



**CENTRO UNIVERSITÁRIO LUTERANO DE PALMAS**

*Recredenciado pela Portaria Ministerial nº 3.607, de 17/10/05, D.O.U. nº 202, de 20/10/2005*  
ASSOCIAÇÃO EDUCACIONAL LUTERANA DO BRASIL

Lucas Henrique Roese

SENTIMENTALL: MÓDULO DE VISUALIZAÇÃO DE INFORMAÇÃO

Palmas – TO  
2016

Lucas Henrique Roese  
SENTIMENTALL: MÓDULO DE VISUALIZAÇÃO DE INFORMAÇÃO

Monografia elaborada e apresentada como requisito parcial para aprovação na disciplina de Trabalho de Conclusão de Curso II (TCC II) do curso de bacharel em Ciência da Computação pelo Centro Universitário Luterano de Palmas (CEULP/ULBRA).

Orientador: Prof.<sup>a</sup> M.Sc. Parcilene Fernandes de Brito

Palmas – TO

2016

Lucas Henrique Roese  
SENTIMENTALL: MÓDULO DE VISUALIZAÇÃO DE INFORMAÇÃO

Monografia elaborada e apresentada como requisito parcial para aprovação na disciplina de Trabalho de Conclusão de Curso II (TCC II) do curso de bacharel em Ciência da Computação pelo Centro Universitário Luterano de Palmas (CEULP/ULBRA).

Orientador: Prof.<sup>a</sup> M.Sc. Parcilene Fernandes de Brito.

Aprovado em: \_\_\_\_/\_\_\_\_/\_\_\_\_

BANCA EXAMINADORA

---

Prof. M.Sc. Parcilene Fernandes de Brito  
Centro Universitário Luterano de Palmas – CEULP

---

Prof. M.Sc. Jackson Gomes de Souza  
Centro Universitário Luterano de Palmas – CEULP

---

Prof. M.Sc. Fabiano Fagundes  
Centro Universitário Luterano de Palmas – CEULP

Palmas – TO  
2016

## RESUMO

Técnicas de Visualização de Informação (VI) foram desenvolvidas para aprimorar e facilitar o acesso, entendimento e visualização de dados. Considerando as necessidades apresentadas pelo grupo de pesquisa em Engenharia Inteligente de Dados do CEULP/ULBRA, esse trabalho teve por objetivo a implementação de um módulo de visualização de informações para a ferramenta SentimentALL. Para tanto, foi utilizada uma base de dados de informações coletadas no site *TripAdvisor*, pré-processada e tratada a partir da utilização da técnica de Análise de Sentimentos implementada em Oliveira (2015). Os estudos das técnicas de VI foram divididos em duas partes, visualização de dados e visualização de informação, de forma a apresentar quais técnicas são voltadas para cada área. Um ponto importante em VI é a interatividade, que foi abordada tendo como base o mecanismo elaborado por Card et al. (1999). O desenvolvimento deste trabalho deu-se em seis etapas, desde reuniões com o especialista de domínio até a construção e entrega do módulo de VI para a SentimentALL. Como resultado deste trabalho, foram desenvolvidas ferramentas para criar dicionários de termos e taxonomias, além de apresentações gráficas que representam a aplicação do dicionário de termos e das técnicas de VI sobre os dados gerados pelo módulo de análise de sentimentos.

**Palavras-chave:** Informação; análise de dados, visualização de informação, taxonomias, dicionário de termos;

## LISTA DE ILUSTRAÇÕES

Figura 1: Tabela simples .....	13
Figura 2: Gráfico de torta .....	14
Figura 3: Gráficos de barra.....	15
Figura 4: Formas de gráficos de linha .....	16
Figura 5: Gráfico de linha .....	17
Figura 6: Gráfico de dispersão .....	18
Figura 7: Coordenadas paralelas .....	19
Figura 8: Mapa de árvore .....	20
Figura 9: Diagrama entidade relacionamento.....	21
Figura 10: Timeline.....	22
Figura 11: Símbolos do gráfico de fluxo .....	22
Figura 12: Gráfico de fluxo .....	23
Figura 13: Rede semântica .....	24
Figura 14: Tabela periódica de métodos de VI.....	25
Figura 15: Mecanismo de Visualização de Informação .....	27
Figura 16: Visão geral da arquitetura do sistema Pulse .....	30
Figura 17: Visualização dos resultados do sistema Pulse.....	31
Figura 18: Visão geral da arquitetura do VISA .....	32
Figura 19: Múltiplas visualizações – Hotéis de Hong Kong.....	34
Figura 20: Metodologia.....	37
Figura 21: Arquitetura ferramenta SentimentALL .....	39
Figura 22: Arquitetura Módulo de Visualização de Informação .....	40
Figura 23: Mecanismo de Visualização de Informação .....	41
Figura 24: Dicionário de termos .....	42
Figura 25: Edição de um item do dicionário de termos .....	43
Figura 26: Tela – Mapa destinos .....	46
Figura 27: Aspectos mais relevantes por destino.....	46
Figura 28: Tela de seleção de destino .....	48
Figura 29: Tela - Aspectos por destino.....	49
Figura 30: Tela - Sentenças por aspecto .....	50
Figura 31: Construtor de taxonomias .....	52
Figura 32: Tela para adicionar aspectos ao nó folha.....	52

Figura 33: Nó folha adicionado à taxonomia .....53

## LISTA DE TABELAS

Tabela 1: Relação de aspectos .....	44
Tabela 2: Dicionário de termo – conjunto de aspectos.....	44
Tabela 3: Tabela de dados resultante .....	44

## **LISTA DE ABREVIATURAS E SIGLAS**

AS – Análise de Sentimentos

VC – Visualização de Conhecimento

VI – Visualização de Informação



## SUMÁRIO

<b>1 INTRODUÇÃO .....</b>	<b>9</b>
<b>2 REFERENCIAL TEÓRICO.....</b>	<b>12</b>
2.1 VISUALIZAÇÃO DE INFORMAÇÃO .....	12
2.1.1 <i>Definição</i> .....	12
2.1.2 <i>Técnicas de visualização</i> .....	12
2.1.3 <i>Interatividade</i> .....	27
2.2 TRABALHOS RELACIONADOS .....	29
<b>3 MATERIAIS E MÉTODOS OU METODOLOGIA.....</b>	<b>36</b>
3.1 POPULAÇÃO E AMOSTRA.....	36
3.2 MATERIAIS .....	36
3.3 PROCEDIMENTOS .....	37
<b>4 RESULTADOS E DISCUSSÃO .....</b>	<b>39</b>
4.1 ARQUITETURA .....	39
4.2 DESENVOLVIMENTO DO MECANISMO DE VISUALIZAÇÃO.....	41
4.2.1 <i>Dicionário de termos</i> .....	41
4.2.2 <i>Visualizações</i> .....	45
4.3 DESENVOLVIMENTO DO CONSTRUTOR DE TAXONOMIAS .....	51
<b>5 CONCLUSÃO .....</b>	<b>54</b>
<b>REFERÊNCIAS.....</b>	<b>55</b>
<b>APÊNDICES .....</b>	<b>60</b>

## 1 INTRODUÇÃO

Visualização da Informação (VI) é definida por Card et al. (1999) como o uso de representações visuais interativas de dados para ampliar a cognição. Isso significa que os dados podem ser transformados, por exemplo, em representações gráficas interativas para auxiliar no entendimento das informações.

De acordo com Yi et al. (2007), os sistemas de VI, em seu núcleo, são compostos por duas partes: representação e interação. O componente de representação tem origem nos campos de computação gráfica, e consiste em mapear os dados para a representação e como essa representação é processada na tela. O componente de interação envolve o diálogo entre o sistema e o usuário, e a forma como o usuário vai explorar os conjuntos de dados para descobrir novos conhecimentos.

O desenvolvimento de sistemas para Visualização de Informação (VI) tem se tornado cada vez mais importante, considerando a grande quantidade de informações que são geradas a todo instante. Segundo o escritor e consultor em big data<sup>1</sup> Bernard Marr (2015), nos últimos dois anos foram criados mais dados e, conseqüentemente, mais informações, do que em toda a história da humanidade e até 2020 cerca de 1,7 megabyte de novas informações serão criadas por segundo para cada habitante do planeta.

Nesse contexto, há a necessidade de sistemas que tenham a função de tornar possível a construção de visualizações que facilitem o processo de compreensão de um amplo conjunto de informações e que o usuário possa de alguma forma interagir com a apresentação.

A partir do estudo das técnicas de VI será possível desenvolver apresentações mais eficientes, que possibilitem ao usuário a abstração de informações a partir de um conjunto de dados. As técnicas podem colaborar com a eficiência do resultado, como a representação gráfica de *timeline*, que apresenta o relacionamento entre eventos (variáveis) com o tempo, criando uma visualização de fácil percepção do período que cada ação ocorreu em um espaço de tempo; gráfico de barra, que é utilizado para representar dados discretos, em que as barras verticais representam categorias e a altura da barra representa seu valor; a representação gráfica de mapa de árvore, que

---

<sup>1</sup> “Big data são conjuntos de dados muito grandes ou complexos, que os aplicativos de processamento de dados tradicionais ainda não conseguem lidar” (WIKIPEDIA, 2015).

é utilizada para apresentar dados hierárquicos na forma de retângulos aninhados ou em camadas e tem por objetivo ajudar o usuário a identificar padrões, comparar nós e sub nós em vários níveis de profundidade.

Uma outra característica importante na VI é a interação, que segundo Ltfi (2009, p. 219, tradução nossa), contribui com o controle que o usuário poderá ter do ambiente, o que pode implicar em um processo mais eficiente de extração da informação, diferentemente de um ambiente que ele só poderá observar de forma passiva.

Nesse contexto são consideradas as necessidades apresentadas pelo grupo de pesquisa em Engenharia Inteligente de Dados do CEULP/ULBRA, mais especificamente a demanda levantada no projeto relacionado ao ambiente SentimentALL de Brito *et al* (2015), este trabalho buscou responder ao seguinte problema de pesquisa: Como utilizar as técnicas de Visualização da Informação para organizar um amplo conjunto de dados sobre produtos turísticos de forma a possibilitar a verificação de padrões e tendências por pesquisadores e profissionais da área?

A hipótese abordada é de que com a competência técnica para implementar os processos de visualização da informação, é possível criar um ambiente de apresentação de informações sobre produtos turísticos que permita a profissionais da área verificar padrões e tendências. Portanto, o objetivo geral deste trabalho é criar um módulo de visualização de informação sobre produtos turísticos na ferramenta de análise de sentimentos – SentimentALL. Neste sentido objetiva-se especificamente:

- desenvolver um modelo de dados para o armazenamento no banco de dados e um modelo de dados para visualização;
- implementar estruturas para o armazenamento dos dicionários de termos;
- implementar construtor da taxonomia;
- implementar recursos de visualização da informação.

Um módulo de VI para a ferramenta SentimentALL permitirá a geração de novos conhecimentos acerca do domínio, pois as análises serão facilitadas, especialmente porque haverá a possibilidade de criação dinâmica de taxonomias para organizar as informações de acordo com uma série de categorias (e.g. por estados, destinos turísticos, por atrações, por facilidades).

O trabalho está estruturado da seguinte forma: o capítulo 2 apresenta o referencial teórico dividido entre Visualização de Informação, que apresenta definições, técnicas e interatividade em VI, e Trabalhos Relacionados, que apresenta

dois projetos (“*Pulse: Mining Customer Opinions from Free Text*” (GAMON et al., 2005) e “*VISA: A Visual Sentiment Analysis System*” (DUAN et al., 2012)) que utilizaram técnicas de VI em seu desenvolvimento. No capítulo 3 são apresentados a metodologia e os materiais utilizados no desenvolvimento deste trabalho. O capítulo 4 apresenta os resultados obtidos e como foi o processo de desenvolvimento do módulo de VI. No capítulo 5 são relacionadas as conclusões deste trabalho, e por fim as referências.

## **2 REFERENCIAL TEÓRICO**

### **2.1 Visualização de informação**

#### **2.1.1 Definição**

De acordo com Däßler e Palm (1998), o uso do termo Visualização de Informação como uma tecnologia que possibilitou a amplificação da visualização de dados deu-se no início dos anos noventa no centro de pesquisa da Xerox. Para Schumann e Muller (2004), a VI, a partir de então, tornou-se uma área de pesquisa autônoma e em ampla expansão. Segundo Freitas (p. 144, 2001), pode ser definida “como uma área da Ciência que tem por objetivo o estudo das principais formas de representações gráficas para apresentação de informações, a fim de contribuir para o entendimento delas”.

Card et al (1999, p. 6, *tradução nossa*) define VI como “o uso de representação visual interativa de dados abstratos suportada por computador para amplificar a cognição”, ou seja, conforme apresentado em Burkhard (2005) o conceito pode ser entendido como uma forma “de explorar e obter novos conhecimentos através da visualização de grandes conjuntos de informações”. Considerando novas definições apresentadas na literatura, Card (2007, p.542) reformulou o conceito de VI apresentando-o como “um conjunto de tecnologias que utilizam computação visual para amplificar a cognição humana através de informações abstratas”.

#### **2.1.2 Técnicas de visualização**

Considerando a sistematização apresentada em Khan e Khan (2011), as técnicas de visualização são organizadas em visualização de dados, que representa uma forma simplificada de interpretar e relacionar dados, e em visualização de informação, que segundo Bederson e Shneiderman (2003) são representações visuais interativas que possibilitam aos usuários ver, explorar e compreender grandes quantidades de informação de uma só vez.

### 2.1.2.1 Visualização de dados

“Visualização de dados é o estudo da representação de dados de uma forma sistemática, incluindo atributos e variáveis para a unidade de informação” (FRIENDLY, 2008, p.2, *tradução nossa*).

Algumas dessas técnicas serão apresentadas a seguir:

#### Tabela

Tabela é um tipo de visualização simples, de fácil entendimento e interpretação. (KHAN e KHAN, 2011) De acordo com Fink (2008) tabelas têm uma estrutura formatada, com dados em linhas e colunas. Além disso, têm diferentes variedades, flexibilidade, formatações, notações, representação e uso. A tabela apresentada na Figura 1 relaciona o dia e a hora com a temperatura e umidade. Para cada hora do dia há um dado sobre a temperatura e umidade daquele instante da medição, a máxima e a mínima da hora.

