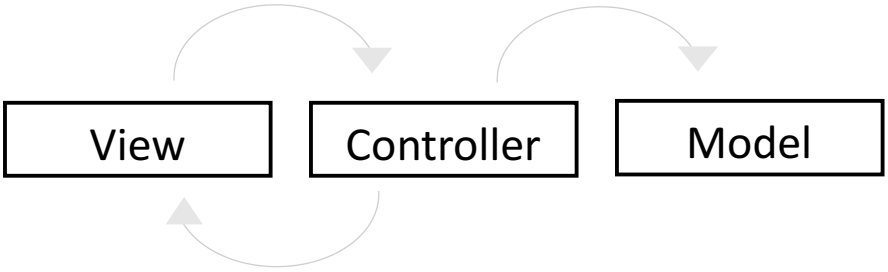


Fonte: Jaqueline Zaia, 2015

4.3 Arquitetura

A arquitetura utilizada no desenvolvimento deste sistema, conforme já discutido no capítulo de Tecnologia, é predominantemente Monolítica, justificado pelas recomendações propostas por Bederson, Grosjean e Meyer (2004), em que arquiteturas monolíticas, apesar de serem mais complexas de serem compreendidas pelo usuário, são mais favoráveis quando utilizadas no desenvolvimento de interfaces bidimensionais e proporcionam mais possibilidades de modificação e personalização dos componentes. Além disso, a proposta deste trabalho é a implementação de uma GUI que permita a geração de visualizações sem necessidade de codificação pelo usuário.

Figura 19 - Representação do Design Pattern MVC



Fonte: Jaqueline Zaia, 2015

Para organização do código, será utilizado o Design Pattern *Model View Controller* (MVC), que se trata de uma organização comumente utilizada no desenvolvimento de Interfaces Gráficas com o Usuário, uma vez que separa as porções da aplicação em *View*, que é toda a interface com o usuário, *Model,* que são os dados a serem representados, e *Controller*, que se trata da forma de interação que será disponibilizada (KRASNER; POPE, 1988), que interagem conforme a figura 19. Uma das principais vantagens da sua utilização é que possibilita o desenvolvimento e manutenção de forma fácil, já que isola a aparência da aplicação do seu funcionamento e dados e suporta facilmente diversas interfaces com linguagens e permissões diversas (LEFF; RAYFIELD, 2001).

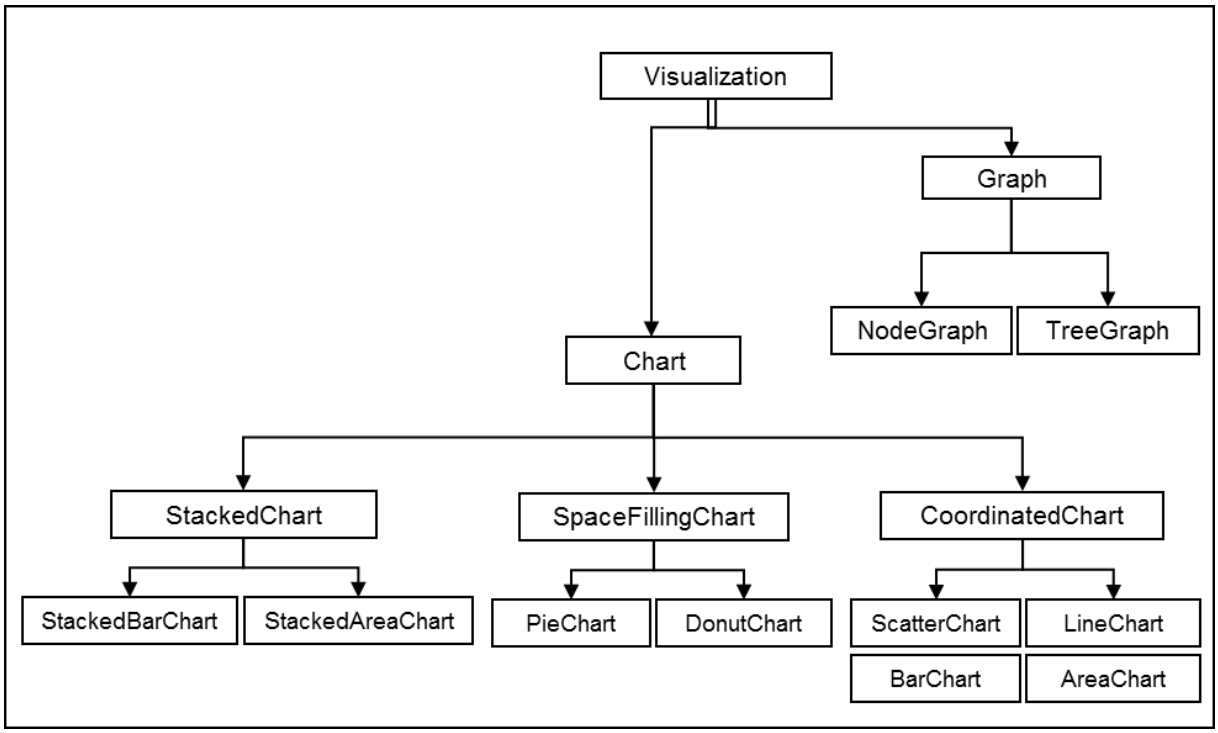
4.4 Técnicas de Visualização

As técnicas de visualização a serem utilizadas nesse protótipo visam a atender às necessidades básicas do usuário de acordo com os tipos de dados já descritos (SHNEIDERMAN, 1996). Dessa forma, são escolhidas, baseado nas limitações do *kit* de ferramentas utilizado e nas técnicas existentes estudadas, as técnicas de visualização e interação implementadas com foco em dados tabulares, podendo ser expandidas em trabalhos futuros. De acordo com a arquitetura desenhada para o sistema, a "Visualização" é uma classe pai que deve possuir as funcionalidades básicas de qualquer visualização. Logo abaixo na hierarquia, as visualizações são divididas entre visualizações de dados n-dimensionais e visualizações de relacionamentos, sendo estas o Gráfico e o Grafo. Enquanto o primeiro difere os tipos de gráfico por quantidade de dimensões dos dados, os tipos de grafo diferem entre grafos de relacionamento e árvores, que são um tipo específico de grafo que representa relacionamentos hierárquicos.

Os tipos de gráficos são divididos em três subclasses, de acordo com o layout que devem apresentar: gráfico de coordenadas, gráfico de preenchimento de espaço e gráfico empilhado. Já no último nível da hierarquia de classes é onde se encontram as configurações voltadas para a aparência e estilo do gráfico, como cor, tamanho e detalhes como ícones, coordenadas e legendas (Figura 20). Além do modelo de dados da visualização, há também classes gráficas pares que na perspectiva de edição são exibidas em abas no menu inferior para possibilitar a edição, que são caracterizadas por seções com componentes de seleção. Para a escolha do tipo de gráfico, é implementada uma seção com um menu gráfico com as opções disponíveis. O mesmo ocorre com relação às configurações, cada configuração possui sua própria seção

gráfica com os recursos disponíveis para aquela característica, todos em forma de botões, caixas de seleção e outros componentes gráficos. Essas seções são dispostas na aba inferior de configurações denominada “Chart”.

Figura 20 – Modelo de classes do módulo de visualizações deste sistema



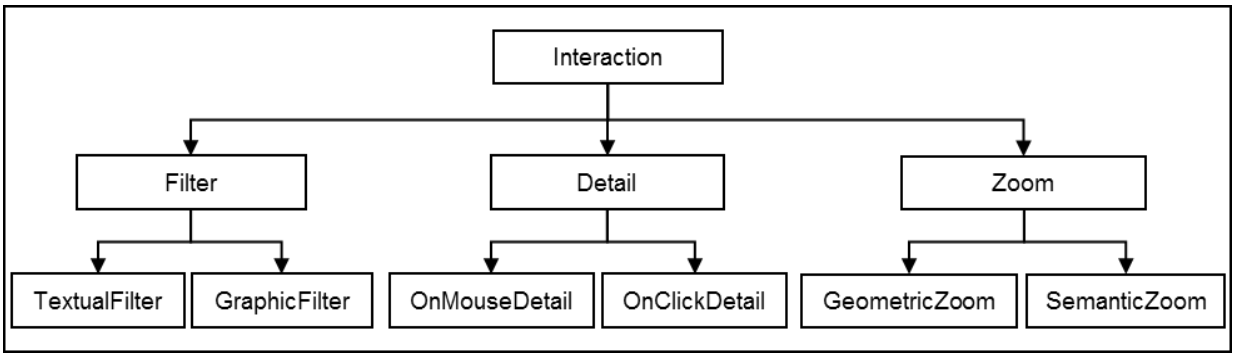
Fonte: Jaqueline Zaia, 2015

4.5 Técnicas de Interação

As técnicas de interação a serem utilizadas nesta solução visam a atender às tarefas básicas propostas por Shneiderman (1996). As duas primeiras tarefas devem ser proporcionadas por meio do Zoom, que permite a navegação nos dados, seja geometricamente, aumentando o tamanho dos itens da visualização, ou semanticamente, aumentando o grau de detalhe da informação. A tarefa de filtragem é proporcionada de duas formas: Textualmente, por meio de uma caixa de busca que filtra os dados, ou graficamente, por meio de interação direta com os itens da visualização. Da mesma forma, os detalhes por demanda que, por meio da navegação ou cliques do mouse, exibem detalhes na própria visualização ou próximo a ela, podendo ser composta de texto, ícones ou outras visualizações. O relacionamento entre dados pode ser feito utilizando a combinação de visualizações, que possibilita ao usuário exibir e relacionar duas

visualizações em uma única página, que proporciona ao usuário duas visões que poder ser dependentes e interagirem entre si. Essas possíveis formas implementadas de interagir diretamente com a visualização são ilustradas no módulo de interação do sistema, conforme a figura 21. Na perspectiva de edição, as configurações de interação podem ser editadas de forma similar às características de visualização, em seções gráficas localizadas na aba inferior denominada “Interaction”.