**Figura 1:** Tabela simples

Data	Hora	Temperatura (°C)			Umidade (%)		
	UTC	Inst.	Máx.	Mín.	Inst.	Máx.	Mín.
26/09/2015	00	30.1	31.6	30.1	34	34	32
26/09/2015	01	29.7	30.6	29.6	34	35	32
26/09/2015	02	28.1	29.8	27.0	40	50	34
26/09/2015	03	25.7	28.6	25.6	52	53	38
26/09/2015	04	25.2	26.3	25.1	53	53	46
26/09/2015	05	25.6	26.4	24.8	49	54	45
26/09/2015	06	25.3	26.0	24.9	57	60	48
26/09/2015	07	23.7	25.4	23.5	65	65	55
26/09/2015	08	23.9	24.1	22.9	63	70	62
26/09/2015	09	23.3	23.9	23.1	65	68	61
26/09/2015	10	25.3	25.3	23.0	60	69	59
26/09/2015	11	28.2	28.6	25.3	57	61	48
26/09/2015	12	31.2	31.2	28.2	46	58	46
26/09/2015	13	33.1	33.2	31.1	40	47	39
26/09/2015	14	35.1	35.1	33.1	33	40	33
26/09/2015	15	35.6	35.8	34.8	32	34	32
26/09/2015	16	36.3	37.2	35.3	30	33	29
26/09/2015	17	37.4	37.7	36.3	27	30	27
26/09/2015	18	38.2	38.3	36.7	23	28	23
26/09/2015	19	38.0	38.6	37.3	22	25	21
26/09/2015	20	37.4	38.1	37.3	21	23	19
26/09/2015	21	35.5	37.4	35.5	27	27	20
26/09/2015	22	32.5	35.5	32.5	32	32	27
26/09/2015	23	32.0	32.5	31.7	32	33	31

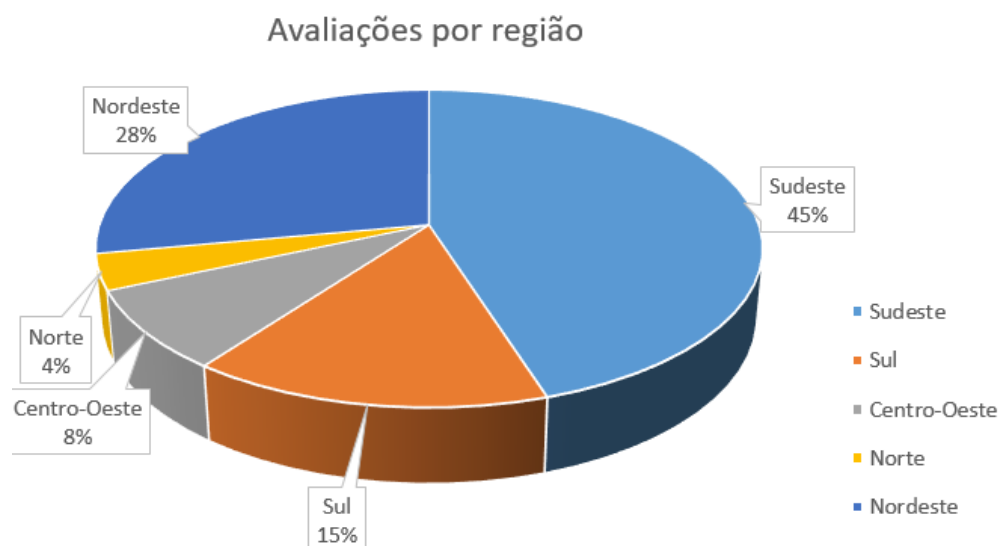
Fonte: Instituto Nacional de Meteorologia - INMET

### Gráfico de Torta

Gráficos de Torta são gráficos divididos em setores, em que cada setor representa o tamanho de alguma informação relacionada. Os gráficos de torta (também conhecidos como gráficos de pizza) são utilizados para mostrar o tamanho relativo das partes de um todo. De acordo com Few (2007) visualização em gráficos de torta são eficientes em alguns casos, por exemplo, comparar um segmento do gráfico com os outros segmentos, porém é difícil comparar diferentes gráficos de torta e diferentes segmentos quando são de dois ou mais gráficos.

No contexto do turismo e AS o exemplo abaixo (Figura 2) mostra um gráfico de torta que representa a proporção da quantidade de avaliações realizadas por região.

**Figura 2:** Gráfico de torta



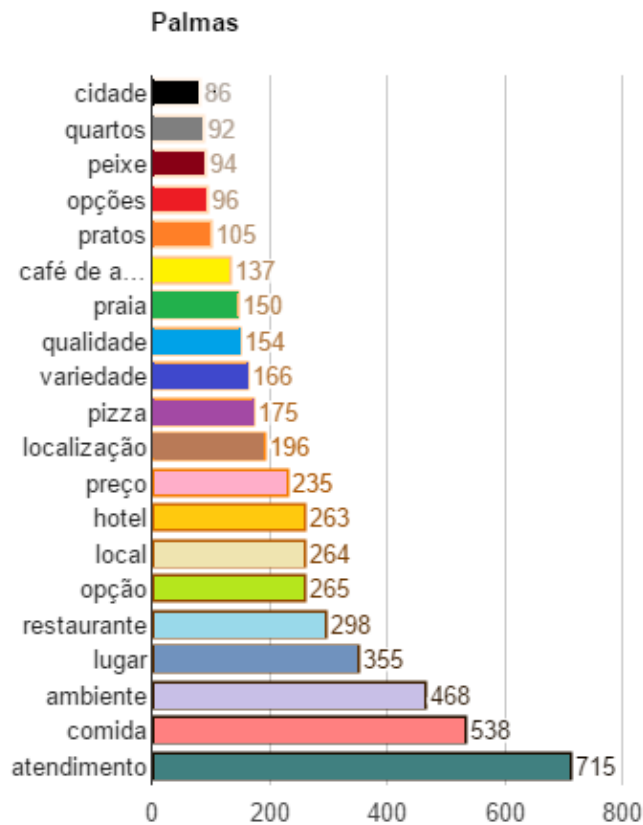
### Gráfico de barra

Também denominado de gráfico de coluna, Gráfico de barra é, muitas vezes, utilizado para representar dados discretos e não dados contínuos. As barras verticais representam categorias, a altura da barra indica seu valor, que pode ser, por exemplo, a soma dos valores da categoria que representa, as categorias podem ser, por exemplo, faixa etária, média de notas entre diversas outras possibilidades (TIBCO SPOTFIRE, 2012). A orientação das barras pode ser vertical ou horizontal.

No contexto do turismo e AS o exemplo abaixo (Figura 3) apresenta um gráfico de barras com os vinte aspectos mais relevantes para um determinado destino. No

gráfico foram utilizadas cores diferentes para cada barra para ajudar o usuário na identificação de cada aspecto.

**Figura 3:** Gráficos de barra

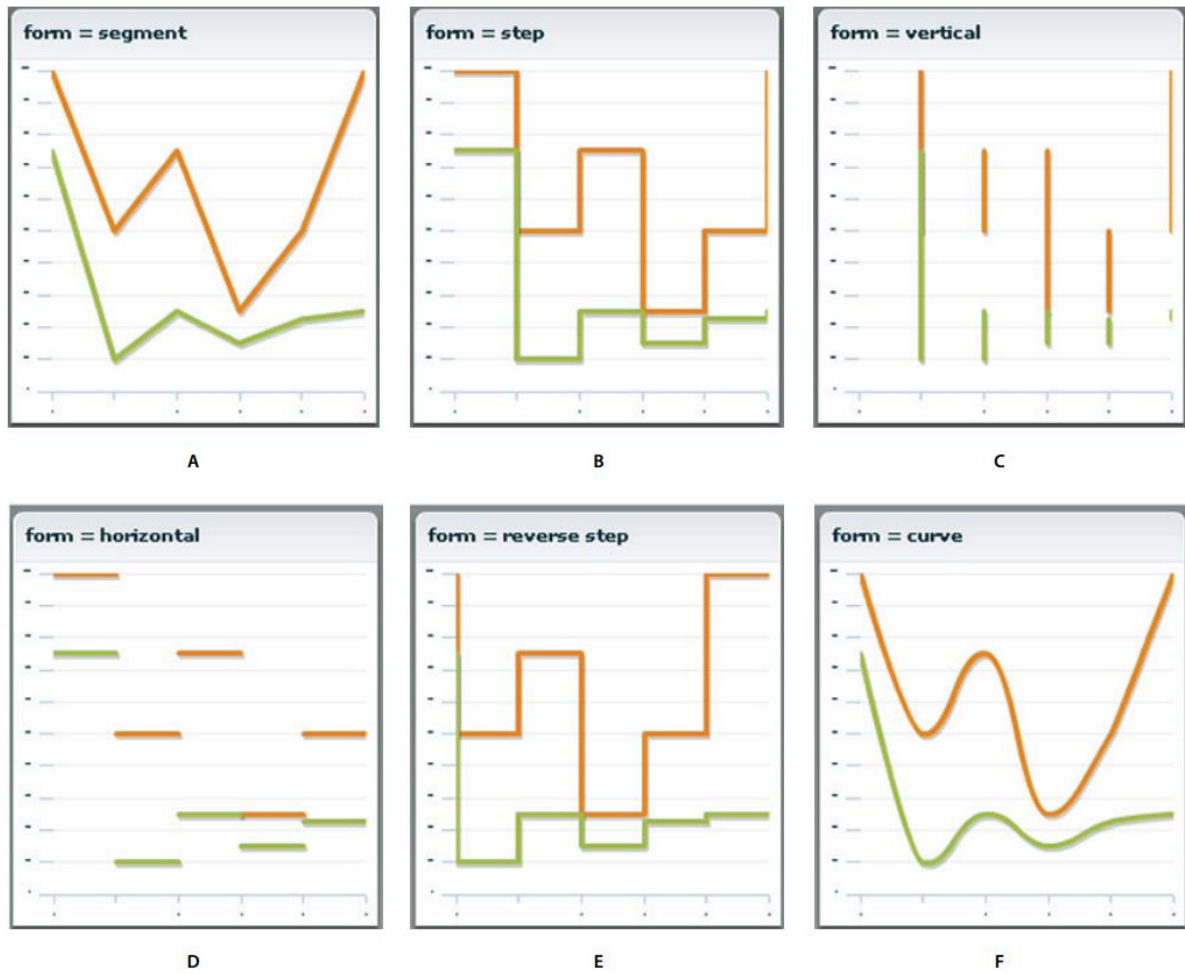


### Gráfico de linha

Gráfico de linha “apresenta os dados como uma série de pontos, em coordenadas cartesianas, ligados por uma linha contínua. Pode ser usado um ícone ou símbolo para representar cada ponto de dados ao longo da linha” (ADOBE FLEX, 2008). “Os gráficos de linhas são ideias para apresentar tendências ao longo do tempo, quaisquer dados que se comportam como uma função em relação ao eixo X pode ser utilizado. Os gráficos de linhas enfatizam o fluxo de tempo e a taxa de alteração” (TIBCO SPOTFIRE, 2012). A Figura 4 apresenta as diversas formas que um gráfico de linha pode assumir: A – segmento; B – passo; C – vertical; D – horizontal; E – passo reverso; F – curva.

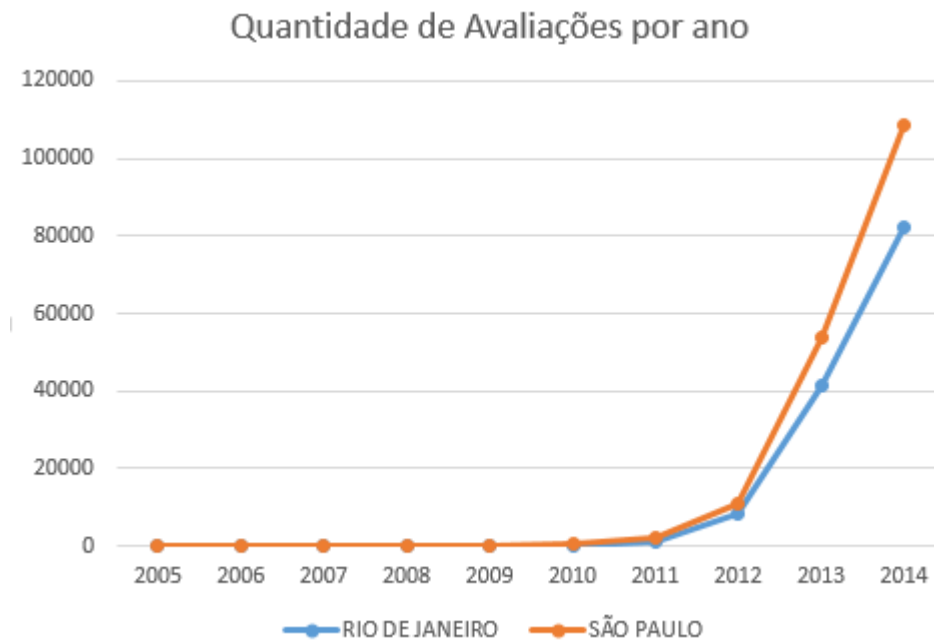


**Figura 4:** Formas de gráficos de linha



Fonte: Adobe Flex (2008)

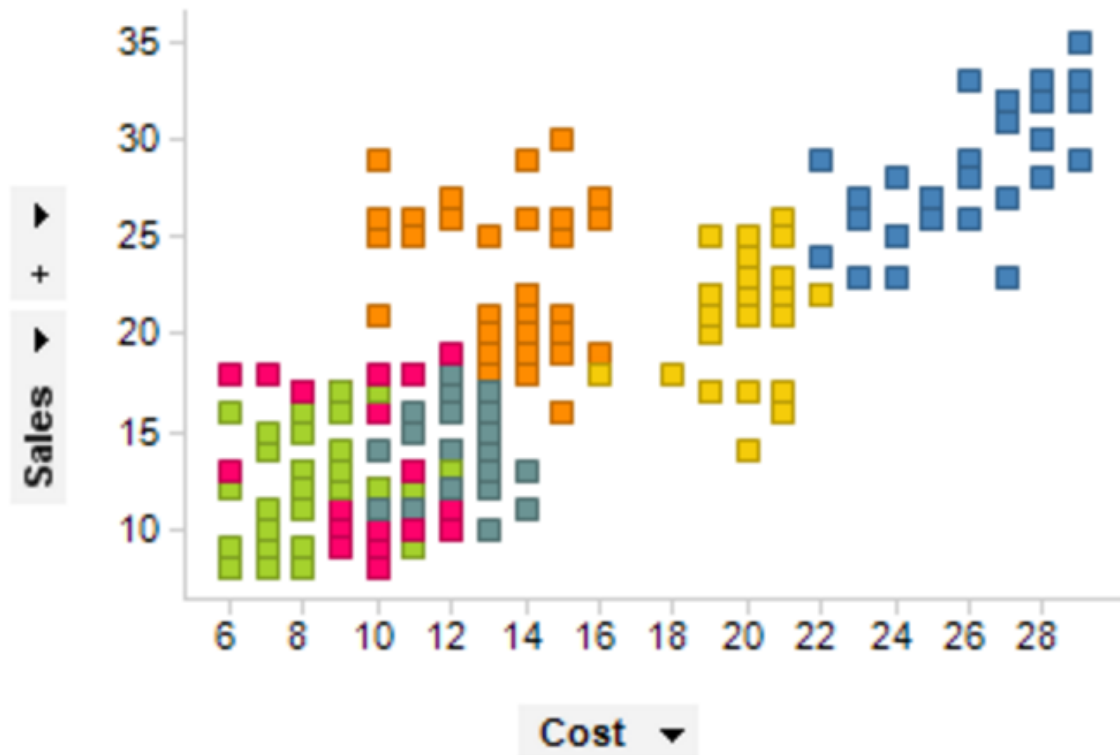
No contexto do turismo e AS o exemplo abaixo apresenta um gráfico de linhas com a quantidade de avaliações que foram feitas em duas cidades, Rio de Janeiro e São Paulo, entre os anos de 2005 e 2014. Através da análise do gráfico é possível perceber o quão grande foi o crescimento na quantidade de avaliações realizadas nos dois destinos.

**Figura 5:** Gráfico de linha

### Gráfico de dispersão

Gráfico de Dispersão “representa dados em coordenadas cartesianas, em que cada ponto tem um valor que determina a sua posição ao longo do eixo x, e um valor que determina a sua posição ao longo do eixo y” (ADOBE FLEX, 2008). O gráfico de dispersão pode mostrar a força que existe na relação entre as variáveis, a direção dessa relação e se existem valores que estão fora do padrão. No exemplo abaixo (Figura 6) são representadas vendas em função do custo para um número de diferentes produtos, para exibir uma baixa correlação positiva.

**Figura 6:** Gráfico de dispersão



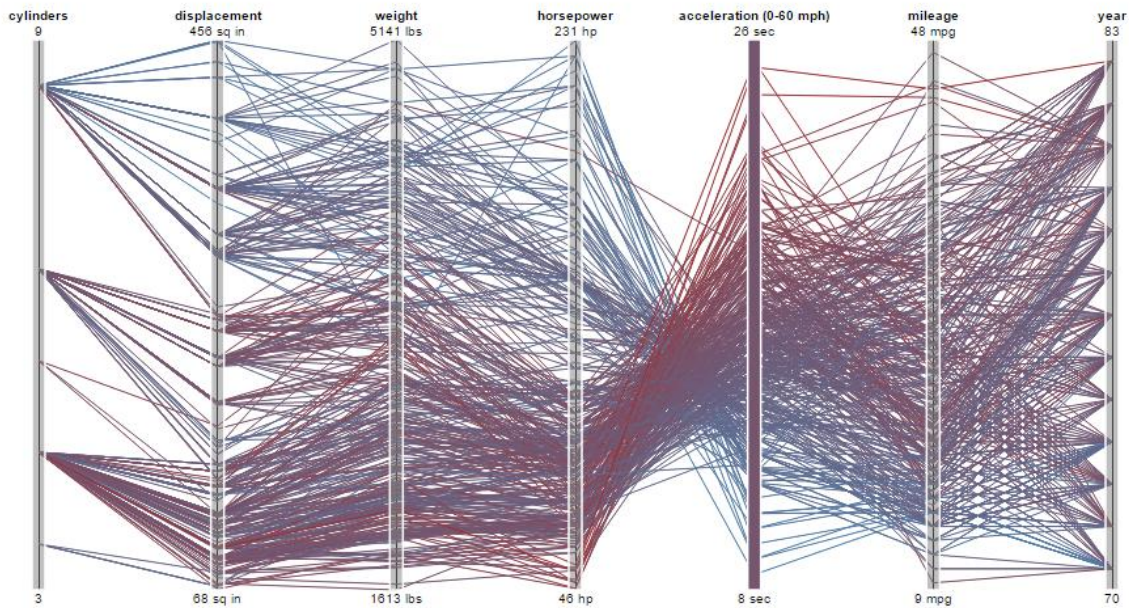
Fonte: Tibco Spotfire (2012)

### 2.1.2.2 Visualização de informação

As técnicas de VI possibilitam uma apresentação coerente de grandes conjuntos de dados, apresentam as informações sob vários pontos de vista e com diferentes níveis de detalhamento, oferecem suporte para comparações e relações, além de contar a história do dado, ou seja, colocá-lo em um contexto (HEARST, 2003).

#### *Coordenadas paralelas*

Coordenadas Paralelas são “uma técnica de visualização usada para plotar elementos de dados individuais através de muitas dimensões” (GEMIGNANI, 2010). Os eixos verticais correspondem as dimensões, as linhas são séries de pontos ligados ao longo das dimensões/eixos e representam um elemento de dado. A maioria dos métodos de visualização não conseguem representar dados multidimensionais, por esse motivo a técnica de coordenadas paralelas é importante, pois pode apresentar inúmeras dimensões de acordo com Few (2006).

**Figura 7: Coordenadas paralelas**

Fonte: Protovis (2015)

Na Figura 7 é apresentado um gráfico de coordenadas paralelas com os dados de centenas de carros, são informações como cilindrada, cavalos de potência, ano, peso entre outros. É possível identificar as linhas que representam os carros, e assim fazer a comparação entre cada uma delas.

### *Mapa de árvore*

De acordo com Shneiderman (2009), Mapa de árvore é uma técnica utilizada para apresentar dados hierárquicos na forma de retângulos aninhados ou em camadas. Para Khan e Khan (2011) mapa de árvore ajuda o usuário a identificar padrões, comparar nós e sub nós em vários níveis de profundidade. Cada item ou objeto de dado é representado por um retângulo, que tem diferentes tamanhos de acordo com os atributos definidos pelo usuário ou propriedades definidas pelo usuário.

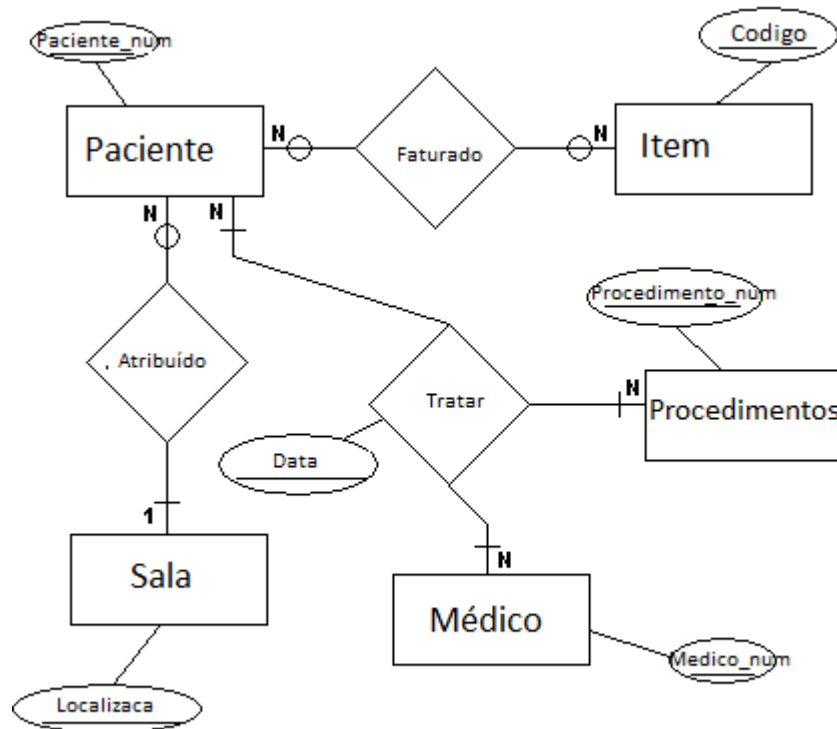
**Figura 8:** Mapa de árvore

No contexto do turismo e AS o exemplo acima (Figura 8) ilustra um gráfico de mapa de árvore que apresentam estados (São Paulo, Rio de Janeiro e Tocantins), cidades (São Paulo, Rio de Janeiro e Palmas) e aspectos (atendimento e café da manhã) e, são dimensionados e coloridos de acordo com a quantidade de vezes que o aspecto aparece nas avaliações.

#### *Diagrama de entidade relacionamento*

O Diagrama de entidade relacionamento é construído através do Modelo Entidade-Relacionamento, que é um artefato de engenharia de software. “É basicamente uma metodologia de modelagem de banco de dados que é usada para gerar o esquema conceitual, esquema semântico e o esquema de banco de dados relacional de um sistema” (KHAN; KHAN, 2011, p.8, tradução nossa). Os elementos chave deste diagrama são Entidade (objeto ou conceito sobre o qual a informação será armazenada); Relacionamento (como duas ou mais entidades compartilham informações); e Atributos (qualquer característica única da entidade).

**Figura 9:** Diagrama entidade relacionamento



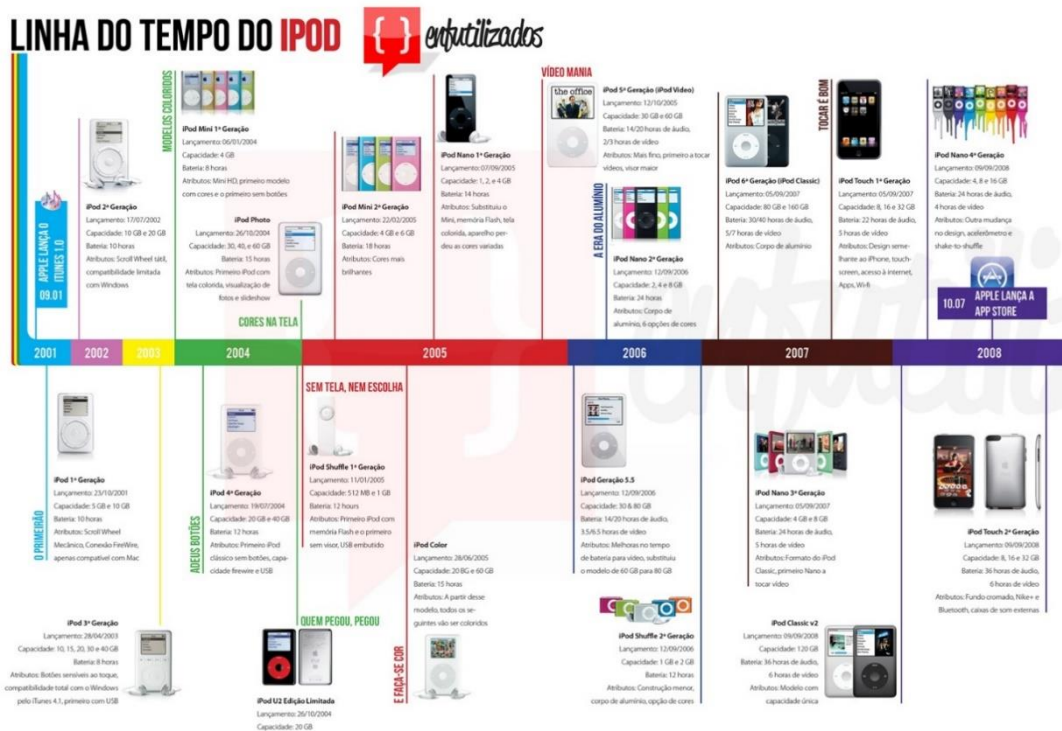
Fonte: Khan e Khan (2011, p.8, tradução nossa)

A Figura 9 representa um diagrama entidade relacionamento de um hospital. Os retângulos são as entidades, as elipses os atributos e os relacionamentos são os losangos.

### *Timeline*

Timeline “é uma representação cronológica de eventos ao longo de uma linha que permite ao leitor entender rapidamente as relações temporais” (ROUSE, 2005) e pode ser linear ou comparativa. Timeline linear apresenta os eventos de acordo como eles ocorreram em um certo período do tempo, pode ser vertical ou horizontal. Timeline comparativa apresenta dois ou mais temas que ocorreram no mesmo período de tempo.

Figura 10: Timeline



A Figura 10 apresenta a linha do tempo linear de um produto, para cada ano são mostrados os eventos que foram relevantes para a mudança tecnológica, como data de lançamento, características gerais e imagens.

Gráfico de Fluxo

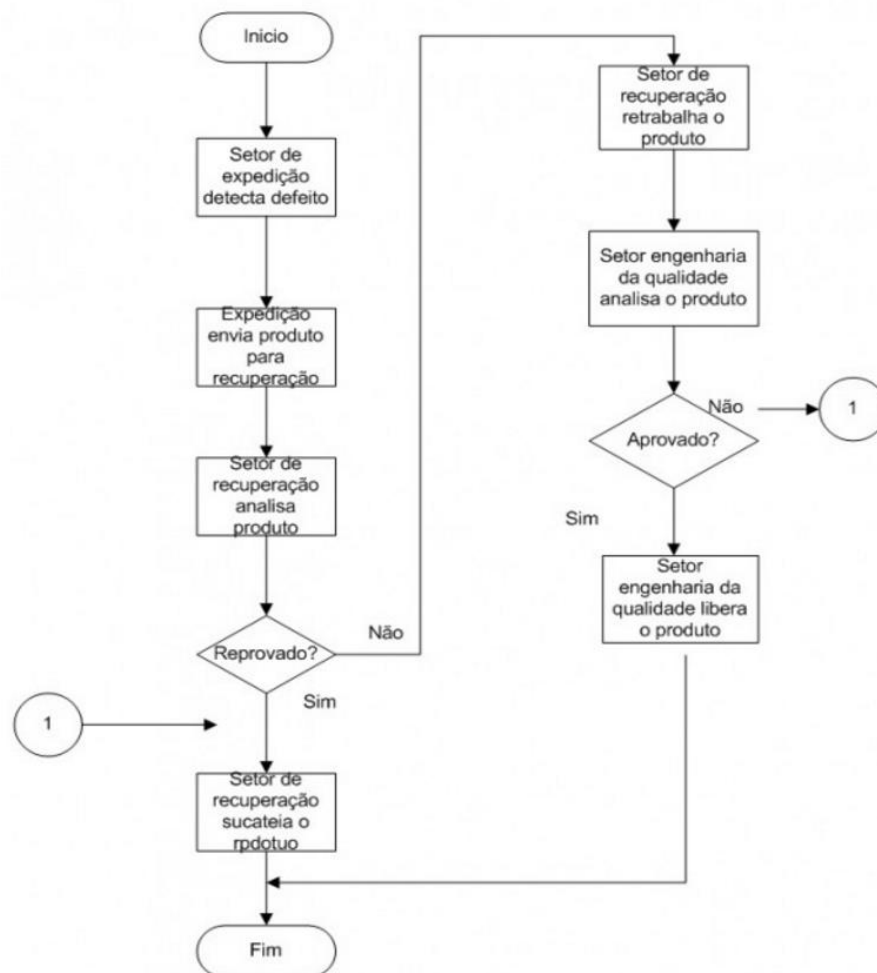
De acordo com Hebb (2015), Gráfico de fluxo é uma representação gráfica ou simbólica das etapas de um processo. As etapas do processo são representadas por símbolos diferentes em cada uma, além do símbolo há uma breve descrição de cada passo. São utilizadas setas para fazer a ligação entre uma (símbolo) e outra, de forma que a orientação das setas indicam a direção do fluxo. Cada símbolo tem um significado diferente, alguns são os representados na Figura 11.

Figura 11: Símbolos do gráfico de fluxo



A Figura 12 é um exemplo de gráfico de fluxo, ela representa um procedimento de controle de produto não conforme.

**Figura 12:** Gráfico de fluxo

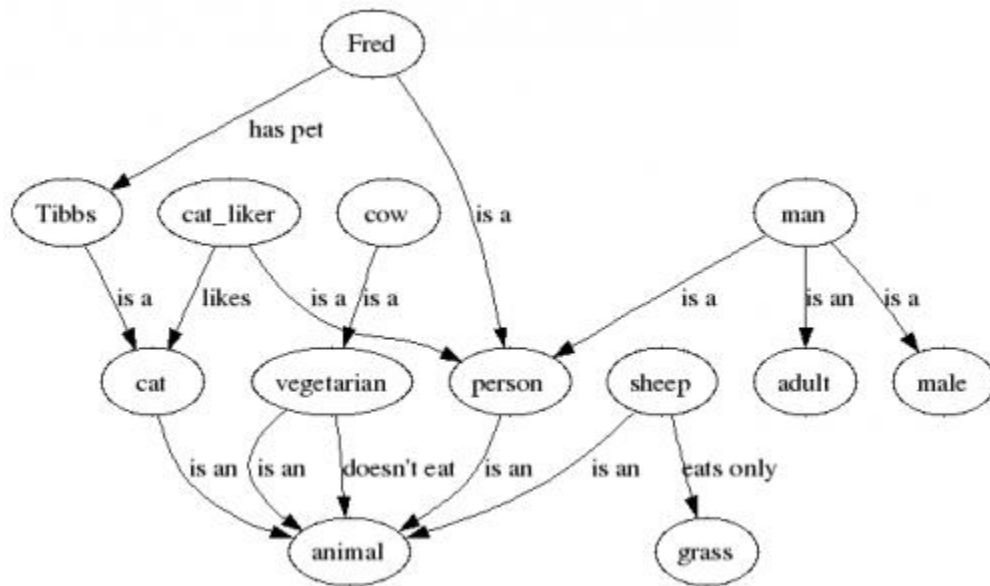


Fonte: Martins (2015)

### *Rede Semântica*

De acordo com Khan e Khan (2011, p. 9, tradução nossa), Rede semântica “é uma representação gráfica da relação lógica entre diferentes conceitos”. “São esquemas de representação do conhecimento, envolvendo nós (objetos ou conceitos) e links (ligação) entre os nós. Os links são direcionados e rotulados, assim uma rede semântica é um grafo direcionado” (MCCANN, 2015, tradução nossa)



**Figura 13:** Rede semântica

Fonte: McCann (2015)

A Figura 13 representa uma rede semântica que relaciona pessoas, animais, sexo, tipo de alimentação. Cada elipse é um objeto, que tem sua relação com outro objeto identificada pela seta de orientação e a frase que mostra o tipo de relação.

### 2.1.2.3 Tabela periódica

Meyer (2009) apresentou em seu trabalho uma tabela periódica de métodos de VI (Figura 14) que foi elaborada por Lengler e Eppler (2007). A tabela (Figura 14) é dividida em grupos, separados por cores, a partir da seguinte categorização:

- Dados: representação de dados quantitativos, em forma estruturada;
- Informação: uso de representação visual e interativa, com o objetivo de ampliar a cognição;
- Conceito: métodos para elaborar conceitos quantitativos, ideias, planos e análises;
- Metáfora: apresenta uma visão da informação através das características-chave da metáfora. São apresentadas graficamente, para organizar e estruturar informações;
- Estratégia: utiliza de forma sistemática representações visuais na análise, desenvolvimento, formulação e comunicação.

- Visualização composta: utiliza diversos formatos de representação gráfica em uma única apresentação.