Figura 21 - Modelo de classes do módulo de interação do sistema



Fonte: Jaqueline Zaia, 2015

Já o histórico de interações na visualização pode ser feito ao utilizar o módulo histórico, que funciona como um cache dos últimos estados da visualização durante a interação, de forma que não seja necessário replicar as mesmas interações para voltar a um estado já antes visualizado dos dados na exibição. E, por fim, a extração de dados, que, apesar de os dados brutos não serem visíveis, é possível por meio do salvamento da visualização, seja utilizando os dados estáticos, ou seja, exatamente os mesmos dados de quando a visualização é criada; ou dinâmicos, como no caso de os dados serem oriundos de um banco de dados relacional, neste caso a visualização deve apresentar os dados atualizados de acordo com uma conexão com o banco de dados.

5 Testes e Resultados

Durante o processo de desenvolvimento, o protótipo do sistema foi submetido a testes funcionais, meramente para avaliar se implementava corretamente os requisitos e funcionalidades especificados. Porém, é necessário avaliar o sistema de maneira mais aprofundada, considerando a sua usabilidade com foco na criação de visualizações. Para isso, foram conduzidos experimentos com voluntários, cujo objetivo foi validar, do ponto de vista da usabilidade, o processo de criação de visualizações de informação de forma interativa e também a exploração da informação gerada como produto deste processo. Cada experimento foi conduzido individualmente com um voluntário em sala privada, e durou em média 50 minutos.

5.1 Metodologia Empregada

A usabilidade é um fator importante na qualidade do software. Por esse motivo, diferentes padrões, guias e critérios têm sido criados nos últimos anos, dados pela necessidade de maior usabilidade nas interfaces (MARIAGE; VANDERDONCKT; PRIBEANU, 2005). Contudo, ferramentas de visualização de informação encaram alguns problemas ao validar suas interfaces. Uma das razões pelas quais isso ocorre é que a visualização depende diretamente do contexto em que é utilizada, dos dados a serem tratados, de quais tarefas são executadas com eles e de quem as está executando (GRANLUND; LAFRENIÈRE; CARR, 2001).

Por isso, as técnicas normalmente usadas para avaliar interfaces com o usuário podem não ser plenamente adequadas ao serem consideradas para a avaliação de interfaces com o usuário voltadas para visualização de informação. Além disso, é comum que durante a exploração da ferramenta, novas perguntas sejam formuladas e respondidas dinamicamente de forma colaborativa, por mais de um usuário. Esse processo de descoberta torna mais difícil a observação e medição do uso da ferramenta (SHNEIDERMAN; PLAISANT, 2006).

Por conta disso, para testar o protótipo de solução desenvolvido conforme discutido no capítulo anterior, foi utilizado um teste em que o processo de criação da visualização é analisado como uma tarefa juntamente com a sua exploração; e não somente a exploração da visualização e seus componentes individualmente. Para isso, o princípio é validar não só a visualização gerada, mas também avaliar juntamente as ferramentas do sistema para criar a visualização do zero, inclusive com foco no aspecto de criação, ainda que não excluindo a avaliação da

exploração de seu resultado. Dessa forma, o usuário executa uma tarefa composta, em que é validada a usabilidade tanto da interface com o usuário durante o processo de criação quanto a visualização resultante dele.

Para auxiliar na validação da usabilidade do sistema, foi utilizado o questionário USE – Usefulness, Satisfaction and Ease of Use (LUND, 2001). Comentários dos voluntários também foram coletados durante e no fim do teste, antes de o questionário ser aplicado, a fim de mapear os detalhes de cada ponto de vista sobre a interação durante as tarefas. Os resultados do experimento são analisados com base no questionário e nos comentários dos voluntários a fim de avaliar a experiência do usuário criando e interpretando visualizações.

5.2 População

A população de voluntários que realizou o teste é composta de cinco profissionais da área de tecnologia e análise de sistemas, sem conhecimento técnico em programação e visualização de informação; e que tiveram pouco um nenhum contato significativo com ferramentas de geração de visualizações. Apesar disso, todos estão familiarizados em suas profissões com o conceito de grupos de dados textuais, seja dados estatísticos, bancos de dados relacionais ou planilhas Excel.

A escolha de uma amostra dessa população se dá para refletir um possível perfil do usuário do sistema, que são em geral analistas de dados, analistas de sistemas, gestores de negócio e outros papéis similares. Isso porque dentre as possíveis tarefas envolvidas no dia a dia destes profissionais está a necessidade de processar, analisar e tirar conclusões baseadas em dados e informações automatizadas. Além disso, é considerado também o conhecimento, área de atuação e tipos de informações que cada um deles lida habitualmente. Isso se justifica pelo fato de que não deve haver viés para diferentes tipos de dados e assuntos. Ou seja, espera-se que idealmente qualquer pessoa com as características descritas deve ter um nível de desempenho e conclusões similares ao utilizar o sistema, independentemente dos dados que são apresentados.

No entanto, é importante mencionar que para esse teste é escolhido um público específico, considerando que esse perfil deve ser similar à maioria dos usuários desse sistema. Usuários sem experiência nenhuma nos pontos mencionados anteriormente ou que possuem domínio no desenvolvimento de visualizações de informação não são considerados para esse teste. Apesar

disso, espera-se que, no caso de usuários sem experiência, haja maior dificuldade na criação dos gráficos e interpretação dos dados, dada uma menor experiência com tarefas de análise, que se justifica em grande parte pela compreensão dos dados e não só pelas funcionalidades do sistema. Neste caso é possível que uma breve explicação seja o suficiente para que o usuário possa prosseguir sem impeditivos. No caso da escolha por usuários experientes, espera-se que não haja prejuízos na execução dos testes, pelo contrário, é possível que esses usuários tenham mais agilidade na utilização do InterVis.

5.3 Procedimento

O procedimento de aplicação dos testes se inicia com uma apresentação breve do projeto a cada voluntário, seus objetivos e como deve acontecer o procedimento de teste. A interface do sistema é exibida com um relatório já criado e mencionado como um exemplo de como devem ocorrer as tarefas.

5.3.1 Termo de Consentimento Livre Esclarecido

Em seguida, é entregue o Termo de Consentimento Livre Esclarecido (TCLE) – onde se encontram as mesmas informações descritas verbalmente, além dos riscos, custos e benefícios envolvidos no procedimento de teste, conforme consta no apêndice A. O voluntário é informado textualmente e verbalmente nessa etapa que pode desistir a qualquer momento, sem necessidade de justificativa e sem prejuízo algum para si.

5.3.2 Pré-teste

Após assinatura do TCLE, dá-se início ao pré-teste, que consiste em dez minutos em que o voluntário fica livre para explorar o sistema. A instrução dada é que o voluntário pode, a seu critério, explorar os dados disponíveis e criar novos documentos. O voluntário também é encorajado a fazer perguntas em qualquer momento que julgar necessário e nesse momento não há supervisão do que está sendo feito, exceto quando solicitada. O objetivo do pré-teste é deixar o voluntário confortável e minimamente familiarizado com a interface do sistema, permitir que tire possíveis dúvidas do funcionamento e esclarecer curiosidades e incertezas sobre o próprio procedimento.

5.3.3 Execução do Teste

Em seguida, o teste é iniciado. São descritas verbalmente em detalhes como devem ser conduzidas as tarefas, que são também entregues textualmente, uma de cada vez. Não é informado previamente ao voluntário quanto tempo deve durar cada tarefa.

5.3.3.1 Os Dados

A descrição das tarefas se inicia com a apresentação da base de dados, que trata de informações de estatísticas populacionais brasileiras de domínio público disponibilizadas pelo Instituto Brasileiro de Geografia e Estatística – IBGE (IBGE, 2016). A escolha dos dados é motivada pela sua natureza pública e de conhecimento geral, em que as informações contidas, como nomes de cidades e estados, são familiares, assim como os conceitos existentes neles como gênero, idade e quantidade. Isso visa a reproduzir uma situação real em que o usuário geralmente conhece o conceito dos dados que deseja analisar; e dessa forma evitar que haja impeditivos na avaliação do uso do sistema pelos voluntários ou grande vantagem de um participante sobre os outros por maior conhecimento do domínio.

É importante notar que poderia ser escolhida qualquer base de dados para a realização dos testes, assim como para a própria utilização do sistema. A premissa em qualquer caso é que o usuário tenha domínio na área de conhecimento a que os dados se destinam, a fim de garantir o melhor aproveitamento das funcionalidades do sistema ao explorar os dados.

Além disso, é apresentado verbalmente o conjunto de dados e quais informações estão contidas ali, como qual é o período de retenção dos dados, onde e como são coletadas e tratadas as informações. Este passo é conduzido de forma detalhada, justificado pela necessidade do usuário em conhecer os dados com os quais deve trabalhar. A duração média dessa etapa foi de 6 minutos e contou com a participação ativa dos voluntários, que fizeram perguntas e comentários.