Figura 14: Tabela periódica de métodos de VI

C continuo													G facilitação gráfica	Visualização de Dados Representação visual de dados quantitativos, em forma esquemática (com ou sem eixos)						
>O< Tb tabela	>O< Cc coordenadas cartesianas													>O< Tu traço de união	>O< Mt mapa metro	O Tm templo	<O> Th template de história	>O< A árvore	>O< Ct caricatura	Visualização de Informação Uso de representação visual interativa de dados para ampliar a cognição. Isto significa que os dados são transformados em uma imagem, que é mapeada para o espaço da tela. A imagem pode ser alterado pelos usuários à medida que trabalha com ela.
>O< Pi gráfico de pizza	>O< L gráfico de linha													>O< De diagrama de comunicação	>O< Py plano de voo	>O< E conceito de esqueleto	O Pt ponte	>O< Fu funil	O Rr retrato rico	Visualização de Conceito Métodos para elaborar (principalmente) conceitos qualitativos, idéias, planos e análises.
>O< B gráfico de barra	>O< Ga gráfico de área	>O< R gráfico de radar	>O< Cp coordenadas paralelas	>O< Ah árvore hiperbólica	>O< De diagrama de ciclo	>O< Lt linha do tempo	>O< Dv diagrama de Venn	<O> Mm mapa mental	<O> Qo quadro de oposição	>O< Cs círculos concêntricos	>O< Ar slide de argumento	>O< Ps diagrama de pasta de Swim	>O< Cg gráfico de Gant	<O> Dp diagrama de perspectivas	>O< D diagrama de dilema	<O> Rp régua de parâmetro	O M mapa de conhecimento	Visualização de Estratégia Utilização sistemática de representações visuais na análise, desenvolvimento, formulação, comunicação e implementação de estratégias nas organizações.		
>O< Hi histograma	>O< Gd gráfico de dispersão	>O< Ds diagrama de Sankey	>O< Li lente de informação	>O< De diagrama de entidade de relacionamento	>O< Rp rede de Petri	>O< Fg fluxograma	<O> Ci clustering	>O< Gc gráfico de camada	>O< Pm técnica de pirâmide de Mintto	>O< Ce cadeias de cause e efeito	>O< Mt mapa de Toulmin	>O< Ad árvore de decisão	>O< Mh método caminho crítico	<O> Cf Concept fan	>O< Co mapa de conceito	>O< Ic icebergue	O Ma mapa de aprendizado	Visualização de Metáfora Metáforas Visuais posicionam graficamente informações para organizar e estruturar informações. Eles também transmitem uma visão sobre a informação representada através das características-chave da metáfora empregada.		
>O< De diagrama de caixa	>O< Eg espectro grama	>O< Md mapa de dados	>O< Ma mapa de árvore	>O< Ac árvore cone	>O< Ss sistema de simulação dinâmica	>O< Df diagrama de fluxo de dados	<O> Rs rede semântica	>O< Mh mapa de sistema suave	>O< Si mapa de sinergia	<O> Df diagrama força de campo	>O< Mi mapa de argumentação de Ibis	>O< Ep cadeia de evento de processo	>O< Cp gráfico de Pert	<O> Mce mapa de conhecimento evocativo	>O< V diagrama de Vee	<O> Hb gráfico céu e inferno	O I mural de informação	Visualização Composta Utilização complementar de diferentes formatos de representação gráfica em um único esquema ou quadro.		
De Visualização de Processo	Ah Visualização de Estrutura	O Visão Geral	□ Detalhe	○ Detalhe E Visão Geral	<> Pensamento Divergente	>> Pensamento Convergente	>O< Fd curva de fornecimento de demanda	>O< Pe gráfico de performance	>O< Me mapa de estratégia	>O< Go gráfico de organização	<=> Cq casa de qualidade	>O< Fb diagrama de feedback	= Af árvore de falha	>O< Qm quadrante mágico	>O< Cv diagrama ciclo da vida	>O< Fp cinco forças de Porter	<=> Sc ciclo s	>O< Su mapa de suporte	○ Is diagrama de Ishikawa	O Rt roteiro de tecnologia
O Ed caixa de Edgeworth	>O< Pf diagrama de portfólio	O Je jogo de estratégia	>O< Om organograma de Mintzberg	<=> Z caixa morfológica de Zwicky	<O> Da diagrama de afinidade	= Dd Diagrama de descoberta de decisão	>O< Mb matriz bcg	>O< Es canvas de estratégia	>O< Va cadeia de valor	<=> Ch ciclo hype	>O< Mes Mapa de classificação de suporte	>O< To torneira	<=> Ds diagrama de spray							

Fonte: Lengler e Eppler (2007, tradução nossa)

Na Figura 14 pode ser verificado que, para cada categoria, a complexidade do método é representada pelo número do período. Quanto maior for o número da linha mais complexo é o método dentro do seu grupo. A tabela original possui um processo de interatividade que, ao passar o mouse sobre cada elemento da tabela são apresentadas mais informações sobre seu método, o processo de cognição necessário e as quais são as informações representadas. Analisando a tabela é possível ter “uma visão sobre uma grande quantidade de métodos diferentes de VI e ajuda a encontrar os métodos necessários para os problemas correspondentes” (Meyer, 2009, p.7, tradução nossa).

### 2.1.2.4 Processo de Visualização

Antes de escolher qual técnica utilizar é necessário considerar algumas questões apresentadas por Mazza (2004) e compiladas a partir dos trabalhos de

Spence (2001, apud MAZZA, 2004, p.10), Card et al. (1999), Hearst (2003) e Reed e Heller (1997, apud MAZZA, 2004, p.10):

- **O Problema:** o que será apresentado, encontrado ou demonstrado?
- **A natureza dos dados:** os tipos de dados podem ser numéricos (e.g. os números inteiros), ordinais (são dados não numéricos que tem uma ordenação convencional, e.g. meses do ano) ou categóricos (dados não ordenados, e.g. nome de países).
- **Número de dimensões dos dados:** dependendo do número de dimensões, as representações são chamadas de manipulação de dados univariada, bivariada, tri-variada e multivariada. Mapear e interpretar até três dimensões é fácil, porque as pessoas percebem o mundo em três dimensões. Lidar com mais de três dimensões representa uma das tarefas mais desafiadoras da VI.
- **Estrutura dos dados:** pode ser linear (e.g. listas alfabéticas, tabelas, *arrays*), temporal (dados que mudam durante o tempo, e.g. probabilidade de chuva), espacial ou geográfica (geralmente é considerado com um assunto de visualização científica e não como de visualização de informação), hierárquica (dados que surgem de taxonomias, gerenciamento de disco, genealogias) e redes (descrevem a estrutura gráfica com nós (uma variável) e link (representam a relação entre os nós)).
- **Tipo de interação:** se a representação gráfica é estática (uma imagem estática mostrada na tela), transformável (usuário pode interagir com a apresentação, utilizando ações de filtro, zoom ou seleção) ou manipulável (usuário pode selecionar quais dados quer que sejam apresentados).

Para sistematizar a VI, Chittaro (2006) definiu seis etapas que devem ser aplicadas no processo: mapeamento, que é a transformação dos dados para a forma visual; seleção, o processo de selecionar os dados que serão apresentados de acordo com os objetivos estabelecidos para a ferramenta; apresentação, que tem relação com a forma de gerenciar e organizar as informações no espaço disponível de forma eficiente; interatividade, as facilidades disponíveis para o usuário, como organizar, explorar e rearranjar os dados; usabilidade, as práticas de desenvolvimento que utilizam conceitos de Interação Homem-Computador, como conhecimento em

percepção visual e cognição; e a última etapa é a avaliação que tem por objetivo avaliar se o método de VI utilizado foi ou não eficaz.

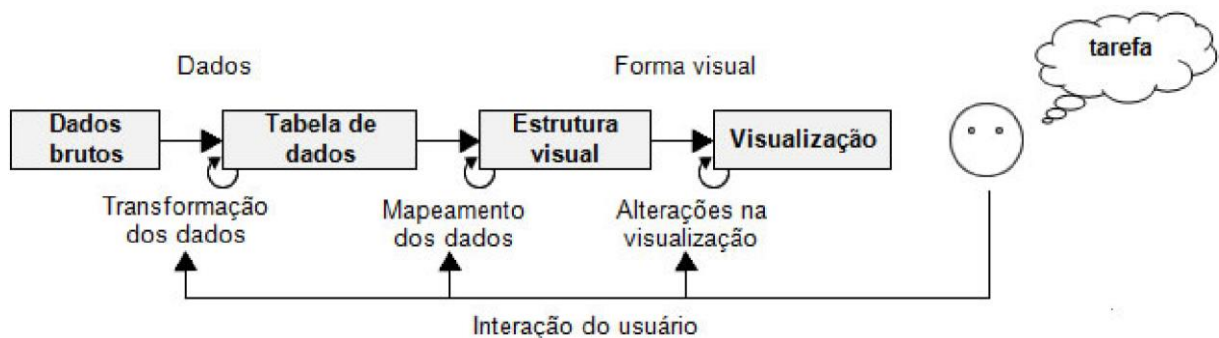
De uma forma geral, para responder questões que requerem o uso de VI, considerando o mapeamento adequado dos dados, os meios para transformá-los em informação eficiente, é necessário conhecimento do contexto em que a VI será aplicada, além de um estudo sobre os processos e métodos que viabilizam a área.

### 2.1.3 Interatividade

Interação é uma parte essencial de VI. Para YI et al. (2007), sem interação, uma técnica de VI torna-se um conjunto de imagens estáticas ou imagens animadas de forma autônoma. Ltifi et al (2009) deixa claro que o ser humano consegue abstrair mais informação quando ele pode, de alguma maneira, manipular a apresentação dos dados. Sendo assim, é fundamental que tais ações estejam disponíveis ao usuário. Os usuários estão interessados não somente em dados comumente disponíveis, mas também em dados abstratos, mesmo sem ter conhecimento prévio sobre eles. A partir desse ponto se torna mais importante o uso da interatividade para exploração, análise e representação dos dados de acordo com Khan e Khan (2011).

Para interação com a visualização o usuário pode utilizar diversas ações, como clique do mouse, passar o mouse sobre um elemento, duplo clique, entre outros. Há diversas maneiras da interação acontecer. Um mecanismo interativo para visualização foi desenvolvido por Card et al (1999). Tal mecanismo descreve a interação do usuário em diferentes etapas (Figura 15).

**Figura 15:** Mecanismo de Visualização de Informação



Fonte: Tradução de Card, Mackinlay e Shneiderman (1999)

Seguindo a interpretação de Moissa et al. (2014), o mecanismo apresentado na Figura 15 permite ao usuário alterar os dados e a estrutura de visualização, de

forma que o sistema realize o reprocessamento de tais dados. A partir da ação do usuário o sistema deve voltar até a etapa do mecanismo que seja adequada, para então reapresentar a visualização.

Alguns passos foram definidos para elaborar visualizações interativas:

### *Seleção*

Segundo Khan e Khan (2011), a seleção interativa de entidades ou de uma parte do conjunto de dados é um dos mais importantes e fundamentais requisitos em uma visualização. A seleção é importante para apresentar mais detalhes, destacar e identificar entidades. Há duas maneiras de realizar a seleção, de forma direta – a apresentação deve permitir essa interação diretamente; e indireta – por exemplo, selecionar os dados em uma tabela fora da apresentação.

### *Filtro*

O uso de filtros permite que o usuário ajuste de forma dinâmica a quantidade de informações que são apresentadas, ou seja, aumentar ou diminuir a quantidade de dados para a apresentação, mantendo o foco no que lhe é mais interessante. Para Ltifi et al (2009) uma vantagem do uso de filtro é a possibilidade de amplificar dinamicamente os detalhes em torno do centro de interesse, sem perder a visão geral do contexto.

### *Linking*

“A ideia é combinar diferentes métodos de visualização para superar as imperfeições de técnicas simples” (LTIFI et al., 2009, p.224, *tradução nossa*). Linking busca reunir as informações que tenham algum tipo de vínculo entre si, de forma a criar outros pontos de vista e apresentar diferentes perspectivas.

### *Zoom*

A ideia é começar com uma visão geral dos dados e permitir a apresentação de detalhes de acordo com a resolução escolhida (LTIFI et al., 2009, p.223, *tradução*

nossa). A navegação por zoom torna possível a exibição de um determinado ponto da visualização com o um nível de detalhamento elevado, porém perdendo a visão do contexto. As ações de *zoom in* e *zoom out*, aproximar e afastar, respectivamente, são de fácil manuseio, mas em alguns casos o usuário pode ficar perdido dentro da apresentação.

### *Distorção*

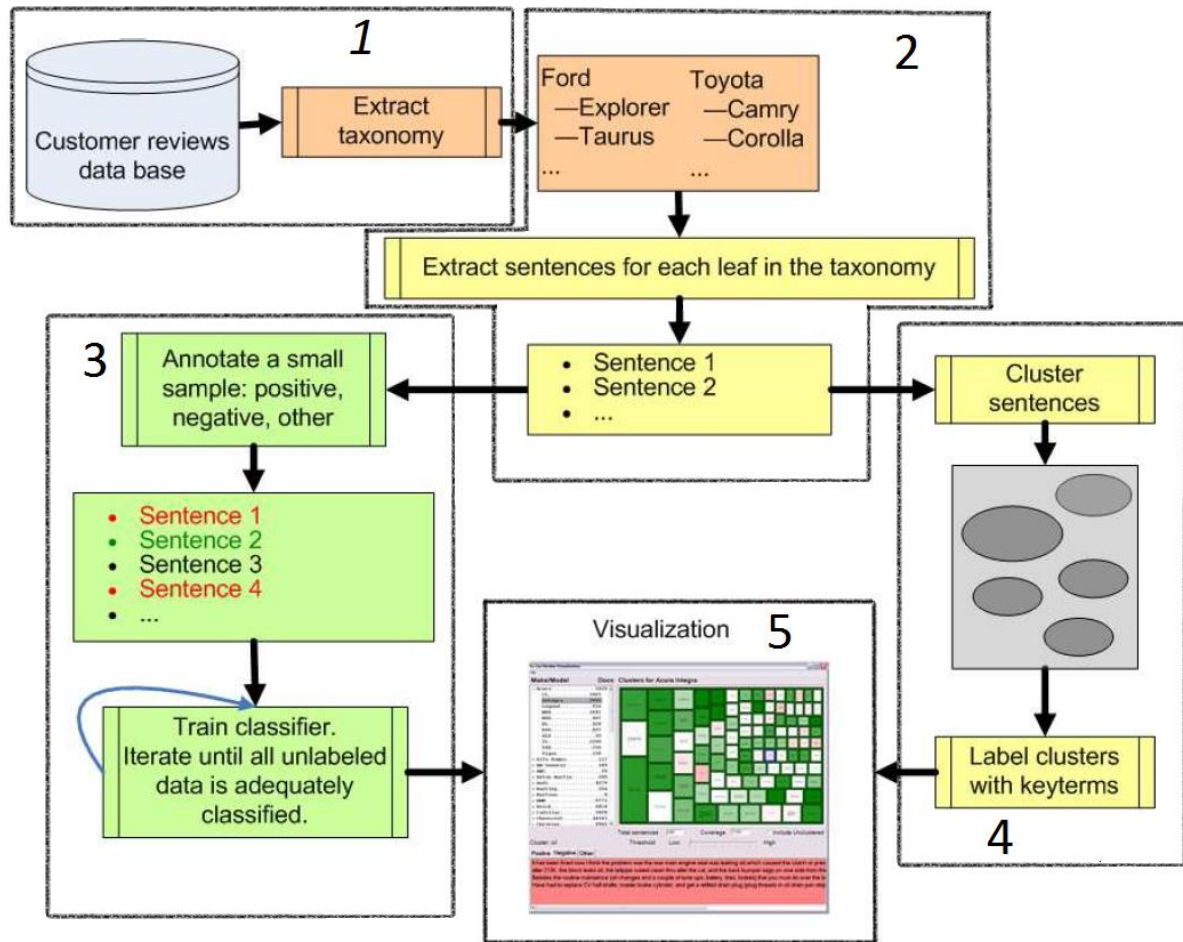
“A ideia básica é mostrar parte dos dados com um nível de detalhes elevado enquanto os demais são apresentados com menos detalhes’ (LTIFI et al., 2009, p.223, *tradução nossa*). São realizadas modificações que criam focos associados a visão do contexto, permitindo assim que, tanto o foco como o contexto sejam exibidos de maneira simultânea. São utilizadas geralmente deformações hiperbólicas ou esféricas para mostrar os dados que tem maior grau de detalhamento, Khan e Khan (2011) chamam tal forma de distorção de “olho de peixe”.