5.3.3.2 As Tarefas

Foram duas tarefas similares que consistem em criar um novo relatório com uma única visualização e atribuir a ela dados e funcionalidades que atendam a certos requisitos, conforme o apêndice B. Cada tarefa é descrita em função do resultado que se deseja alcançar no ponto de

vista da área de aplicação e é composta de dois a quatro requisitos. Os requisitos sugerem quais perguntas a visualização deve responder, sem entrar em detalhes de quais funcionalidades podem ser utilizadas, a fim de deixar que os voluntários procurem e testem os recursos que julgarem ideais para mostrar determinada informação. Pôde ser observada a predileção de cada voluntário por determinadas formas de abstração, que se repetiam nas tarefas seguintes, mas diferiam entre eles.

Na primeira tarefa, o objetivo é criar uma visualização em que seja possível identificar a) o crescimento ou diminuição da população no decorrer dos anos e b) a variação das faixas de idade no decorrer dos anos. O primeiro requisito infere que é necessário visualizar o número populacional em relação ao tempo em anos, enquanto o segundo requisito acrescenta uma classificação ao número populacional, que é a faixa etária. Tanto a quantidade de pessoas, quanto o ano de referência e a faixa etária são informações encontradas em colunas de uma das tabelas da base de dados. Essa tarefa é a primeira, visto que tende a ser a tarefa mais simples com apenas dois requisitos, ambos relacionados à quantidade de pessoas. Espera-se que a quantidade de pessoas deve ser a primeira característica a ser identificada juntamente com o ano; e que a visualização resultante disso seja uma linha do tempo, cuja quantidade é representada pelo eixo Y e os anos são representados pelo eixo X. Até então o tipo de gráfico a ser selecionado seria indiferente. Já com o segundo requisito, espera-se que os voluntários percebam a necessidade de um gráfico com séries de dados empilhadas, sendo uma para cada faixa de idade, o que torna a escolha pelo layout determinante. Nesta tarefa, não é solicitado que haja ainda nenhum tipo de interação ou componentes adicionais ao gráfico, a fim de introduzir ao voluntário os recursos básicos e proporcionar confiança na execução das tarefas.

Na segunda tarefa, são inseridos dois requisitos que implicam na necessidade de incluir um recurso de interação, juntamente com mais dois requisitos semelhantes à tarefa anterior. Nesta, o voluntário é convidado a criar uma visualização da população por regiões do Brasil diferenciadas por gênero. Na visualização, deve ser possível: a) identificar a quantidade de pessoas em cada região do Brasil, b) identificar a quantidade de homens e mulheres em cada região do Brasil, c) identificar a quantidade de pessoas em cada estado do Brasil e d) identificar a quantidade de homens e mulheres em cada estado do Brasil.

É possível observar nos requisitos que os dois primeiros e os dois últimos são similares, porém os dois primeiros identificam a quantidade de pessoas (homens e mulheres) por região e os dois seguintes fazem o mesmo por estado. Novamente, a primeira característica a ser

identificada é a quantidade de pessoas, seguida da classificação por região ou estado. Nessa tarefa, espera-se que um gráfico de pontos ou barras seja criado, em que o eixo X mostra as regiões ou estados e o eixo Y mostra a quantidade de pessoas; a decisão de gerar o gráfico por região ou estado fica a critério do voluntário, já que é possível em ambos os casos identificar as características pedidas nos requisitos ao inserir recursos de interação, que devem ser o próximo passo.

Novamente é necessária a utilização de duas séries de dados para diferenciar os gêneros de forma similar ao que ocorreu na primeira tarefa. O próximo passo é identificar a necessidade de criar recursos interativos para responder às perguntas dos dois próximos requisitos. As opções possíveis são a criação de um filtro e um componente de detalhe ativado pelo movimento do mouse, que possam responder às perguntas dos dois últimos requisitos.

5.3.3.3 Avaliação de Usabilidade Pelos Participantes

A avaliação da usabilidade do protótipo pelos participantes do experimento ocorre de duas formas: uma avaliação verbal seguida do preenchimento do questionário USE. A avaliação verbal durou em média dez a quinze minutos e tratou-se de uma conversa informal sobre o sistema. Não houve perguntas diretas aos voluntários sobre o sistema, mas sim sobre as tarefas. Os voluntários foram estimulados e comentar passo a passo as dificuldades e facilidades que no julgamento dele foram determinantes na execução das tarefas.

Ao fim da avaliação verbal, é apresentado o questionário USE, conforme mostrado na tabela 1 do apêndice C. O questionário é discutido no artigo de Lund (2001) e está disponível online[1]. Este é composto com 30 perguntas avaliando 4 áreas (utilidade, facilidade de uso, facilidade de aprender e satisfação) por meio de uma escala Likert de 1 a 7, em que 1 representa a pior avaliação e 7 a melhor. Como todos os participantes eram fluentes em inglês, o questionário foi apresentado em sua forma original em inglês para evitar imprecisões na tradução. Inicialmente são descritos verbalmente aos participantes os grupos de perguntas do questionário. Então, cada voluntário é deixado sozinho para concluir o preenchimento do questionário, que é feito no computador, o que garante anonimato de cada grupo de respostas.

[1] http://garyperlman.com/quest/quest.cgi?form=USE

No fim do teste, que é composto de 30 questões, o usuário pode preencher três aspectos negativos e três positivos da experiência, também anônimos.

Durante a avaliação é ressaltado o fato de que o desempenho do voluntário não é avaliado em nenhum momento do teste, mas que todas as perguntas são voltadas para o sistema. A justificativa é que o sistema deve atender às necessidades do usuário e não o contrário, que, além de reduzir o risco de constrangimento ou pressão sobre o participante do teste, reafirma o princípio da transparência de Rutkowski (1982), em que a ferramenta não deve ser percebida pelo usuário, que deve estar concentrado somente na tarefa.

5.4 Resultados

Os testes funcionais mostraram que o protótipo foi desenvolvido como especificado.

No experimento com os usuários, no pré-teste observou-se que no geral eles tendem a criar um novo documento, abrem nos modos de exibição e edição, exploram superficialmente o menu e se mostram dispostos a iniciar as tarefas antes que acabe o tempo estabelecido para o pré-teste. Também ocorreram diversas perguntas e comentários, como quantas bases de dados podem ser combinadas, como navegar entre as perspectivas ou para que servem determinadas funcionalidades.

Para o primeiro requisito da primeira tarefa, o tipo de gráfico era indiferente, porém foi possível notar a preocupação dos voluntários em selecionar todos os layouts disponíveis antes de passar para o próximo requisito. Já no segundo requisito, em que a seleção do layout afetava a tarefa, ficou claro para os voluntários quais dentre as opções já exploradas serviam e quais não para o propósito, assim como a utilidade em inserir legenda e escolher as cores. Todos os voluntários concluíram com sucesso a tarefa dentro do tempo previsto.

Na realização da segunda tarefa, a identificação da principal característica a ser visualizada (a quantidade de pessoas) ocorreu com facilidade para todos os participantes. Um dos voluntários leu todos os requisitos antes de iniciar a criação e teve dúvidas em iniciar pelo estado ou pela região, identificando a similaridade dos requisitos. Os demais seguiram os requisitos sequencialmente, da mesma forma que na primeira tarefa, criando o gráfico por regiões. Os voluntários não tiveram dificuldades em identificar a necessidade de criar duas séries (homens e mulheres) para realizar essa tarefa e criaram o gráfico com mais rapidez que na primeira, sugerindo que estavam aprendendo sobre o uso do sistema rapidamente. Na criação

de recursos interativos necessária na segunda tarefa, para dois voluntários foi a primeira vez que a aba de interação foi explorada em detalhes e surgiram dúvidas sobre se seria correto utilizar um recurso interativo para isso.

Todos os voluntários concluíram todas as tarefas atendendo a todos requisitos, porém na segunda tarefa surgiram alguns questionamentos relacionados com as interações, como a possibilidade não existente de mover as caixas de busca e detalhes dentro da visualização; a dificuldade em perceber o funcionamento da caixa de detalhes, que depende da interação direta do usuário sobre a visualização para que alguma informação seja exibida; e a forma de utilizar o zoom na visualização uma vez que a funcionalidade de zoom é acionada. Ao final dos testes observou-se que as tarefas tiveram um tempo de execução de somente cinco a dez minutos cada (inicialmente o tempo previsto era entre 10 e 15 minutos).

Houve diversos comentários dos voluntários durante o teste que são considerados como importantes na avaliação do InterVis, que tratam sugestões, críticas e perguntas relevantes no ponto de vista da usabilidade e expansão de funcionalidades e que, em alguns casos, foram recorrentes. No início do teste, foi possível notar com frequência perguntas sobre por onde iniciar, como inserir novas bases de dados, quais os eixos X e Y como adicionar um dado no gráfico, para onde arrastar um dado. Essas perguntas foram rapidamente respondidas, às vezes pelo próprio voluntário, porém mostram a falta de instruções claras no protótipo, ou de alguma outra forma rápida de transmitir instruções iniciais ao usuário, como por exemplo um breve tutorial interativo. Por outro lado, em grande parte das dúvidas, o próprio voluntário respondia sua pergunta rapidamente explorando o sistema e descobrindo intuitivamente qual era a resposta.

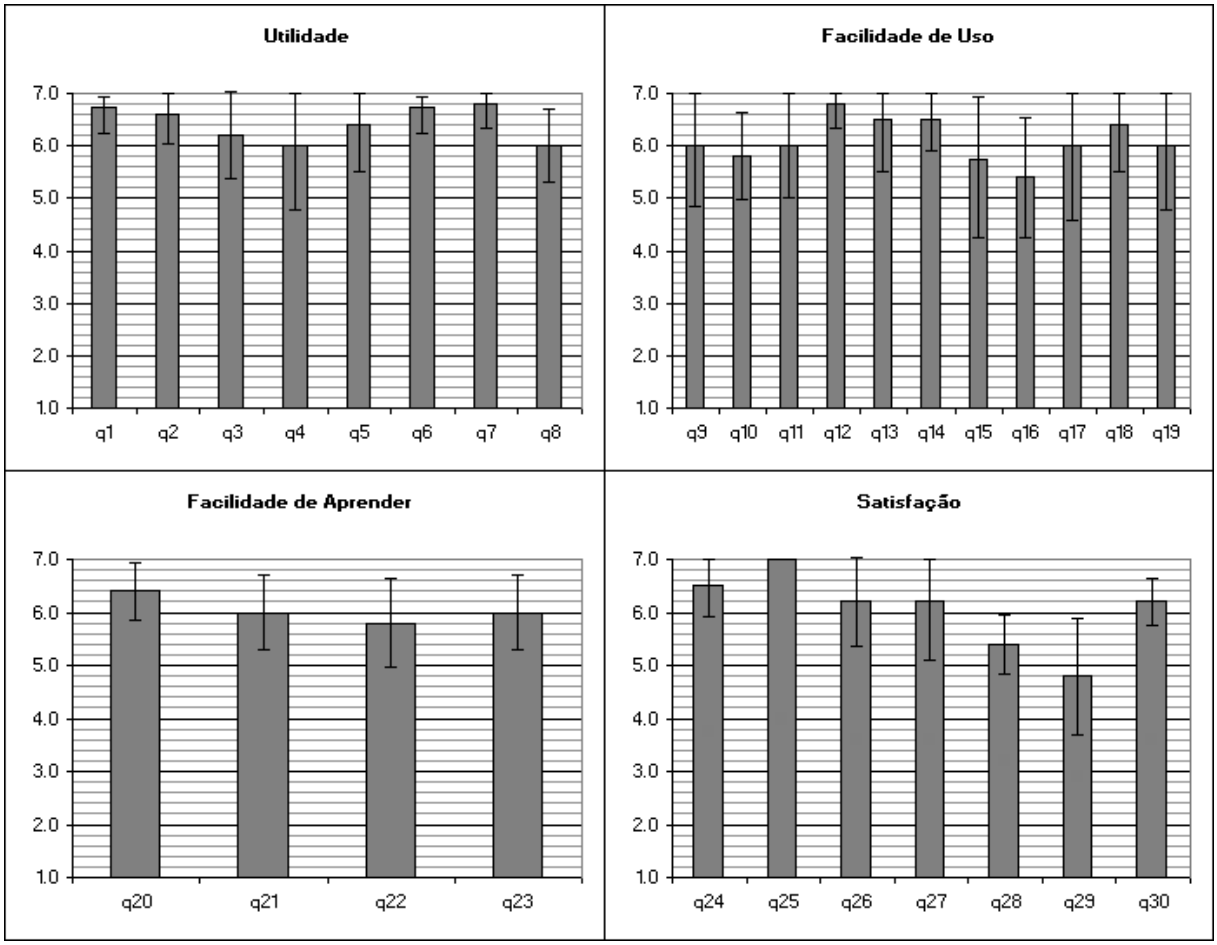
Algumas observações foram feitas de comportamentos recorrentes entre os voluntários, e mostram que melhorias ainda podem ser feitas, principalmente com relação à disposição dos recursos na tela. A primeira observação é quanto à escolha dos dados e das bases de dados. Os voluntários, a princípio, demoraram para explorar a estrutura dos dados, em forma de árvore, disponibilizado em um painel horizontal estreito à esquerda da tela. Essa dificuldade, em conjunto com as perguntas sobre onde e como relacionar os dados com a visualização, mostra que os dados não estão dispostos de forma clara para que o usuário as explore intuitivamente. Frequentemente os voluntários, quando queriam sair da perspectiva de edição voltando para a perspectiva inicial ou para a visualização, levavam o mouse até o canto superior esquerdo da tela, quando os botões para essa ação se encontram no canto superior direito. O mesmo acontece

na perspectiva de exibição, em que, quando os voluntários querem editar uma visualização sendo exibida, buscam o botão para fazê-lo no canto superior esquerdo da tela, quando este encontra-se centralizado na barra inferior de recursos. Esses erros recorrentes mostram que existe um hábito, possivelmente baseado no uso de outras ferramentas, em encontrar um menu no canto superior esquerdo.

Houve também duas sugestões relevantes. A primeira sugestão é relacionada com uma expansão da funcionalidade de filtro e diz respeito ao tipo de dado que pode ser filtrado. O InterVis possibilita que determinado dado do gráfico seja filtrado por um valor textual ou numeral. A sugestão é que o filtro possa também abranger intervalos, de forma que dados numéricos e datas possam ser filtradas baseado em determinado intervalo de dados. A segunda está relacionada com a facilitação do aprendizado de uso do InterVis e se manifestou de duas maneiras. A primeira foi a sugestão de caixas de texto explicativas em cada uma das funcionalidades quando o mouse é posicionado sobre elas ou botões explicativos junto das funcionalidades; a fim de ajudar o usuário a identificar qual o comportamento das funcionalidades antes que haja perda de tempo com o seu uso. A segunda foi a proposta de uma solução menos específica que não só explicasse cada funcionalidade, mas que orientasse passo a passo a criação de uma visualização, com recomendações sobre quais os melhores layouts e recursos em função do que o usuário deseja fazer, para que houvesse instrução e conduzisse o usuário a tomar as melhores decisões ao criar suas visualizações.

Durante a avaliação verbal ao fim das tarefas, os comentários foram sempre pontuais sobre alguma funcionalidade e como ela ajudou ou atrapalhou. Houve também algumas sugestões baseadas na possibilidade de atender a outros requisitos que não constavam na tarefa, que mostra que os voluntários absorveram o processo de gerar visualizações e tiveram visão crítica de propor funcionalidades em função de novos possíveis problemas. Esse comportamento remete ao que Andrienko et. Al. (2007) apresentam ao afirmar que o sistema deve ser capaz de responder perguntas que o usuário faz; e que esse processo é dinâmico e cíclico na medida em que o usuário explora mais os dados e surgem novas perguntas.

Figura 22 - Resultado do questionário USE separado pelos quesitos avaliados.



Fonte: Jaqueline Zaia, 2016

A avaliação do protótipo pelos participantes através do questionário USE mostrou resultados bastante positivos, conforme pode ser observado na tabela 2 do apêndice C. Dentre os quatro quesitos avaliados no questionário, o que obteve a pior avaliação foi Satisfação, com média 6 em uma escala de 1 a 7. Facilidade de uso e facilidade de aprendizado foram avaliadas com média 6.1 e utilidade teve a avaliação mais alta, com média 6.4. A figura 22 mostra as médias das respostas dos cinco participantes (no eixo das ordenadas) para cada questão do USE (de q1 a q30, as abscissas) como barras, e o desvio padrão como barras de erro. As questões estão agrupadas nos quatro quesitos.

Dentre as questões sobre utilidade, as que tiveram respostas mais baixas foram as questões 4 e 8, referentes respectivamente ao controle que o usuário tem e se o sistema apresenta todas as funcionalidades esperadas. Como nem todos os tipos de layout, gráficos, interação e funcionalidades foram implementados no protótipo, justamente por ser somente um protótipo, a avaliação mais baixa nesses aspectos já era esperada (e na verdade foi até mais positiva que

o esperado). No entanto durante os testes surgiram diversas sugestões de mudanças ou novas funcionalidades que não tinham sido especificadas anteriormente, de modo que mesmo se o experimento tivesse sido feito com o sistema em um estágio mais avançado de desenvolvimento, ainda não teriam sido implementadas. As questões sobre utilidade com avaliação mais alta referem-se à eficiência do uso do sistema, à economia de tempo que ele proporciona e às necessidades que satisfaz, mostrando que os participantes de fato valorizaram a utilidade do InterVis.

Sobre facilidade de uso, a questão com avaliação mais alta refere-se ao sistema permitir a realização de tarefas com o menor número de ações possível, novamente reforçando a eficiência percebida pelos participantes. A questão com menor avaliação refere-se a inconsistências na interface, refletindo mais uma vez os comentários dos usuários e observações de uso indicando que certos elementos deveriam estar mais claros e posicionados em regiões diferentes no layout da aplicação. Já em relação à facilidade de aprendizado, os participantes consideraram como item com maior avaliação o pouco tempo necessário para aprender a usar o sistema, e como pior a facilidade de aprender (possivelmente devido às inconsistências já mencionadas). Por fim, em relação à satisfação, todos os usuários avaliaram com nota máxima que recomendariam o sistema e a pior avaliação refere-se a sentir que precisam do sistema. Como nenhum dos participantes trabalhava diretamente com Visualização de Informação, ainda que estivessem bem familiarizados com a manipulação de grandes quantidades de dados, essa resposta também não foi uma surpresa.

5.5 Riscos

Quando se trata de pesquisas envolvendo pessoas, sempre há riscos éticos que devem ser levados em consideração. Além disso, existem também os riscos à validade dos testes, que podem pôr em dúvida os resultados de trabalhos embasados nestes e que precisam ser validados.

5.5.1 Riscos Éticos

Riscos éticos estão presentes em qualquer tipo de pesquisa que envolva pessoas em sua participação e podem variar desde a veiculação de materiais colhidos sem a autorização dos participantes até danos graves à saúde. Apesar de os testes na área de engenharia de software

apresentarem riscos reduzidos em relação a testes em outras áreas, como as áreas da saúde, esses também têm potencial de causar danos aos envolvidos (FILGUEIRAS; SILVA, 2008).