## **2.2 Trabalhos relacionados**

Dois trabalhos serão apresentados nesta seção, “*Pulse: Mining Customer Opinions from Free Text*” (GAMON et al., 2005) e “*VISA: A Visual Sentiment Analysis System*” (DUAN et al., 2012). Ambos desenvolveram ferramentas que envolvem a visualização de informação para um processo de análise de sentimentos.

O trabalho desenvolvido por Gamon et al. (2005) teve como objetivo a criação de uma ferramenta para a exploração de uma grande quantidade de textos sobre *feedbacks* de clientes, de forma que o usuário pudesse examinar os dados de uma forma geral ou com maior nível de detalhes. A descrição da técnica utilizada para agrupamento de sentenças, abordagem inicial sobre classificação de sentimentos e uma nova interface para o usuário foram discutidas neste trabalho. *Pulse* é o nome do sistema desenvolvido e sua arquitetura é apresentada na Figura 16.

**Figura 16:** Visão geral da arquitetura do sistema Pulse



Fonte: Gamon et al. (2005)

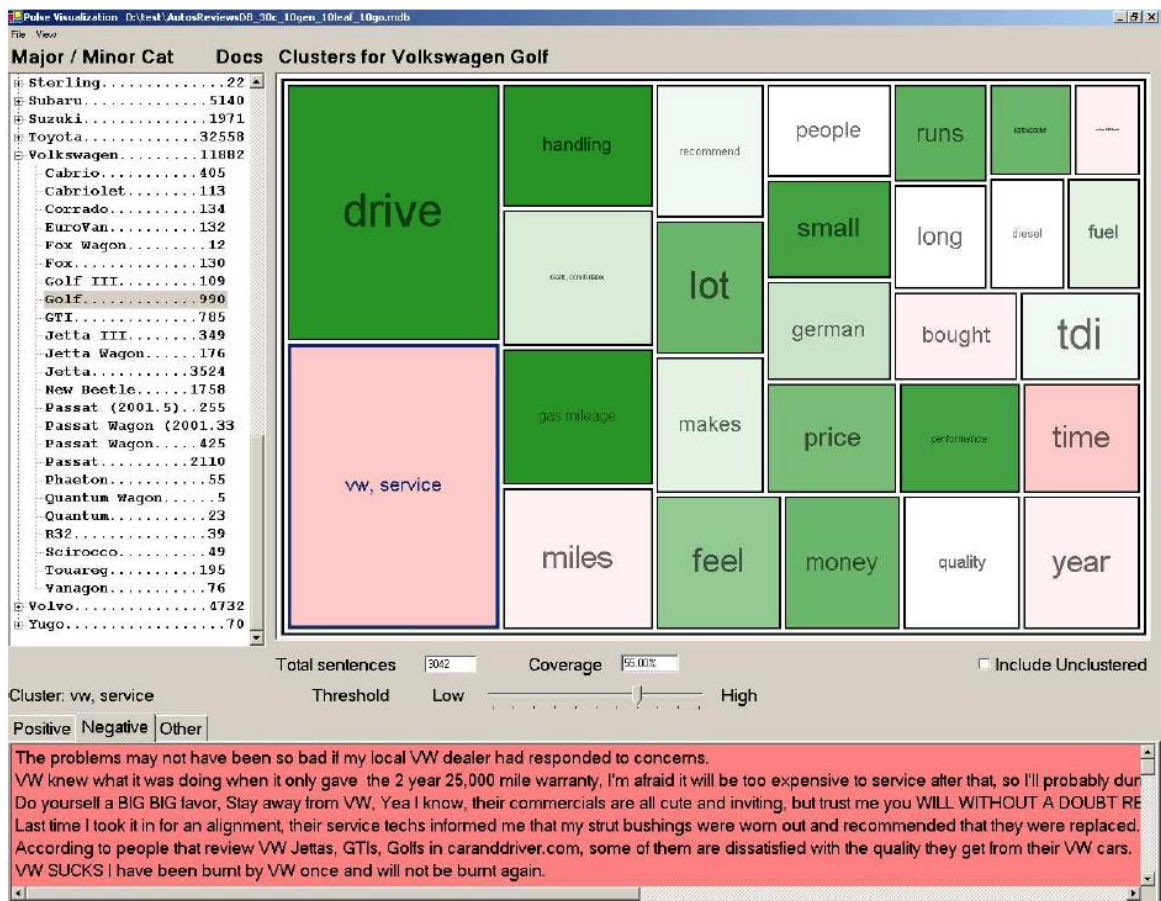
Para os testes do sistema foi utilizada uma base de dados com comentários sobre carros. Como ilustra a Figura 16, o funcionamento do sistema, de uma forma geral, se dá da seguinte maneira: é feita a extração da taxonomia das categorias principais (marca/modelo) em (1), a partir de uma consulta ao banco de dados. Em seguida as frases são extraídas dos comentários de cada marca e modelo e processadas de acordo com as duas dimensões de informação que o sistema quer expor ao final do estágio de visualização, sentimentos e tópicos (2). Para treinar a classificação de sentimentos um pequeno conjunto de dados é avaliado manualmente (positivo, negativo e outros) e utilizado com os demais para inicializar a classificação (3). Em seguida o componente de agrupamento forma grupos de conjuntos de sentenças que correspondem a um nó folha na taxonomia (um modelo de carro específico, por exemplo), então os grupos são rotulados com os termos chave mais



proeminentes (4). Após todas as sentenças terem sido classificadas e agrupadas, o componente de visualização exibe os resultados (5).

O componente de visualização precisa apresentar de forma simultânea as duas dimensões da informação: sentimento e tópico. Para apresentar os dados é utilizada a técnica de mapa de árvore, sendo que cada bloco representa um grupo, o tamanho do bloco representa a quantidade de sentenças presentes e a cor representa a média dos sentimentos de cada bloco (Figura 17). As cores são vermelho (grupos negativos), verde (grupos positivos) e branco (quantidade de positivos e negativos iguais ou a maioria foi classificada como “outros”). Cada bloco é rotulado com um termo chave.

**Figura 17:** Visualização dos resultados do sistema Pulse



Fonte: Gamon et al. (2005)

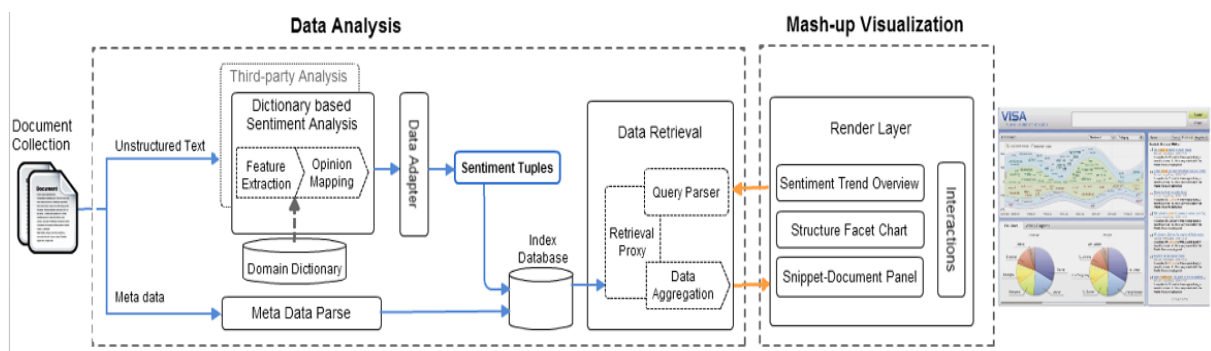
A Figura 17 apresenta a forma como os dados são disponibilizados ao usuário. A taxonomia das marcas e modelos são apresentadas à esquerda. À direita é visualizado o mapa de árvore, sendo que cada retângulo representa um grupo de sentenças, o tamanho de cada item indica a frequência de sentenças dentro do grupo,



e as cores mostram a polaridade de cada bloco, e na parte inferior são apresentadas as sentenças. Neste exemplo, o usuário escolheu o carro *Volkswagen Golf*, assim, os blocos apresentados são referentes a este objeto. O usuário, ao clicar no bloco “vw, service”, tem as sentenças apresentadas, separadas pela polaridade na parte inferior da tela. Há diversas opções de interação que o usuário pode utilizar para modificar os dados apresentados a ele.

Já o trabalho elaborado por Duan et al. (2012), teve como objetivo a construção de uma tupla de sentimentos genéricos, baseada em um modelo de dados e consumindo diferentes facetas de dados de sentimentos, com pontos de vista de múltiplas coordenadas, logo é escalável para trabalhar com a maioria dos sistemas de análise de sentimentos. O *VISA* foi desenvolvido a partir de uma ferramenta de visualização de texto genérico chama *TAIRA*. Diversas melhorias e mudanças foram introduzidas, como um novo modelo de dados para suportar um maior refinamento da análise de sentimentos. No centro do modelo de dados está a tupla de sentimentos, que é composta por quatro elementos: característica, aspecto, opinião e polaridade. Esse modelo serve como uma interface de dados entre o *backend* do sistema de análise de sentimentos e o *frontend* de visualização.

**Figura 18:** Visão geral da arquitetura do *VISA*



Fonte: Duan et al. (2012)

A Figura 18 apresenta a arquitetura do *VISA* de uma forma geral. O sistema tem como entrada uma coleção de documentos com dados brutos e constrói um sumário de sentimentos visuais como os dados de saída. O sistema é composto por dois componentes principais: análise de dados e visualização. O componente de análise de dados extrai as informações de sentimentos do texto e estrutura os resultados para apresentação. O componente de visualização recebe os dados pré-

formatados e cria um conjunto de visualizações para ajudar a entender os resultados da análise de sentimentos.

O componente de análise de dados inicialmente separa os dados em texto não estruturado e metadados estruturados, então eles são processados em dois caminhos diferentes. O texto não estruturado é analisado pela ferramenta de análise de sentimentos, que pode ser a ferramenta padrão de dicionário do VISA ou alguma ferramenta externa, então o resultado é transformado em um modelo de dados genéricos, ou seja, as tuplas de sentimentos, que consistem em quatro elementos – aspecto, característica, opinião e polaridade. No outro caminho, os metadados são analisados em uma lista de pares campo-valor para utilização posterior. Depois desse processamento, as tuplas de sentimentos e os metadados são armazenados e indexados pelo sistema de indexação Lucene. O componente de recuperação de dados pode recuperar tais informações posteriormente e o componente de agregação de dados combina as tuplas de sentimentos com os metadados em um formato que pode ser consumido pela visualização.

A análise de dados é dividida em dois modelos: modelo de índice de dados e modelo de visualização de dados. O modelo de índice define o formato dos dados para serem armazenados no bando de dados, e consiste em um conjunto de documentos. Entre os dados estruturados alguns são nativos dos dados brutos e outros derivados automaticamente do conjunto de dados utilizado. O modelo de visualização define o formato dos dados para visualização de sentimentos, e consiste em três tipos de dados principais, dados de tendência, níveis de detalhes dos dados e fragmento de dados.

A análise de sentimentos neste sistema é baseada em dicionário, e utiliza três domínios específicos. Os dicionários são utilizados para quatro propósitos: definir as características que são frequentemente mencionadas e discutidas, agrupar essas características em aspectos, extrair palavras de opinião e, por fim, extrair palavras adversativas, pois seu uso pode mudar a polaridade das palavras de opinião.

O funcionamento da análise de sentimentos dá-se da seguinte maneira: no início o documento é dividido em dois segmentos e cada um é *“tokenizado”*, então cada *token* é procurado no dicionário de entidades para seus aspectos e características associados. Caso não seja encontrado ou não seja encontrada alguma palavra de opinião próxima, o *token* é descartado. Caso contrário será utilizado o dicionário de oposição para determinar a polaridade das palavras de opinião. Depois

que todos os segmentos são processados as tuplas de sentimentos são armazenadas como vetores em cada documento.

O componente de visualização tem como resultado final as telas que apresentam as informações ao usuário. As informações são apresentadas de diversas maneiras através de múltiplas visualizações. As imagens a seguir mostram alguns dos resultados obtidos pelo sistema em dois casos de estudo. O primeiro dele é no domínio de viagens, foram coletadas 11.032 análises sobre 173 hotéis em Hong Kong através do site TripAdvisor.com. O segundo contexto é relacionado a 665.360 *tweets* do *Twitter*<sup>2</sup> relacionados ao tópico *National Football League* (NFL).

**Figura 19:** Múltiplas visualizações – Hotéis de Hong Kong



Fonte: Duan et al. (2012)

<sup>2</sup> <https://twitter.com/>

Observando a Figura 19 é possível observar algumas formas de visualização de dados que são utilizadas, além das opções de interação que são disponibilizadas ao usuário para que ele possa filtrar, selecionar e destacar os dados de forma que possa obter maior número de informações possíveis.

Neste projeto foi utilizado um processo de visualizações de múltiplas coordenadas. Na visualização principal são apresentadas as tendências dos sentimentos analisados, na parte inferior há uma *dashboard* com gráficos de torta que apresentam a distribuição de classificação dos clientes e o tipo de viagem nos comentários. A outra parte de visualização apresenta um fragmento das sentenças analisadas, sendo possível vê-las por completo ao clicar sobre cada uma.

O uso de interação é bastante grande neste trabalho, na visualização das tendências é possível escolher a polaridade, a quantidade de itens selecionados, marcar datas, entre outras opções, na *dashboard* há opção de selecionar o tipo de gráfico (barra ou torta), tipo de viagem e classificação do cliente, na área de visualização dos fragmentos é possível ordenar as sentenças por total, positivos e negativos.

Os dois trabalhos apresentados nesta seção utilizaram técnicas de VI diferentes, mas ambos são importantes para o estudo e desenvolvimento deste trabalho. No trabalho elaborado por Gamon et al (2005) há uma semelhança em relação a apresentação das sentenças, algo que será utilizado para este desenvolvimento, por ser importante para o grupo de pesquisa, assim como a utilização de múltiplas visualizações e interação que o trabalho elaborado por Duan et al. (2012) construiu. Com o estudo destes trabalhos é possível extrair boas ideias para serem utilizadas neste projeto.

## 3 MATERIAIS E MÉTODOS OU METODOLOGIA

### 3.1 População e Amostra

Para este trabalho, considerou-se como população cem destinos (cidades) e os aspectos (características) extraídos de um milhão quatrocentos e quinze mil, quatrocentos e setenta e seis avaliações do site TripAdvisor.

As avaliações foram analisadas pelo módulo de análise de sentimentos da ferramenta SentimentALL, e como resultado desse processo foi gerada uma base de dados com os aspectos avaliados como positivos e negativos.

### 3.2 Materiais

A primeira etapa trabalhada foi de estudo para a elaboração do referencial teórico, foram utilizadas como fontes bibliográficas artigos, dissertações, manuais técnicos e páginas da web.

Para o desenvolvimento do módulo de visualização de informação foi utilizado o padrão cliente-servidor, sendo que no lado do servidor foi utilizada a linguagem de programação PHP, que é uma linguagem de *script* de código aberto, especialmente voltado para o desenvolvimento web e que pode ser embutida dentro do código HTML. Seu foco é na programação no lado do servidor. A linguagem PHP é de extrema simplicidade, facilitando seu uso para os iniciantes, mas é também muito consistente para programadores profissionais, pois disponibiliza inúmeras funções e flexibilidade em relação ao uso de sistemas operacionais e banco de dados. (PHP.NET, 2016 online).

O framework utilizado no servidor foi o Silex, que é um micro framework PHP baseado em componentes do Symfony2 e Pimple e inspirado pelo Sinatra, tem uma estrutura minimalista que deixa o programador com liberdade, fornece um conjunto reduzido de serviços para o desenvolvimento de um aplicativo web e utiliza conceito de REST. (SENSIOLABS.ORG, 2016 online)

No lado do cliente foi utilizado o framework AngularJs, que “permite estender o vocabulário HTML para a sua aplicação. O ambiente resultante é extraordinariamente expressivo, legível e rápido para se desenvolver.” (GOOGLE, 2016 online). Para a criação dos gráficos e utilização do mapa, foram utilizadas APIs do Google, Chart e

Maps respectivamente. Como forma de armazenamento dos dados foi utilizado o Sistema de Gerenciamento de Banco de Dados (SGBD) “SQL Server 2014” (MICROSOFT, 2016).