Durante a preparação do teste executado neste projeto, alguns riscos foram identificados e minimizados. Um dos riscos identificados é a possível exposição da identidade dos voluntários na veiculação da pesquisa, que pode ser dar pela identificação direta ou indireta pelo comportamento e linguagem empregada. A fim de minimizar esse risco, todo o processo de teste é realizado sem a identificação do voluntário e todo o material gerado é tratado antes de iniciada a sua avaliação, de forma a preservar cada forma de expressão e evitar o reconhecimento das pessoas.

Outra situação passível de ocorrer durante o teste do InterVis é o caso de levar o voluntário à frustração caso não consiga realizar alguma das tarefas, independentemente do motivo. Para minimizar esse risco, é informado ao voluntário durante todo o procedimento que a avaliação não é referente ao seu desempenho pessoal e que não é necessário concluir o teste, caso se sinta desconfortável. Além disso, as tarefas são pensadas para serem breves e são testadas funcionalmente antes do teste a fim de estabelecer uma média de tempo que não cause muito estresse para o voluntário e garanta que não haverá problemas técnicos que o impeçam de realizar as tarefas.

5.5.2 Riscos à Validade do Teste

Além dos riscos para os envolvidos, existem também os riscos e vieses que interferem na validade do teste, os quais devem ser considerados na preparação e avaliação do teste, e minimizados sempre que possível. Dentre os possíveis riscos à validade do teste neste projeto, é identificada a possibilidade de haver opiniões e comportamentos enviesados justificados pela seleção de uma população composta de especialistas da área de tecnologia. Apesar de o sistema ser voltado para essa população, o fato de o teste executado se restringir a ela abre a possibilidade de que funcionalidades, sejam as mais básicas ou as mais complexas, não sejam avaliadas de forma apropriada. Isso pode ocorrer, pois o seu uso pelos voluntários pode ser restrito pelo hábito de utilizar outras ferramentas ou ter algum conhecimento específico não considerado para a avaliação, por exemplo.

Além disso, o teste pode conter variações relacionadas com o julgamento dos voluntários, que é subjetivo e pode ser influenciado de acordo com os mais diversos fatores,

como a forma como o teste é conduzido com cada um dos voluntários, se são utilizadas afirmações e frases encorajadoras ou se as descrições verbais do procedimento são mais detalhadas, por exemplo. As experiências mais recentes do voluntário e a sensação de novidade causada pela apresentação de uma nova solução ou mesmo sua intenção de "ser simpático" ou agradar o avaliador também podem influenciar os resultados de forma positiva ou negativa. A fim de minimizar o risco de julgamento, o procedimento é conduzido de forma similar com todos os voluntários, sendo utilizados os mesmos discursos, descrições e materiais no mesmo período de tempo. A população relativamente pequena de participantes também agravaria essas variações provenientes do julgamento subjetivo e, em trabalhos futuros, planeja-se realizar experimentos com populações maiores. Apesar disso, os desvios relativamente baixos nas respostas do questionário USE e nos tempos de execução da tarefa indicam que, mesmo para essa população pequena, essa variação foi relativamente baixa.

6 Conclusão

Com a existência de bases de dados textuais extensas com informações e padrões importantes contidos nelas, a área de visualização de informação tem se tornado cada vez mais útil e necessária. Diversas novas soluções comerciais e pesquisa acadêmica têm voltado sua atenção para a busca de técnicas, bibliotecas e ferramentas para melhor visualização de informação a partir de dados brutos. Porém, uma das grandes dificuldades na área de visualização de informação está relacionada com a busca da melhor forma de visualizar cada tipo de dado em função da tarefa que o usuário deseja desempenhar. Nesse contexto, pode-se observar duas naturezas de conhecimento envolvidas na criação de visualização de informação. A primeira é o conhecimento na área da aplicação dos dados em que o usuário deseja analisar, a fim de garantir que as informações necessárias estejam sendo exibidas da melhor forma possível para análise naquele domínio. A segunda é o conhecimento técnico nas áreas de programação e de visualização de informação, que garantem um melhor aproveitamento no uso dos recursos disponíveis para exibir tais informações. Portanto, para a criação de uma visualização de informação ideal, torna-se necessário aproximar essas duas naturezas de conhecimento.

Um dos caminhos para fazer isso é proporcionando o suporte técnico necessário ao usuário – que já tem o conhecimento na sua área de aplicação. Dessa forma, ele tem autonomia para gerar suas próprias visualizações da maneira como achar conveniente, sem que tenha que se preocupar com detalhes técnicos da visualização, como a forma como ela deve funcionar. Existem diversos trabalhos relacionados que propõem formas de facilitar a criação de visualizações de informação pelo usuário. Porém, enquanto alguns dão foco em criar uma interface de visualização de informação com o usuário e acabam limitando sua utilização a um domínio de aplicação; outros buscam flexibilizar a quantidade de técnicas e acabam criando soluções que exigem certo preparo e conhecimento específico para suas utilizações. Neste contexto, é possível encontrar soluções muito similares, mas com focos diferentes, que vão desde a performance para grandes volumes de dados e utilização de dados dinâmicos até exibição e interação de estruturas de dados específicas.

Seguindo o caminho de fornecer ao usuário final recursos técnicos simples para que ele seja capaz de gerar visualizações, o objetivo desse trabalho foi desenvolver e testar um sistema para criação e exploração interativa de visualizações de informação a partir de dados dinâmicos,

que visa a permitir que o usuário final possa gerar e editar visualizações de acordo com a sua necessidade, independentemente da natureza da informação que deve ser analisada.

Pode-se citar como principal contribuição deste trabalho para a área de visualização a análise da combinação dos recursos essenciais para a geração de uma visualização de informação; assim como a avaliação dos pontos altos e baixos de concentrar tais recursos em uma única ferramenta de forma simplificada e próxima da realidade de usuários de qualquer área de aplicação.

Para atingir tal objetivo, foi conduzido um estudo sobre as técnicas de visualização e interação disponíveis, a fim de selecionar aquelas mais frequentes e eficazes. Técnicas de avaliação de GUI's no geral, e técnicas para testar visualizações em específico também foram estudadas e discutidas no trabalho para permitir uma melhor avaliação do protótipo desenvolvido do sistema. Além disso, foi feita uma comparação entre os *kits* de ferramentas de suporte a visualizações de informação; e outra comparação entre os sistemas similares para geração interativa de visualizações, com o intuito de mapear características em comum, sua recorrência e suas vantagens e desvantagens. Esse mapeamento pode ser considerado como outra contribuição deste trabalho.

Uma vez mapeadas as características consideradas necessárias para atingir o objetivo, o trabalho foi conduzido em três etapas: a codificação do protótipo do InterVis, a realização de testes com os usuários e a análise dos resultados.

Para o desenvolvimento do InterVis, foram reaproveitados recursos visuais, modelos de arquitetura e algoritmos de outras bibliotecas. Alguns pontos relevantes foram destacados e implementados, como a importância em omitir os dados brutos do usuário final para evitar a exposição de dados sensíveis; e a necessidade da inclusão de funcionalidades interativas na visualização de forma também interativa, por meio da GUI. O objetivo dessa etapa foi construir um protótipo para ser utilizado nos testes reproduzindo uma amostra significativa das principais características e técnicas mapeadas e classificadas anteriormente.

Na etapa de testes com os usuários, um experimento foi conduzido com voluntários individualmente, em que cada usuário pôde realizar um pré-teste para se familiarizar com o InterVis e, em seguida, realizou tarefas compostas de geração e análise da visualização gerada. O objetivo dos testes foi avaliar a usabilidade da ferramenta InterVis tanto na criação de visualizações como em sua exploração, com tarefas que exigiam a combinação desses dois aspectos para que fossem completadas. Durante os testes, os voluntários foram incentivados a

comentar a experiência; e no final dos testes, foi respondido por eles o questionário USE (LUND, 2001) seguido de uma avaliação verbal do sistema e da experiência em usá-lo.

Na terceira etapa foi feita a consolidação do resultado dos testes com o usuário, em que o resultado do questionário, juntamente com os comentários anônimos foram analisados e as características recorrentes e relevantes dos comentários foram tabuladas. Nessa etapa foi possível identificar os pontos fortes e fracos do InterVis, assim como sugestões de melhorias e trabalhos futuros. O resultado final indica que InterVis é uma ferramenta que, apesar de ainda permitir diversos aprimoramentos, atende a todos os objetivos e requisitos propostos no início deste trabalho, bem como os princípios de usabilidade estabelecidos no objetivo, sendo uma ferramenta de fácil utilização e aprendizado, e que provê uma experiência satisfatória para o usuário, como indicam os resultados dos testes e do USE.

Apesar das avaliações de usabilidade positivas, alguns pontos levantados podem ser considerados como melhorias e trabalhos futuros, de forma a agregar e refinar funcionalidades e recursos, assim como avaliar e analisar outras características da ferramenta. Os pontos mais relevantes são descritos a seguir.

Uma ferramenta instrutiva: O principal ponto observado diz respeito ao aprendizado do usuário. Isso por que não houve instrução direta aos voluntários nos testes sobre como utilizar a ferramenta propositalmente, a fim identificar com maior clareza situações em que o uso da ferramenta é pouco intuitivo. Apesar de o resultado dos testes em relação ao aprendizado do usuário ter sido satisfatório, pôde ser observada a dificuldade do usuário em explorar a diversidade de recursos que a ferramenta oferece, mesmo nesse protótipo com funcionalidades mais limitadas. Entende-se que isso ocorre não só pelos problemas já apontados de usabilidade da GUI, mas pela falta de experiência dos voluntários em utilizar visualização de informação. Isso sugere uma possível forma de melhoria do software, em que seja possível não só prover os recursos que o usuário necessita, mas também dar instrução para que ele possa tomar melhores decisões e ter senso crítico ao utilizá-los da forma que melhor lhe atenderem durante a criação da visualização, independentemente de ser um usuário experiente ou não na utilização do InterVis.

Uma ferramenta mais rápida: Embora o resultado dos testes do ponto de vista da usabilidade tenha sido satisfatório, seria interessante a aplicação de outros testes para avaliar também características que não eram o foco dessa solução. É o caso de testes de desempenho na criação de visualizações com grandes volumes de dados, na integração de diferentes fontes

de dados e testes técnicos e na eficiência da aplicação para objetivos mais específicos, como áreas de aplicação que exigem a combinação de diversas estruturas de dados e visualizações diferentes. Além disso, são necessários testes com uma população maior.

Uma ferramenta mais completa: Outro ponto de melhoria está no refinamento dos recursos disponíveis no InterVis, que representam pouco se considerarmos todos os recursos visuais e algoritmos existentes de visualização de informação. Dado que o sistema segue um modelo cujas funcionalidades podem ser estendidas, agregar esses novos recursos e funcionalidades deve ser tarefa mais simples que seu desenvolvimento inicial, principalmente se for liberado como sistema de código aberto. Além disso, outra grande lacuna no sistema atualmente é que ele não transmite aos usuários sem conhecimentos técnicos específicos de visualização de informação. Diversas recomendações úteis sobre a criação de visualizações que já são bem conhecidas na área (como que tipos de visualização ou interação se adequam melhor a determinados tipos de dado ou tarefa, por exemplo), não estão explicitamente presentes no protótipo. Um sistema de recomendação adaptativo capaz de transmitir esse conhecimento para o usuário da melhor forma possível seria não só uma ferramenta útil como também instrutiva, além de um desafio e contribuição interessantes do ponto de vista de pesquisa.

Pode-se concluir que a solução proposta e prototipada neste trabalho, o InterVis, é uma ferramenta bastante satisfatória do ponto de vista da usabilidade que, apesar ainda possuir detalhes a serem aprimorados, atende aos objetivos e requisitos definidos inicialmente para dar suporte à criação e exploração de visualizações de informação interativas de forma eficiente, simples e satisfatória, sem exigir do usuário conhecimentos de programação ou visualização de informação. Além disso, tem a possibilidade de ser enriquecida e integrada a outras ferramentas de forma descomplicada e transparente. O sistema une as características essenciais de uma visualização de informação de forma interativa sem excluir a possibilidade da sua expansão, dessa forma assumindo um papel de peça chave para a visualização de informações. Contribui ainda para a área de visualização como um exemplo de quais características, recursos, funcionalidades e elementos um sistema desse tipo deve ter, com a análise comparativa de ferramentas semelhantes e demonstrando uma forma de avaliação que combina nas mesmas tarefas a necessidade de criação e exploração de visualizações, algo incomum na literatura. Um artigo completo sobre este trabalho foi publicado e apresentado no HCI International 2016 (ZAIA; BERNARDES, 2016).

Referências

ABELLO, James; VAN HAM, Frank; KRISHNAN, Neeraj. Ask-graphview: A large scale graph visualization system. Visualization and Computer Graphics, IEEE Transactions on, v. 12, n. 5, p. 669-676, 2006.

ADOBE, 2015. Disponível em: <http://www.adobe.com/>. Acesso em: 15 Julho 2015.

AHLBERG, Christopher; SHNEIDERMAN, Ben. Visual information seeking: Tight coupling of dynamic query filters with starfield displays. In: Proceedings of the SIGCHI conference on Human factors in computing systems. p. 313-317. ACM, 1994.

ALENCAR, Aretha Barbosa. Mineraçao e visualizaçao de coleçoes de séries temporais. Tese de Doutorado. Instituto de Ciências Matemáticas e de Computação. 2007.

ANDREWS, Keith. Evaluating information visualisations. In: Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization. p. 1-5. ACM, 2006.

ATW Java API. 2015. Disponível em: <https://docs.oracle.com/javase/8/docs/technotes/guides /awt/index.html>. Acesso em: 15 Julho 2015.

BEDERSON, Benjamin B.; GROSJEAN, Jesse; MEYER, Jon. Toolkit design for interactive structured graphics. Software Engineering, IEEE Transactions on, v. 30, n. 8, p. 535-546, 2004.

BEDERSON, Benjamin B.; HOLLAN, James D. Pad++: a zooming graphical interface for exploring alternate interface physics. In: Proceedings of the 7th annual ACM symposium on User interface software and technology. p. 17-26. ACM, 1994.

BEDERSON, Benjamin B.; MEYER, Jon; GOOD, Lance. Jazz: an extensible zoomable user interface graphics toolkit in Java. In: Proceedings of the 13th annual ACM symposium on User interface software and technology. p. 171-180. ACM, 2000.

BERTIN, Jacques. Semiology of graphics: diagrams, networks, maps. 1983.

CARD, Stuart. Information Visualization, in JACKO, Julie A. (Ed.). Human Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications. CRC press, 2012.

CARD, Stuart K.; MACKINLAY, Jock D.; SHNEIDERMAN, Ben. Readings in information visualization: using vision to think. Morgan Kaufmann, 1999.

CARR, David A. Guidelines for designing information visualization applications. Proceedings of ECUE, v. 99, p. 1-3, 1999.

CATARCI, Tiziana et al. Visual query systems for databases: A survey. Journal of Visual Languages & Computing, v. 8, n. 2, p. 215-260, 1997.

CHANG, Remco et al. Legible cities: Focus-dependent multi-resolution visualization of urban relationships. Visualization and Computer Graphics, IEEE Transactions on, v. 13, n. 6, p. 1169-1175, 2007.

CHEN, Chaomei. Searching for intellectual turning points: Progressive knowledge domain visualization. Proceedings of the National Academy of Sciences, v. 101, n. suppl 1, p. 5303-5310, 2004.

CHEN, Chaomei. Top 10 unsolved information visualization problems. Computer Graphics and Applications, IEEE, v. 25, n. 4, p. 12-16, 2005.

COCKBURN, Andy; KARLSON, Amy; BEDERSON, Benjamin B. A review of overview+ detail, zooming, and focus+ context interfaces. ACM Computing Surveys (CSUR), v. 41, n. 1, p. 2, 2008.

COLLINS, Christopher M.; CARPENDALE, Sheelagh. VisLink: Revealing relationships amongst visualizations. Visualization and Computer Graphics, IEEE Transactions on, v. 13, n. 6, p. 1192-1199, 2007.

COLLINS, Christopher; CARPENDALE, Sheelagh; PENN, Gerald. Docuburst: Visualizing document content using language structure. In: Computer graphics forum. Blackwell Publishing Ltd, p. 1039-1046. 2009.

CRAIG, Paul; KENNEDY, Jessie. Coordinated graph and scatter-plot views for the visual exploration of microarray time-series data. In: Information Visualization, 2003. INFOVIS 2003. IEEE Symposium on. IEEE, p. 173-180. 2003.

DASGUPTA, Aritra; KOSARA, Robert. Adaptive privacy-preserving visualization using parallel coordinates. Visualization and Computer Graphics, IEEE Transactions on, v. 17, n. 12, p. 2241-2248, 2011.

DE ALMEIDA BIOLCHINI, Jorge Calmon et al. Scientific research ontology to support systematic review in software engineering. Advanced Engineering Informatics, v. 21, n. 2, p. 133-151, 2007.

DE BRITTO ALMEIDA, Caio Sacramento; APOLINÁRIO JR, Antônio Lopes. OpenedEyes: A web standards-based generic framework for multidimensional information visualization. SBC Journal on Interactive Systems, v. 4, n. 1, p. 25-36, 2013.

DE OLIVEIRA, Maria Cristina Ferreira; LEVKOWITZ, Haim. From visual data exploration to visual data mining: a survey. Visualization and Computer Graphics, IEEE Transactions on, v. 9, n. 3, p. 378-394, 2003.

EXCEL. 2016. Disponível em: < https://products.office.com/pt-br/excel>. Acesso em: 20 julho 2016.

FEKETE, Jean-Daniel. The infovis toolkit. In: Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on. IEEE, 2004.

FEW, Stephen. Information dashboard design. O'Reilly, 2006.

FILGUEIRAS, Lucia Vilela Leite; SILVA, Bárbara Dariano. Ética no envolvimento de seres humanos na Engenharia de Software. Scientia 19.2, 2008.

FLUIT, Christiaan. Autofocus: Semantic search for the desktop. In: Information Visualisation, 2005. Proceedings. Ninth International Conference on. IEEE, p. 480-487. 2005.

FOX, David Strickman. Tabula rasa: a multiscale user interface system. New York University, 1998. FREITAS et al., 2002.

FREITAS, Carla Maria Dal Sasso et al. Introdução à visualização de informações. Revista de informática teórica e aplicada. Porto Alegre. Vol. 8, n. 2 (out. 2001), p. 143-158, 2001.

FURNAS, George W. Generalized fisheye views. ACM, 1986.

FURNAS, George W.; BEDERSON, Benjamin B. Space-scale diagrams: Understanding multiscale interfaces. In: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM Press/Addison-Wesley Publishing Co., p. 234-241. 1995.