### 3.3 Procedimentos

Para o desenvolvimento deste projeto foram realizadas seis etapas (Figura 20):

**Figura 20: Metodologia**



Os dados utilizados neste trabalho são conjuntos de aspectos que, neste contexto, são características de destinos turísticos. Os aspectos foram extraídos de comentários de usuários do site TripAdvisor, a partir da ferramenta SentimentALL de Brito *et al* (2015) e que foram avaliados como positivos ou negativos pelo módulo de análise de sentimentos desta ferramenta.

A primeira etapa do processo foi a realização de reuniões com os especialistas do domínio, com o objetivo de compreender o contexto e suas necessidades e demais questões pertinentes ao desenvolvimento deste trabalho. Nesta etapa foram respondidas, para cada uma das visualizações desejadas, as questões apresentadas por Mazza (2004). Tais questões foram relevantes a fim de estabelecer quais informações deveriam ser apresentadas e quais técnicas de VI seriam utilizadas. As

reuniões com os especialistas do domínio aconteceram durante todo o desenvolvimento do projeto, não sendo limitadas à fase inicial do desenvolvimento.

A segunda etapa correspondeu ao desenvolvimento do modelo de dados. Dois modelos foram necessários a princípio, um modelo para os dados que foram armazenados no banco de dados e outro modelo de dados para as ferramentas de VI. Para ambos os casos foi utilizada a estrutura de dados em *arrays*.

A terceira etapa consistiu em desenvolver o banco de dados. Nesta etapa foram criadas as tabelas e relacionamentos necessários para que os dados pudessem ser armazenados corretamente, e pudessem ser recuperados de forma eficiente quando necessário.

A quarta etapa foi a implementação dos recursos de VI com interatividade. Para tal implementação foi feito o mapeamento dos dados: os dados foram analisados e transformados de tabelas relacionadas em estruturas gráficas. Em seguida foi realizada a seleção dos dados, as cem cidades e seus vinte aspectos mais relevantes. O próximo passo foi a apresentação, nesta etapa as informações foram dispostas na tela de maneira que ficassem de fácil entendimento pelo usuário. Nesta fase do desenvolvimento foram aplicadas as técnicas de interatividade, conforme apresentadas na seção 2.1

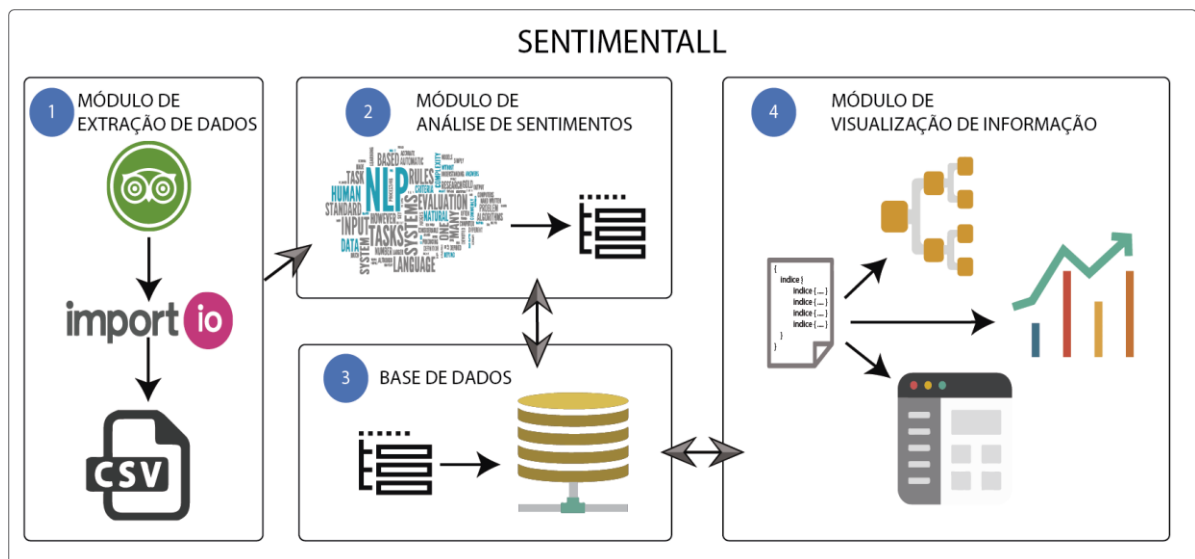
Na quinta etapa foram construídas as estruturas para armazenamento e apresentação do dicionário de termos. E na sexta etapa foi realizada a implementação do construtor de taxonomia, que permite que sejam criadas e editadas taxonomias de forma interativa e prática.

## 4 RESULTADOS E DISCUSSÃO

### 4.1 Arquitetura

A arquitetura apresentada na Figura 21 fornece uma visão geral de toda a ferramenta SentimentALL. Com isso, evidenciam-se as relações entre os processos de aquisição dos dados, de análise de sentimento e de visualização da informação.

**Figura 21:** Arquitetura ferramenta SentimentALL



Conforme apresentado na Figura 21, a arquitetura da ferramenta SentimentALL foi dividida em 3 módulos, sendo que no módulo de extração de dados (1) foi utilizada a ferramenta ImportIO para capturar as informações do site *TripAdvisor*<sup>3</sup>, gerando um arquivo no formato csv. Os arquivos gerados no primeiro módulo foram processados por uma ferramenta de análise de sentimentos no módulo 2; posteriormente, os resultados da análise foram armazenados em uma base de dados (3), que foi estruturada de forma que os dados pudessem ser relacionados entre si para facilitar futuras consultas.

A partir da base de dados, o módulo de visualização de informação (4) acessa os dados, depois aplica as equivalências do dicionário de termos (essas equivalências são definidas pelo usuário) sobre os dados retornados da base e então apresenta as

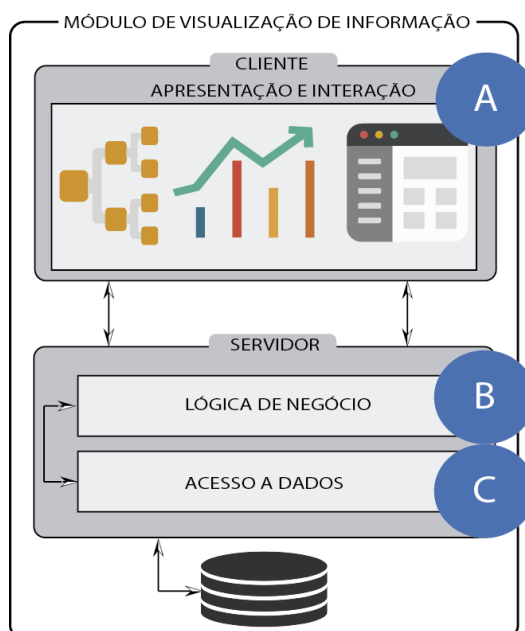
<sup>3</sup> <http://www.tripadvisor.com.br>



informações através das técnicas de VI aplicadas, ainda no módulo de VI também é disponibilizada a ferramenta para criar taxonomias.

O módulo de visualização de informação é dividido em duas partes principais, que são o servidor e o cliente. No servidor são realizados o acesso a base de dados e a lógica de negócio, no cliente são implementados os componentes visuais e de interação com o usuário. A seguir será apresentada a arquitetura deste módulo de forma mais detalhada.

**Figura 22:** Arquitetura Módulo de Visualização de Informação



A arquitetura do módulo de visualização de informação apresentada na Figura 22 utiliza o conceito de API *REST*, em que são separados cliente e servidor, e eles se comunicam através de métodos do *HTTP*.

A camada inferior na Figura 22 – C representa a interação da aplicação com a base de dados responsável pelas consultas e inserções, de acordo com as funções da camada de lógica de negócio (Figura 22-B). A partir deste ponto, a camada de lógica de negócio faz o tratamento e mapeamento dos dados, utilizando funções que verificam no dicionário de termos quais dados foram agrupados e devem ser enviados a interface como um único item, então transforma-os em *arrays* de objetos, enviando-os para a camada de apresentação (Figura 22-A).

Quando a camada de lógica de negócios recebe os dados da camada de apresentação, transforma-os para que possam ser armazenados na base de dados utilizando funções recursivas para analisar cada nível de profundidade do *array* recebido.

Já a camada de apresentação recebe os dados estruturados e os aplica nas ferramentas de renderização para gerar os gráficos ou as demais formas de exibição. Quando há inserção de dados pela camada de apresentação, eles são enviados para a camada de lógica de negócio e lá são novamente processados.

## 4.2 Desenvolvimento do Mecanismo de Visualização

As visualizações criadas neste trabalho seguiram uma adaptação do mecanismo de VI apresentado na seção 2.1.3. São quatro etapas relacionadas a partir de três ações que preparam os dados para o próximo passo. A Figura 23 ilustra a adaptação criada para este projeto. As principais mudanças se dão no momento da transformação, em que é aplicado o dicionário de termos, e no mapeamento, que utiliza o modelo de dados estabelecido no projeto.



Fonte: Adaptação do modelo de Card, Mackinlay e Shneiderman (1999)

Para iniciar o desenvolvimento das visualizações foi implementado um dicionário de termos, que será apresentado a seguir.

### 4.2.1 Dicionário de termos

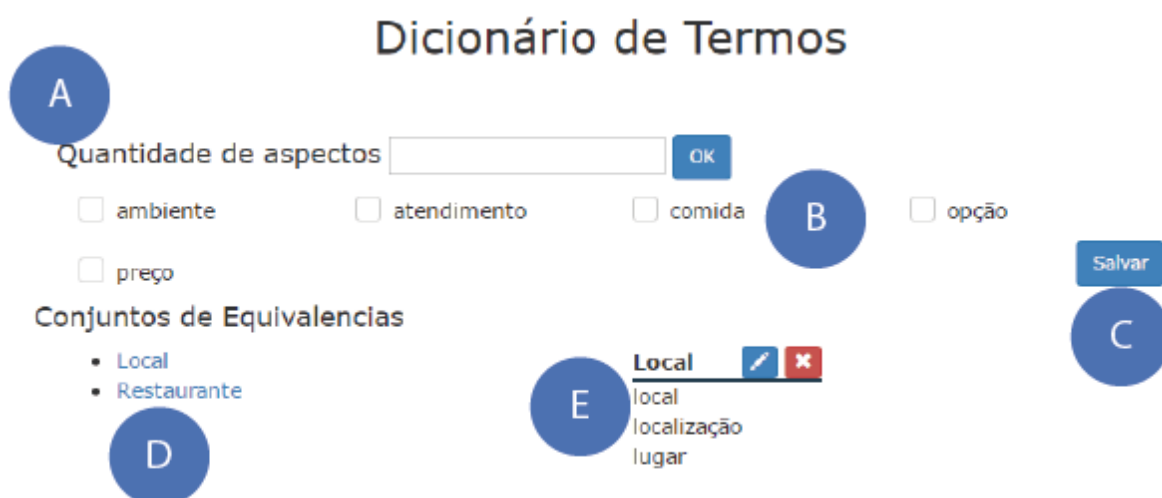
O dicionário de termos tem como objetivo criar uma maneira de selecionar a quantidade de aspectos com que se deseja trabalhar em todas as visualizações, e

principalmente eliminar a redundância, agrupando os aspectos que têm significado comum, de forma que possa ser atribuído um nome para este conjunto. Assim, sempre que esses aspectos forem utilizados, eles serão apresentados como um único item e com o nome dado ao conjunto.

Para o desenvolvimento desta funcionalidade foram criadas duas tabelas no banco de dados (a estrutura do banco de dados é apresentada na seção de apêndices), uma para armazenar os aspectos com maior incidência nas avaliações, de acordo com a quantidade escolhida pelo usuário, e outra tabela para armazenar os conjuntos de aspectos relacionados. Além das tabelas na base de dados foram criadas funções para consultar os dados da base e armazenar os dados inseridos pelo usuário, e também foi criada a estrutura de interface utilizada para apresentar e editar os conjuntos de aspectos.

A seguir são apresentadas as interfaces desenvolvidas para o dicionário de termos.

**Figura 24:** Dicionário de termos



Na Figura 24 é apresentada a tela do dicionário de termos.

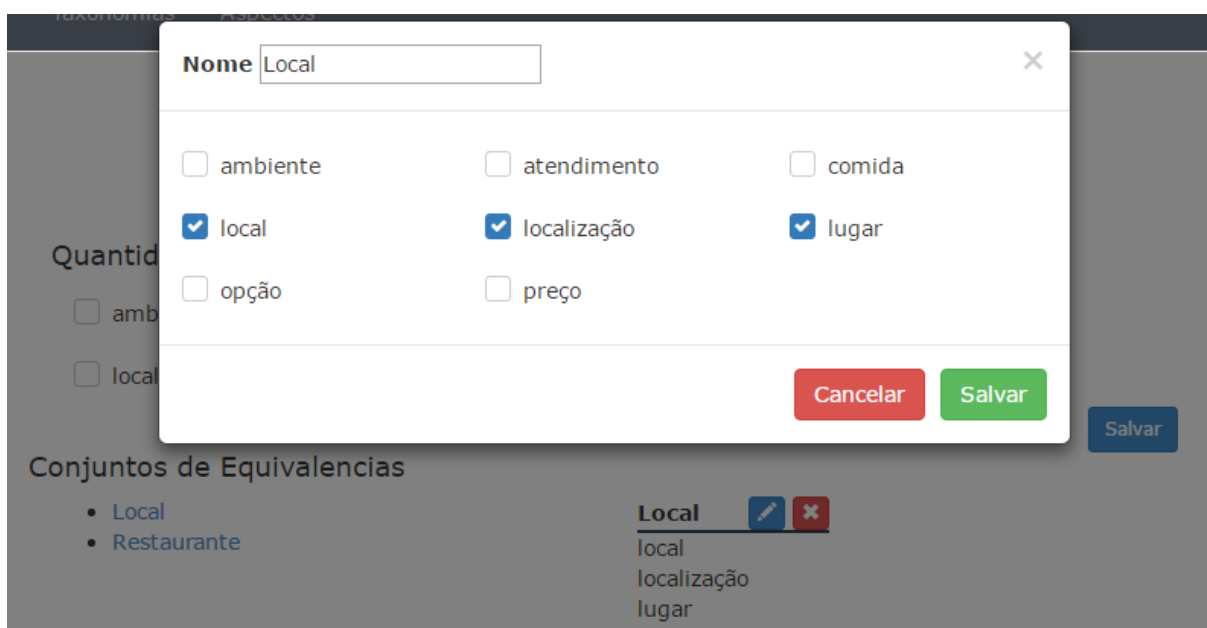
Primeiramente deve ser feita a seleção da quantidade de aspectos desejados (Figura 24 - A), a partir de então é feita uma consulta ao banco de dados através da camada de acesso a dados, que retorna a lista dos aspectos e apresenta-os (Figura 24 - B), de acordo com a quantidade selecionada.

Para criar um novo conjunto, o usuário deve selecionar os itens que deseja e então clicar no botão 'Salvar' (Figura 24 - C), em seguida inserir um nome para aquele

conjunto de aspectos. Ao finalizar, clicando em salvar, será apresentado na lista dos conjuntos de equivalências (Figura 24 - D) este novo grupo.

Para o usuário consultar e editar um conjunto, deve clicar sobre o item que deseja, então será apresentado ao lado direito (Figura 24 - E) o conjunto com seu rótulo destacado em negrito e os aspectos relacionados logo abaixo. Há, também, as opções de edição e exclusão.

**Figura 25:** Edição de um item do dicionário de termos



A Figura 25 apresenta a tela de edição de um conjunto de equivalência. Para editar, o usuário clica sobre o ícone de edição e uma nova tela será apresentada com a relação de todos os aspectos não agrupados, dentre os listados na tela anterior, e aqueles que já fazem parte deste conjunto, assim o usuário pode selecionar ou remover aqueles itens que devem ou não fazer parte do conjunto.