GANSNER, Emden R.; HU, Yifan; KOBOUROV, Stephen. GMap: Visualizing graphs and clusters as maps. In: Pacific Visualization Symposium (PacificVis), 2010 IEEE. IEEE, p. 201-208. 2010.

GRANGER, Clive William John; NEWBOLD, Paul. Forecasting economic time series. Academic Press, 2014. HABER, Robert B. Visualization techniques for engineering mechanics. Computing Systems in Engineering, v. 1, n. 1, p. 37-50, 1990.

GRANLUND, A.; LAFRENIÈRE, D.; CARR, D. A pattern-supported approach to the user interface design process. Proceedings of HCI International, Vol. 1. 2001.

HARGER, John R.; CROSSNO, Patricia J. Comparison of open-source visual analytics toolkits. In: IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, 2012.

HAVRE, Susan et al. Themeriver: Visualizing thematic changes in large document collections. Visualization and Computer Graphics, IEEE Transactions on, v. 8, n. 1, p. 9-20, 2002.

HEER, Jeffrey; CARD, Stuart K.; LANDAY, James A. Prefuse: a toolkit for interactive information visualization. In: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, p. 421-430. 2005.

HERMAN, Ivan; MELANÇON, Guy; MARSHALL, M. Scott. Graph visualization and navigation in information visualization: A survey. Visualization and Computer Graphics, IEEE Transactions on, v. 6, n. 1, p. 24-43, 2000.

HOCHHEISER, Harry; SHNEIDERMAN, Ben. Dynamic query tools for time series data sets: timebox widgets for interactive exploration. Information Visualization, v. 3, n. 1, p. 1-18, 2004.

HORNBÆK, Kasper; BEDERSON, Benjamin B.; PLAISANT, Catherine. Navigation patterns and usability of zoomable user interfaces with and without an overview. ACM Transactions on Computer-Human Interaction (TOCHI), v. 9, n. 4, p. 362-389, 2002.

IBGE. Instituto Brasileiro de Geografia e Estatística. 2016. Disponível em: <http://www.ibge.gov.br/home/ >. Acesso em: 15 janeiro 2016.

ISO/IEC. 9241-11 Ergonomic Requirements for OfTice Work with Visual Display Terminals (VDTs) - Part 11: Guidance on usability. ISO/IEC 9241-11, 1998.

JAVA3D, Java 3D API, Java SE Desktop Technologies. 2015. Disponível em: <http://www.oracle.com/technetwork/articles/javase/index-jsp-138252.html>. Acesso em: 14 Julho 2015.

JAVED, Waqas; ELMQVIST, Niklas. Stack zooming for multi-focus interaction in time-series data visualization. In: Pacific Visualization Symposium (PacificVis), 2010 IEEE. IEEE, p. 33-40. 2010.

JOHNSON, Brian; SHNEIDERMAN, Ben. Tree-maps: A space-filling approach to the visualization of hierarchical information structures. In: Visualization, 1991. Visualization'91, Proceedings., IEEE Conference on. IEEE, p. 284-291. 1991.

KEIM, Daniel et al. Challenges in visual data analysis. In: Information Visualization, 2006. IV 2006. Tenth International Conference on. IEEE, p. 9-16. 2006.

KEIM, Daniel et al. Information visualization and visual data mining. Visualization and Computer Graphics, IEEE Transactions on, v. 8, n. 1, p. 1-8, 2002.

KEIM, Daniel A. Visual exploration of large data sets. Communications of the ACM, v. 44, n. 8, p. 38-44, 2001.

KENNEDY, Jessie B.; MITCHELL, Kenneth J.; BARCLAY, Peter J. A framework for information visualisation. ACM SIGMOD Record, v. 25, n. 4, p. 30-34, 1996.

KITCHENHAM, Barbara. Procedures for performing systematic reviews. Keele, UK, Keele University, v. 33, n. 2004, p. 1-26, 2004.

KNAUTZ, Kathrin; SOUBUSTA, Simone; STOCK, Wolfgang G. Tag clusters as information retrieval interfaces. In: System Sciences (HICSS), 2010 43rd Hawaii International Conference on. IEEE, p. 1-10. 2010.

KRASNER, Glenn E. et al. A description of the model-view-controller user interface paradigm in the smalltalk-80 system. Journal of object oriented programming, v. 1, n. 3, p. 26-49, 1988.

LANDAUER, Thomas K. Research methods in human-computer interaction. Handbook of human-computer interaction, p. 905-28, 1988.

LEE, Bongshin et al. How users interact with biodiversity information using TaxonTree. In: Proceedings of the working conference on Advanced visual interfaces. ACM, p. 320-327. 2004.

LEFF, Avraham; RAYFIELD, James T. Web-application development using the model/view/controller design pattern. In: Enterprise Distributed Object Computing Conference, 2001. EDOC'01. Proceedings. Fifth IEEE International. IEEE, p. 118-127. 2001.

LUND, A. M.: Measuring usability with the USE questionnaire. Usability interface 8.2. 2001.

MACEACHREN, Alan M. et al. Geographic visualization: Designing manipulable maps for exploring temporally varying georeferenced statistics. In: Information Visualization, 1998. Proceedings. IEEE Symposium on. IEEE, p. 87-94, 156. 1998.

MARIAGE, C.; VANDERDONCKT, J.; PRIBEANU, C. State of the art of web usability guidelines. The handbook of human factors in web design, pp.688-700. 2005.

MCCORMICK, Bruce Howard; DEFANTI, Thomas A.; BROWN, Maxine D. Visualization in scientific computing. IEEE Computer Graphics and Applications, v. 7, n. 10, p. 69-69, 1987.

MICHOTTE, Albert. The perception of causality. 1963.

MOERE, Andrew Vande. Time-varying data visualization using information flocking boids. In: Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on. IEEE, p. 97-104. 2004.

O'MADADHAIN, Joshua et al. Analysis and visualization of network data using JUNG. Journal of Statistical Software, v. 10, n. 2, p. 1-35, 2005.

OPEN INVENTOR, Silicon Graphics International Corp. 2015. Disponível em: <http://www.sgi.com/software/inventor/>. Acesso em: 24 Julho 2015.

ORACLE. Oracle Business Analytics. 2015. Disponível em: <http://www.oracle.com/br/solutions/business-analytics/overview/index.html>. Acesso em: 20 julho 2015.

PARASURAMAN, A.; ZEITHAML, V.; BERRY, L. SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality. Retailing: critical concepts, v. 64, n. 1, p. 140, 2002.

PERLIN, Ken; FOX, David. Pad: an alternative approach to the computer interface. In: Proceedings of the 20th annual conference on Computer graphics and interactive techniques. ACM, p. 57-64. 1993.

PIETRIGA, Emmanuel; APPERT, Caroline. Sigma lenses: focus-context transitions combining space, time and translucence. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, p. 1343-1352. 2008.

PLAISANT, Catherine. The challenge of information visualization evaluation. In: Proceedings of the working conference on Advanced visual interfaces. ACM, 2004. p. 109-116.

PLAISANT, Catherine et al. LifeLines: visualizing personal histories. In: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 1996. p. 221-227.

PLAISANT, Catherine; CARR, David; SHNEIDERMAN, Ben. Image-browser taxonomy and guidelines for designers. Software, IEEE, v. 12, n. 2, p. 21-32, 1995.

POWERPOINT. Microsoft PowerPoint. 2016. Disponível em: <https://products.office.com/en/powerpoint >. Acesso em: 22 junho 2016.

RHYNE, Theresa-Marie et al. Information and scientific visualization: Separate but equal or happy together at last. 2003.

ROBERTSON, George G.; MACKINLAY, Jock D.; CARD, Stuart K. Cone trees: animated 3D visualizations of hierarchical information. In: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 1991. p. 189-194.

RUTKOWSKI, C. An introduction to the human applications standard computer interface. Byte, 7(10), p. 291-314. 1982.

SAS Visual Analytics. 2015. Disponível em: <http://www.sas.com/pt_br/software/business-intelligence/visual-analytics.html>. Acesso em 20 julho 2015.

SCHROEDER, Will J.; LORENSEN, Bill; MARTIN, Ken. The visualization toolkit. Kitware, 2004.

SEAH, Boon-Siew et al. FUSE: a system for data-driven multi-level functional summarization of protein interaction networks. In: Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium. ACM, 2012. p. 847-850.

SHNEIDERMAN, Ben. The eyes have it: A task by data type taxonomy for information visualizations. In: Visual Languages, 1996. Proceedings., IEEE Symposium on. IEEE, 1996. p. 336-343.

SHNEIDERMAN, B.; PLAISANT, C. Strategies for evaluating information visualization tools: multi-dimensional in-depth long-term case studies. Proceedings of the 2006 AVI workshop on Beyond time and errors: novel evaluation methods for information visualization, pp. 1-7. 2006.

SPRENGER, Thomas Carl; BRUNELLA, R.; GROSS, Markus H. H-BLOB: a hierarchical visual clustering method using implicit surfaces. In: Proceedings of the conference on Visualization'00. IEEE Computer Society Press, 2000. p. 61-68.

STASKO, John; ZHANG, Eugene. Focus+ context display and navigation techniques for enhancing radial, space-filling hierarchy visualizations. In: Information Visualization, 2000. InfoVis 2000. IEEE Symposium on. IEEE, 2000. p. 57-65.

STOLTE, Chris; TANG, Diane; HANRAHAN, Pat. Polaris: A system for query, analysis, and visualization of multidimensional relational databases. Visualization and Computer Graphics, IEEE Transactions on, v. 8, n. 1, p. 52-65, 2002.