Após o dicionário de termos ter sido criado, o mecanismo de visualização pode ser aplicado de forma mais eficiente, porque as redundâncias foram eliminadas. Desta forma os dados que são retornados de uma consulta à base de dados (camada de acesso a dados da arquitetura) (Figura 23 - 1), são processados, utilizando o dicionário de termos, e transformados em uma tabela de dados (Figura 23 - 2).

Quando os aspectos selecionados para determinada visualização são carregados é verificado se eles pertencem a algum grupo de aspectos, e, em caso positivo, seus valores (quantidade de vezes que aparecem nas sentenças) são somados aos demais aspectos que compõem o grupo. Quando todos os aspectos retornados tiverem sido verificados, o grupo terá um único valor, e é este valor,

juntamente com o rótulo do grupo que irá ser enviado para a interface. Por exemplo: A consulta à base de dados, na camada de acesso a dados, apresentada na arquitetura, retorna os seguintes dados:

**Tabela 1:** Relação de aspectos

Aspecto	Total
atendimento	200
local	50
localização	30
localidade	10
comida	100

E no dicionário de termos há o seguinte conjunto armazenado:

**Tabela 2:** Dicionário de termo – conjunto de aspectos

Rótulo	Aspectos
local	Local, localidade, localização

Assim, após o processo de transformação dos dados, o resultado será o seguinte:

**Tabela 3:** Tabela de dados resultante

Aspecto	Total
atendimento	200
<b>Local</b>	<b>90</b>
comida	100

O item que representa um grupo de aspectos do dicionário de termos está destacado em negrito na Tabela 3.

A partir da tabela de dados é feito o mapeamento dos dados para a estrutura visual. Neste trabalho foi estabelecida uma estrutura de dados única para todas as

ações da ferramenta O padrão estabelecido trabalha com um *array* de objetos, com um par de chaves, por exemplo, a tabela de dados resultante é mapeada da seguinte forma:

```
[{aspecto: atendimento, total:200}, {aspecto: local, total: 90}, {aspecto: comida, total: 100}]
```

Neste exemplo, o par de chaves é {aspecto, total}. Este padrão foi utilizado pela facilidade em acessar cada item da lista dentro da estrutura dos gráficos que foram desenvolvidos.

Após este processamento, a estrutura visual é finalizada com a aplicação dos dados nos gráficos, tabelas e demais componentes de visualização. Ao finalizar as estruturas visuais a visualização fica disponível ao usuário.

A seguir serão apresentadas as visualizações que foram desenvolvidas neste trabalho.

#### **4.2.2 Visualizações**

Na Figura 26 é apresentado um mapa com os cem destinos marcados sobre a sua localização geográfica. Quando o usuário clica sobre um marcador (Figura 26 - A) é apresentado um gráfico de barras com os aspectos mais relevantes para aquele destino ( Figura 27).

Figura 26: Tela – Mapa destinos

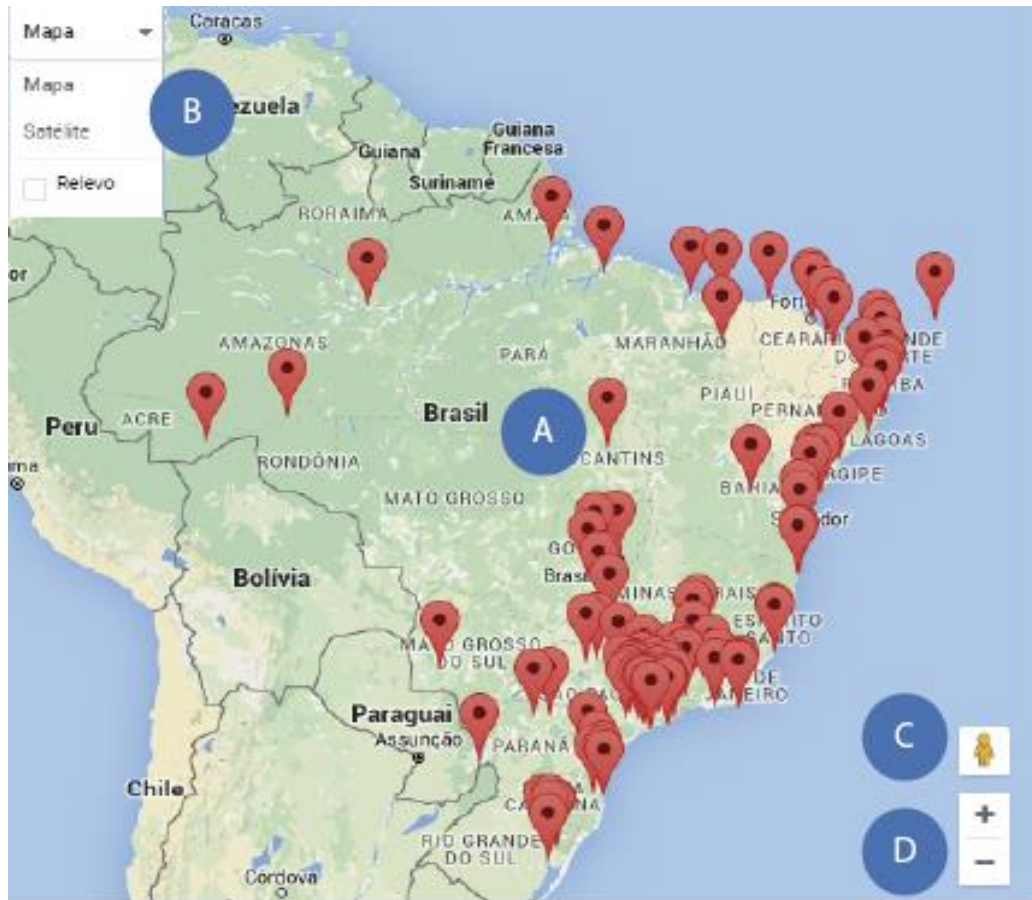
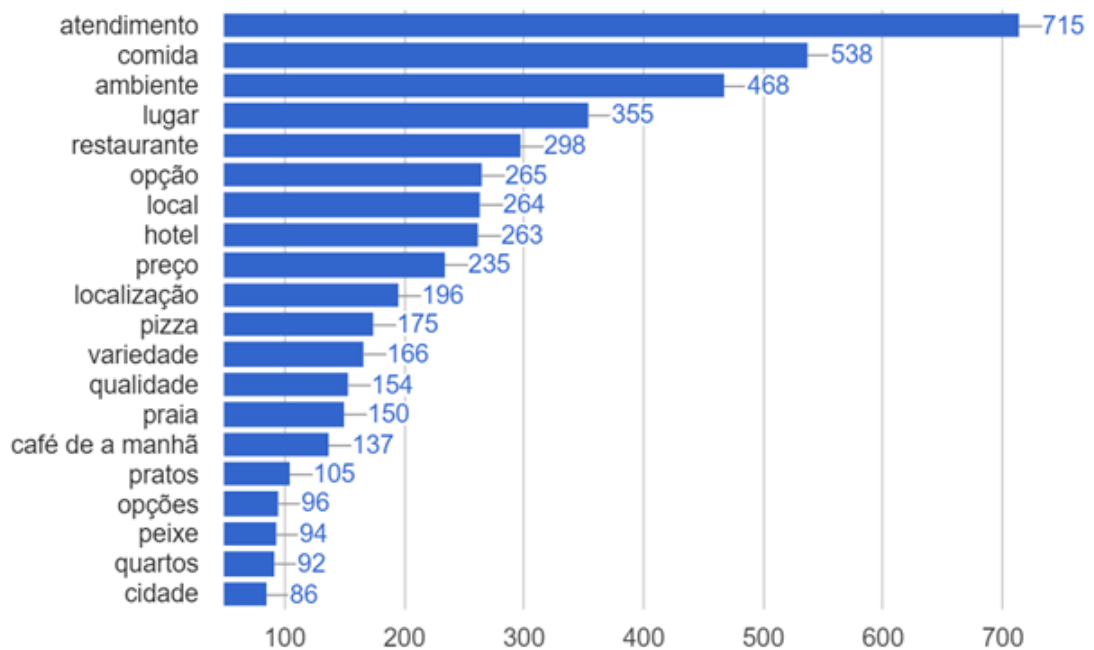


Figura 27: Aspectos mais relevantes por destino

Palmas

Aspectos mais relevantes



Algumas funções de interação são disponíveis para o usuário, que pode alterar a forma de visualização da superfície do mapa (Figura 26 - B), por exemplo, visualizar como uma imagem de satélite. Além disso, o usuário pode utilizar a função de zoom (Figura 26 - C) para afastar ou aproximar a imagem de um determinado ponto no mapa, como também pode utilizar a função do *Google Street View* (Figura 26 - D), que mostra imagens do local onde o cursor é posicionado, podendo navegar entre as ruas da localidade.

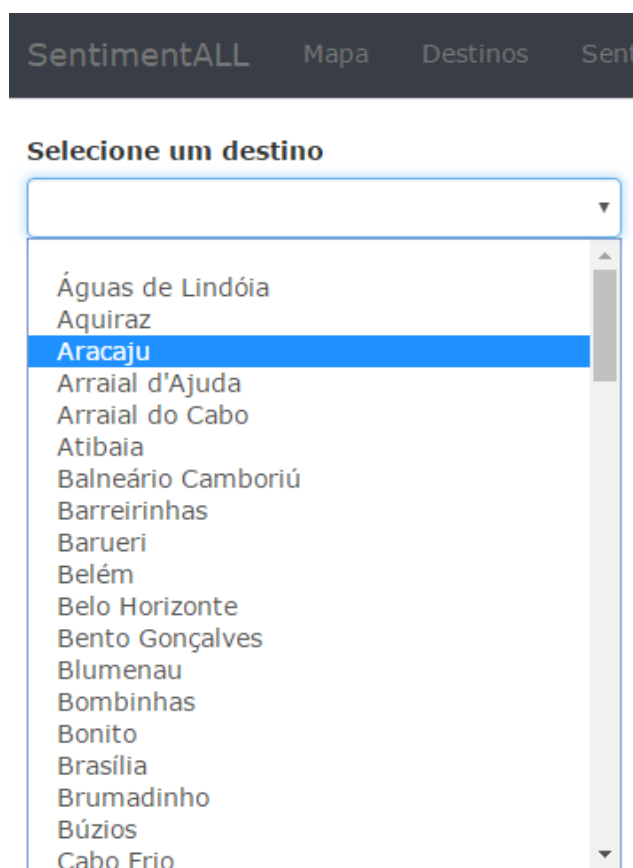
As técnicas de VI utilizadas para esta visualização foram escolhidas com base nos resultados desejados pelo especialista do domínio, na estrutura e dimensão dos dados e no tipo de interação, porque cada técnica de VI se adapta melhor com determinada estrutura e dimensão dos dados e também com o tipo de interação. Para o mapa, a estrutura dos dados é linear e são utilizadas duas dimensões, o tipo de interação é transformável, pois o usuário pode interagir utilizando zoom, filtro e seleção.

Foi escolhido a representação em mapa porque permite ao usuário identificar a posição geográfica de cada destino, as opções de interação disponíveis neste tipo de visualização e também por permitir que o usuário possa analisar, por exemplo, qual região teve mais destinos avaliados.

Para o gráfico de barras, apenas foi modificado o tipo de interação, que neste caso é estática. Foi escolhido este tipo de representação porque permite relacionar um aspecto com seu quantitativo e apresentar a proporção (tamanho das barras) entre os aspectos.

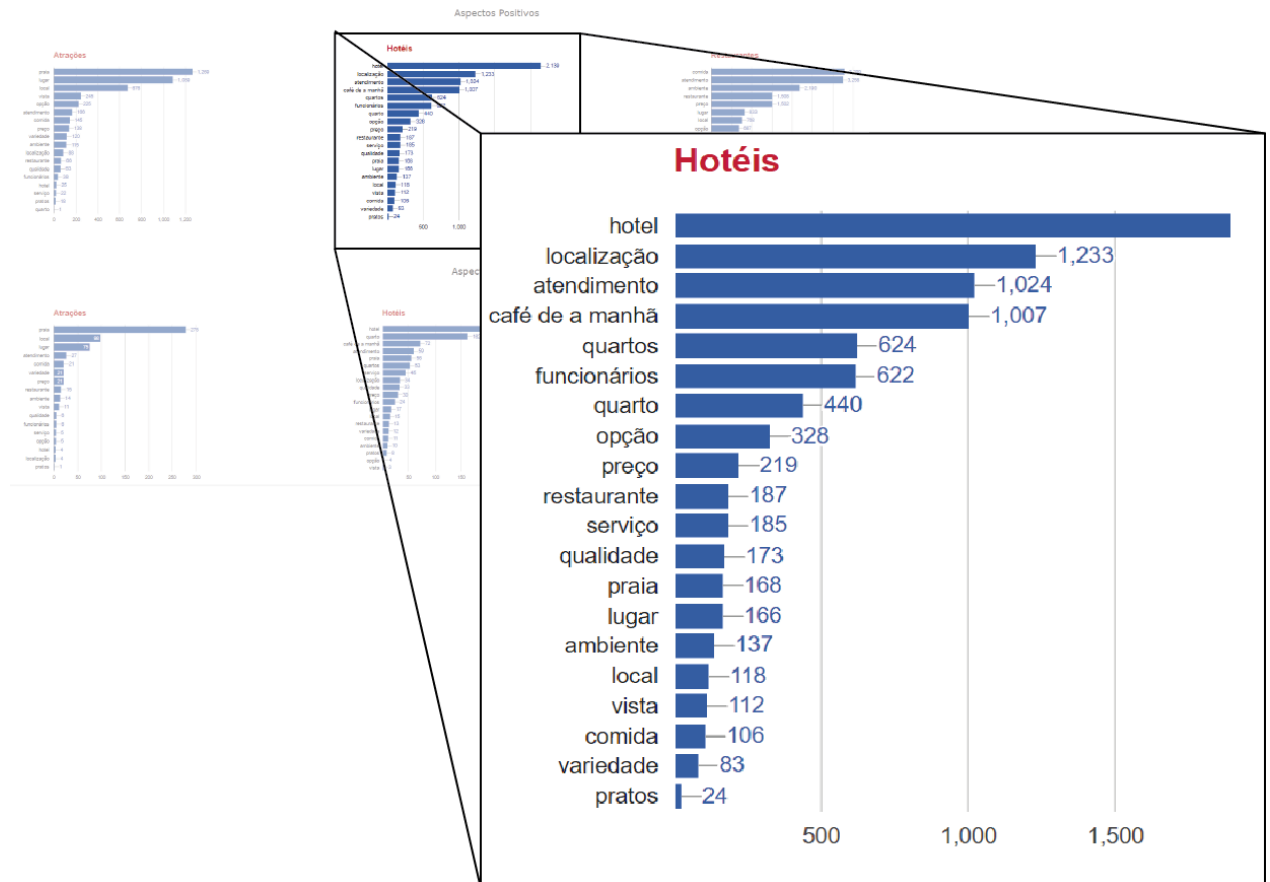
Outra visualização desenvolvida tem como objetivo apresentar os aspectos mais relevantes de um destino selecionado, separando-os por polaridade positiva e negativa, sendo que dentro de cada uma das polaridades há uma organização das informações por atrações, hotéis e restaurantes.



**Figura 28:** Tela de seleção de destino

A Figura 28 apresenta a tela para selecionar o destino que se deseja consultar dentre os cem destinos listados pela ferramenta. Quando o usuário selecionar o destino que deseja, serão apresentados os gráficos com os aspectos relacionados a cidade em questão. A Figura 29 apresenta o resultado da seleção.

**Figura 29:** Tela - Aspectos por destino



A representação gráfica utilizada na Figura 29 foi a de gráfico de barras e foi escolhida pelo fato dos dados serem lineares, de duas dimensões, e a necessidade de relacionar um aspecto com o seu valor, indicando de forma visual a proporção (tamanho das barras) entre cada aspecto, nesta visualização a interação acontece quando é feita a seleção do destino desejado.

A Figura 30 apresenta a visualização das sentenças relacionadas a um aspecto e a uma palavra opinativa, que é o termo dentro da sentença que expressa opinião e que é utilizada na ferramenta de análise de sentimentos para definir a polaridade positiva ou negativa.