SWING Java API. 2015. Disponível em: <http://docs.oracle.com/javase/7/docs/api/javax/swing/package-summary.html>. Acesso em: 15 Julho 2015.

TABLEAU. 2016. Disponível em: <http://www.tableau.com/>. Acesso em: 20 julho 2016.

TANAKA, Yoichi; OKADA, Yoshihiro; NIIJIMA, Koichi. Treecube: Visualization tool for browsing 3d multimedia data. In: Information Visualization, 2003. IV 2003. Proceedings. Seventh International Conference on. IEEE, 2003. p. 427-432.

THAKUR, Sidharth; RHYNE, Theresa-Marie. Data vases: 2d and 3d plots for visualizing multiple time series. In: Advances in Visual Computing. Springer Berlin Heidelberg, 2009. p. 929-938.

TOMINSKI, Christian; SCHULZE-WOLLGAST, Petra; SCHUMANN, Heidrun. 3d information visualization for time dependent data on maps. In: Information Visualisation, 2005. Proceedings. Ninth International Conference on. IEEE, 2005. p. 175-181.

TVERSKY, Barbara; MORRISON, Julie Bauer; BETRANCOURT, Mireille. Animation: can it facilitate?. International journal of human-computer studies, v. 57, n. 4, p. 247-262, 2002.

VAN WIJK, Jarke J.; VAN SELOW, Edward R. Cluster and calendar based visualization of time series data. In: Information Visualization, 1999.(Info Vis' 99) Proceedings. 1999 IEEE Symposium on. IEEE, 1999. p. 4-9, 140.

VISUMAP. 2016. Disponível em: <http://www.visumap.net/>. Acesso em: 20 julho 2016.

WATTENBERG, Martin. Baby names, visualization, and social data analysis. In: Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on. IEEE, 2005. p. 1-7.

YAFFEE, Robert A.; MCGEE, Monnie. An introduction to time series analysis and forecasting: with applications of SAS® and SPSS®. Academic Press, 2000.

ZAIA, Jaqueline; BERNARDES JR., João Luiz. A Graphical System for Interactive Creation and Exploration of Dynamic Information Visualization. In International Conference on Human Interface and the Management of Information, pp. 214-225. Springer International Publishing, 2016.

Apêndice A – Termo de Consentimento Livre Esclarecido

Você está sendo convidado (a) como voluntário (a) a participar da pesquisa: InterVis - Um Sistema Gráfico para Criação e Exploração Interativa de Visualizações de Informação Dinâmicas.

1 Justificativa, Objetivos e Procedimentos

O motivo que nos leva a estudar o problema de geração de visualizações de informação é que muitas vezes é necessário conhecimento técnico em programação e desenvolvimento de componentes visuais tornando esse processo custoso e difícil de ser personalizado. A pesquisa se justifica pela necessidade de facilitar a criação de visualizações de informação dinâmicas por pessoas sem conhecimento técnico. O objetivo desse projeto é desenvolver uma ferramenta de criação de visualizações de informação dinâmicas que não demande ao usuário conhecimento técnico em programação. O procedimento de coleta de dados será da seguinte forma: Será proposta uma lista de tarefas exploratórias em um contexto de dados de conhecimento público e será solicitado que o voluntário as execute utilizando a ferramenta proposta. Durante a execução sugere-se que o voluntário faça comentários da forma como achar conveniente para reportar a experiência. No fim da execução um questionário é aplicado. A duração de todo o experimento não deve passar de uma hora.

2 Desconfortos, Riscos e Benefícios

Não existem riscos, desconfortos ou benefícios.

3 Garantia de Esclarecimento, Liberdade de Recusa e Garantia de Sigilo

Você será esclarecido (a) sobre a pesquisa em qualquer aspecto que desejar. Você é livre para recusar-se a participar, retirar seu consentimento ou interromper a participação a qualquer momento. A sua participação é voluntária e a recusa em participar não irá acarretar qualquer penalidade ou perda de benefícios.

A pesquisadora irá tratar a sua identidade com padrões profissionais de sigilo. Os resultados das conversas e questionário serão enviados para você e permanecerão confidenciais.

Seu nome ou o material que indique a sua participação não será liberado sem a sua permissão. Você não será identificado (a) em nenhuma publicação que possa resultar deste estudo. Uma cópia deste consentimento informado será arquivada e outra será fornecida a você.

4 Custos da Participação, Ressarcimento e Indenização por Eventuais Danos

A participação no estudo não acarretará custos para você e não será disponível nenhuma compensação financeira adicional. No caso você sofrer algum dano decorrente dessa pesquisa, não existe nenhum tipo de compensação financeira.

5 Declaração do (a) Participante

Eu, _____ fui informada (o) dos objetivos da pesquisa acima de maneira clara e detalhada e esclareci minhas dúvidas. Sei que em qualquer momento poderei solicitar novas informações e motivar minha decisão se assim o desejar. A estudante Jaqueline Zaia de Sousa e o seu professor orientador João Luiz Bernardes Jr. certificaram-me de que todos os dados desta pesquisa serão confidenciais.

Também sei que caso existam gastos adicionais, estes serão absorvidos pelo orçamento da pesquisa. Em caso de dúvidas poderei chamar a estudante Jaqueline Zaia de Sousa ou o professor orientador João Luiz Bernardes Jr; ou o Comitê de Ética em Pesquisa em Seres Humanos da Escola de Artes, Ciências e Humanidades (EACH) da Universidade de São Paulo (USP), sito à Av. Arlindo Béttio, 1000, Sala T14 - I1, Ermelino Matarazzo – São Paulo, SP.

Declaro que concordo em participar desse estudo. Recebi uma cópia deste termo de consentimento livre e esclarecido e me foi dada a oportunidade de ler e esclarecer as minhas dúvidas.

Nome	Assinatura do Participante	Data
Nome	Assinatura do Pesquisador	Data
Nome	Assinatura da Testemunha	Data

Apêndice B – Instruções para Tarefas

Este projeto tem como objetivo desenvolver um software para geração interativa de visualizações de informação dinâmicas, ou seja, em que é possível se interagir. O resultado que se espera alcançar é que um software para esse fim seja eficiente na criação de visualizações. Para isso, são feitos testes qualitativos de usabilidade da ferramenta, a fim de levantar os pontos fortes e fracos dela. Os testes são baseados na execução de tarefas pelo voluntário e visam avaliar exclusivamente a ferramenta e não o desempenho do mesmo na tarefa, conforme o seguinte roteiro:

- Apresentação breve sobre a pesquisa;
- Assinatura do Termo de Consentimento Livre Esclarecido;
- Apresentação da ferramenta e período para adaptação e dúvidas;

1 Tarefa #1 – Idade da População

Dados: Population's Age – Dados estatísticos da quantidade de pessoas no Brasil e as respectivas faixas de idade.

Requisitos: Gerar uma visualização em que:

- É possível identificar o crescimento da população no decorrer dos anos;
- É possível identificar as faixas de idade.

2 Tarefa #2 – População das Regiões por Gênero

Dados: Brazil's States – Dados dos estados como população, região e gênero.

Requisitos: Gerar uma visualização em que:

- É possível identificar a população total por região;
- É possível visualizar a quantidade de homens e mulheres em cada região;
- É possível identificar a população total por estado;
- É possível visualizar a quantidade de homens e mulheres em cada estado.

Apêndice C – Perguntas e Resultados do Questionário

Tabela 1 – Perguntas que compõem o questionário USE

Número da Pergunta	Pergunta
	USEFULNESS
1	It helps me be more effective
2	It helps me be more productive
3	It is useful
4	It gives me more control over the activities in my life
5	It makes the things I want to accomplish easier to get done
6	It saves me time when I use it
7	It meets my needs
8	It does everything I would expect it to do
	EASE OF USE
9	It is easy to use
10	It is simple to use
11	It is user friendly
12	It requires the fewest steps possible to accomplish what I want to do with it
13	It is flexible
14	Using it is effortless
15	I can use it without written instructions
16	I don't notice any inconsistencies as I use it
17	Both occasional and regular users would like it
18	I can recover from mistakes quickly and easily
19	I can use it successfully every time
	EASE OF LEARNING
20	I learned to use it quickly
21	I easily remember how to use it
22	It is easy to learn to use it
23	I quickly became skillful with it
	SATISFACTION
24	I am satisfied with it
25	I would recommend it to a friend
26	It is fun to use
27	It works the way I want it to work
28	It is wonderful
29	I feel I need to have it
30	It is pleasant to use

Fonte: Lund (2001)

Tabela 2 – Resultados da aplicação do questionário USE no teste com os voluntários

Questão	Nota do Voluntário 1	Nota do Voluntário 2	Nota do Voluntário 3	Nota do Voluntário 4	Nota do Voluntário 5
1	7	7	6	7	
2	7	6	7	7	6
3	7	6	5	7	6
4	4	7	6	7	6
5	7	6	7	7	5
6	7	6	7	7	
7	7	7	7	6	7
8	5	7	6	6	6
9	7	5	5	7	
10	7	5	5	6	6
11	7	5	5	6	7
12	7	6	7	7	7
13	7	7	5	7	
14	7	7	6	6	
15	5	7	4	7	
16	5	7	4	5	6
17	7	7	5	4	7
18	7	7	5	7	6
19	7	6	4	7	6
20	7	6	6	7	6
21	7	5	6	6	6
22	7	5	6	6	5
23	7	6	6	6	5
24	7	6	6	7	
25	7	7	7	7	
26	7	5	6	6	7
27	7	5	5	7	7
28	5	6	5	6	5
29	3	5	5	6	5
30	7	6	6	6	6

Fonte: Jaqueline Zaia, 2016