**Figura 30:** Tela - Sentenças por aspecto

Selecione um aspecto

Local A Enviar

Pesquisar C Q

Sentença	Aspecto	Palavra Opinativa
local muito bom, porções boas variadas, chope gelado repostado a toda hora.	local	bom
experiência unica em Sampa, esse restaurante é o melhor da cidade é um local perfeito para ocasioes especiais.	local	perfeito
local agradável e a vista é linda da cidade de são Paulo, pois, está localizado no 41º andar na cobertura do edificio Itália, creio que seja um dos mais altos da cidade, permitindo uma vista de 360 graus.	local	agradável
excelente local e refeição.	local	excelente
alias por ganhar um premio "framboesa" entre uns dos piores restaurantes franceses, vale ressaltar que o local realmente é um charme!	local	charme
vá com paciência, mas o local é muito bom.	local	bom
o local é muito agradável para um jantar com os amigos.	local	agradável
local muito gostoso, comida muito boa e se for comparado os preços com os demais bares e restaurantes em Campos não é nada de absurdo.	local	gostoso
comida boa, bom atendimento, local privilegiado.	local	bom
muito bom! atendimento muito bom, cerveja muito bem servida, local agradável e a porção de salsicha excelente!	local	agradável

**1000 registro(s)**

«
<
1
2
3
4
5
...
100
>
»

Na visualização apresentada na Figura 30, quando o usuário seleciona um aspecto (Figura 30 - A), são apresentadas mil sentenças em uma tabela (Figura 30 - B) com paginação. Para o usuário não precisar ficar navegando entre as páginas da tabela para encontrar alguma frase, ele pode utilizar o campo de pesquisa (Figura 30 - C), que fará a consulta nos dados da tabela e apresentará os resultados na própria tabela.

A quantidade de mil sentenças foi definida por apresentar uma quantidade suficiente de sentenças para análise do usuário, sem perder a relevância. Quando utilizada uma quantidade maior, foi observado que há muitas sentenças semelhantes, assim não trazem novas informações para o usuário.

Foi escolhida a apresentação em forma de tabela pelo conjunto de dados que seria apresentado, pois era necessário criar uma estrutura que relacionasse a sentença, o aspecto e a palavra opinativa. A tabela, nesse contexto, permitiu uma visualização mais direta das relações entre esses elementos, assim o pesquisador tem acesso rápido às sentenças que tem determinados aspectos, bem como as palavras opinativas usadas pelo autor do comentário. Similar às visualizações apresentadas anteriormente, a estrutura dos dados é linear, são duas dimensões e o tipo de interação é estático.

Na sequência será apresentado o desenvolvimento do construtor de taxonomias, mostrando as interfaces geradas, a necessidade de criar novas tabelas na base de dados e um exemplo para ajudar no entendimento da ferramenta.

### 4.3 Desenvolvimento do Construtor de Taxonomias

O objetivo do construtor de taxonomia é fornecer um mecanismo para o usuário criar taxonomias de produtos turísticos de forma interativa. Isso permitirá a organização dos aspectos extraídos de um destino turístico a partir de um determinado contexto. Com o mecanismo, é possível criar nós com seus respectivos rótulos (são os itens da taxonomia, e cada um desses itens recebe um nome), adicionar os aspectos que têm relação com o objeto principal da taxonomia (nó raiz – primeiro item) em um nó folha (último item da hierarquia), poder mover os nós de forma a alterar a ordem da taxonomia, e poder consultar e editar as taxonomias criadas.

Por exemplo, uma taxonomia para Produto Turístico:

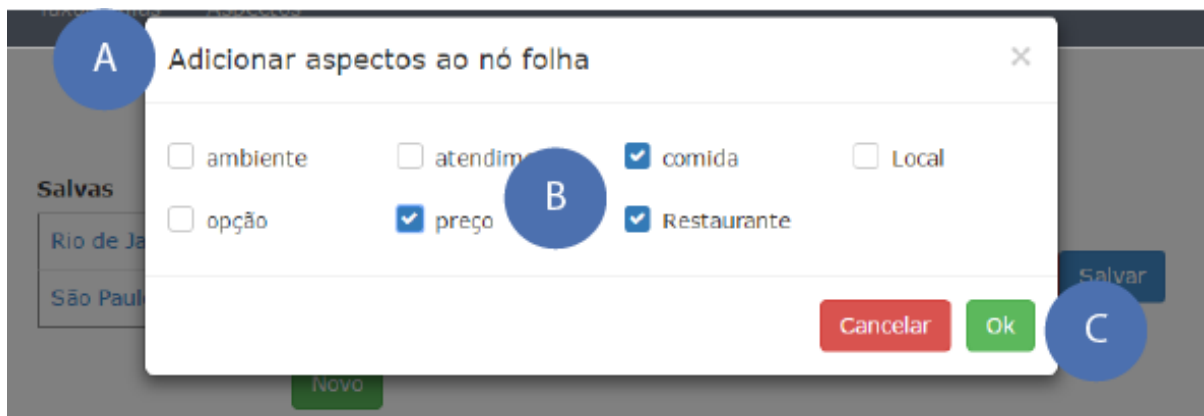
- **Produto Turístico – Nó raiz**
  - Atrativos
    - Lugar
      - Naturais; Usos e costumes; infraestrutura – **Nó folha**

Para que fosse possível criar as taxonomias, foi necessário adicionar duas tabelas na base de dados, sendo uma para armazenar os nós e outra para armazenar os aspectos relacionados no nó folha. Além das tabelas foram criadas funções para acessar e salvar os dados, e também foi criada a estrutura de interface para apresentar e criar novas taxonomias.

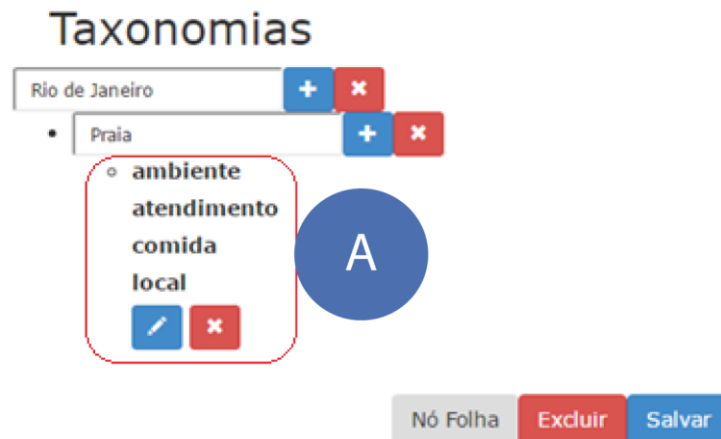
A seguir são apresentadas as interfaces desenvolvidas para o construtor de taxonomias.

**Figura 31:** Construtor de taxonomias

Na Figura 31 é apresentada a tela de cadastro e edição de uma taxonomia, além de apresentar a relação das taxonomias salvas na base de dados (Figura 31-A). Ao clicar sobre um dos itens salvos, é apresentada a taxonomia no lado direito da figura (Figura 31-B), que pode ser editada, adicionando ou removendo nós, ou excluída. O usuário poderá também clicar na opção 'Novo' (Figura 31-C) e criar uma nova taxonomia.

**Figura 32:** Tela para adicionar aspectos ao nó folha

De acordo com a Figura 32, para adicionar o nó folha à taxonomia, o usuário clica no botão 'Nó folha' (Figura 31-D) e então uma nova tela (Figura 32-A) é apresentada, listando os aspectos relacionados na base de dados (Figura 32-B), então o usuário poderá selecionar aqueles que são pertinentes ao contexto e clicar em 'ok' (Figura 32-C), feito isso será apresentado um nó com os aspectos listados, e o usuário irá movê-lo para a posição desejada dentro da taxonomia (Figura 33-A).

**Figura 33: Nó folha adicionado à taxonomia**

Este capítulo abordou os resultados obtidos durante o desenvolvimento do módulo de visualização de informação para a ferramenta SentimentALL. O desenvolvimento incluiu a apresentação da arquitetura da ferramenta, a arquitetura do módulo desenvolvido neste trabalho, o mecanismo de VI com o dicionário de termos aplicado nas visualizações e o desenvolvimento do construtor de taxonomias. Foram apresentadas também todas as visualizações geradas neste módulo.

## 5 CONCLUSÃO

Este trabalho teve como resultado a construção de um módulo de visualização de informações no âmbito do turismo, utilizando as técnicas de VI. Através das visualizações geradas os usuários podem ter acesso a novas informações, que até então não eram possíveis pela falta deste módulo. Com o uso das técnicas de VI foi possível proporcionar apresentações interativas que permitem que as informações sejam apresentadas de acordo com as necessidades do usuário.

A partir das visualizações geradas foi possível analisar os aspectos relacionando-os com cidades, com suas polaridades (positivo e negativo), por tipo de destino (hotel, restaurante e atração) e com outros aspectos. Isso facilita, ao usuário, a identificação, de acordo com seu interesse, de padrões e características obtidas a partir dos dados analisados.

Com o desenvolvimento do dicionário de termos, o usuário pode criar grupos de aspectos que são equivalentes, tornando assim os resultados obtidos através da análise dos aspectos nos gráficos mais avançada, uma vez que a redundância é eliminada. Para projetos futuros que envolvam a equivalência de palavras, pode-se tornar esse método mais eficaz desenvolvendo um módulo que realize buscas automáticas através do uso dos radicais das palavras (*stemming*), evitando que o usuário tenha que fazer manualmente o processo de equivalência entre os termos apresentados.

Com o construtor de taxonomia é possível criar taxonomias diferentes para cada destino e vincular quais os aspectos que tem relação com o destino em questão. Desta forma o usuário pode comparar a taxonomia criada no sistema com a disponível no site *TripAdvisor*, e a partir deste ponto chegar às suas conclusões. Assim como no dicionário de termos, ter um conjunto de radicais das palavras da língua portuguesa disponíveis tornará o processo mais eficiente, pois, assim, a forma como os aspectos são vinculados a taxonomia será automática, porque o usuário poderá apenas informar os termos chaves e o sistema retornará todos os aspectos que tem o radical do termo.

Neste trabalho foi utilizada uma estrutura de dados para as visualizações que tem como padrão um par de chaves {aspecto, total}, porém, existe uma forma que pode ser mais eficiente, que consiste em utilizar o aspecto em si como a chave e atribuir a ele o valor correspondente a chave total.

## REFERÊNCIAS

ADOBE FLEX, **Advanced Data Visualization Developer Guide**, 2008.

BEDERSON, B. B.; SHNEIDERMAN, B. **The craft of information visualization: readings and reflections**, San Francisco, CA: Morgan Kaufmann Publishers, 2003.

BRITO, P. F.; CHRISTHIE, W.; SOUZA, J. G.; SILVA, E.M. **Projeto SentimentALL**. Fábrica de Software: CEULP/ULBRA, 2015.

BURKHARD, R. A. **Knowledge visualization**: The use of complementary visual representations for the transfer of knowledge. A model, a framework, and four new approaches. Eidgenössische Technische Hochschule ETH Zurich, 2005.

BURKHARD, R. A. Learning from architects: the difference between knowledge visualization and information visualization. In **IV '04: Proceedings of the Information Visualisation, Eighth International Conference**, p. 519–524, Washington, DC, USA: IEEE Computer Society, 2004.

CARD, S. K.; MACKINLAY, J. D. & SHNEIDERMAN, B. **Readings in Information Visualization: using vision to think**. Morgan Kauffman, 1999

CARD, S; Information visualization, in **The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications**, A. Sears and J.A. Jacko, Lawrence Erlbaum Assoc Inc, 2007.

CHITTARO, L.. Visualizing Information on Mobile Devices. **Computer**, [s.l.], v. 39, n. 3, p.40-45, mar. 2006. Institute of Electrical & Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/mc.2006.109>.

DÄßLER, R.; PALM, H. **Virtuelle Informationsräume mit VRML: Informationen recherchieren und präsentieren in 3D**. Heidelberg: dpunkt-Verlag, 1998.

DUAN, D. et al. **VISA: A Visual Sentiment Analysis System**. Hangzhou, 2012.



FEW, S. **Multivariate Analysis Using Parallel Coordinates**, September 12, 2006.

\_\_\_\_\_. **Save the Pies for Dessert**. Ago, 2007.

\_\_\_\_\_. **Visual Business Intelligence: Why Do We Visualize Quantitative Data**. 2014. Disponível em: <<https://www.perceptualedge.com/blog/?p=1897>>. Acesso em: 03 dez. 2015.

FINK, A. **How to conduct surveys: A step-by-step guide**, Sage Publications, Inc, 2008.

FREITAS, C. M. D. S.; CHUBACHI, O. M.; LUZZARDI, P. R. G.; CAVAR, A. Introdução à Visualização de Informações. In: **RITA – Revista de Informática Teórica e Aplicada, Instituto de Informática UFRGS**, Porto Alegre, RS, v. VIII, n. 2, p. 143-158, outubro, 2001.

FRIENDLY, M. **Milestones in the history of thematic cartography, statistical graphics, and data visualization**, 2009. Disponível em: <<http://www.math.yorku.ca/SCS/Gallery/milestone/milestone.pdf>>.

GAMON, Michael et al. Pulse: Mining Customer Opinions from Free Text. **Lecture Notes In Computer Science**, [s.l.], p.121-132, 2005. Springer Science + Business Media. [http://dx.doi.org/10.1007/11552253\\_12](http://dx.doi.org/10.1007/11552253_12)

GEMIGNANI, Z. **Better Know a Visualization: Parallel Coordinates**. 2010. Disponível em: <<http://www.juiceanalytics.com/writing/writing/parallel-coordinates?rq=parallel-coordinates%2F>>

GOOGLE. **AngularJS**. Disponível em: < <https://angularjs.org/>>. Acesso em: 09 jun. 2016.

HEARST, M. **Information Visualization: Principles, Promise, and Pragmatics**. 2003. Disponível em: <<http://bailando.sims.berkeley.edu/talks/chi03->

tutorial.ppt&usg=AFQjCNFXTzmNHfbftRdc0x9KlrLfl8hIUw&sig2=wEOdfaNIkR0t0UBDGXmXeQ>. Acesso em: 11 dez. 2015.

HEBB, N. **What is a Flow Chart?** Disponível em: <<http://www.breezetre.com/articles/what-is-a-flow-chart.htm>>. Acesso em: 28 nov. 2015.

KELLER, T.; TERGAN, S.O. **Visualizing knowledge and information: An introduction.** In **Knowledge and Information Visualization**, pages 1–23, Berlin/Heidelberg, Germany: Springer, 2005.

KHAN, M.; Khan S. S. Article: Data and Information Visualization Methods, and Interactive Mechanisms: A Survey. In: **International Journal of Computer Applications** 34(1): 1-14, nov. 2011. Disponível em: <<http://research.ijcaonline.org/volume34/number1/pxc3875722.pdf>>

LEGLER, R.; EPPLER, M. J. Towards a periodic table of visualization methods for management. In **GVE 2007: Graphics and Visualization in Engineering**. Acta Press, 2007.

LIU, B. Sentiment analysis and subjectivity. In: **INDURKHYA, N; DAMERAU, F J. Handbook of Natural Language Processing**. 2. ed. Boca Raton, Fl: Chapman And Hall/crc, 2010. Cap. 26. p. 627-666.

LTIFI, H.; BEN AYED, H.M; LEPREUX, S.; ALIM, M.A. Survey of Information Visualization Techniques for Exploitation in KDD, in: **7th ACS/IEEE International Conference on Computer Systems and Applications**, A1CCSA-2009 (May 10-13), Rabat, Morocco, pp. 218-225, maio, 2009.

MARR, B. **Big Data: 20 Mind-Boggling Facts Everyone Must Read**. 2015. Disponível em: <<http://www.forbes.com/sites/bernardmarr/2015/09/30/big-data-20-mind-boggling-facts-everyone-must-read/>>. Acesso em: 30 nov. 2015.

MARTINS, Rosemary. **Fluxograma de Processo**. Disponível em: <<http://www.blogdaqualidade.com.br/fluxograma-de-processo/>>. Acesso em: 04 dez. 2015.

MAZZA, R. **Using Information Visualisation to Facilitate Instructors in Web-based Distance Learning**. 2004. 231 f. Dissertação (Ph.D in Communication Sciences) - Faculty of Communication Sciences University of Lugano. 2004

MCCANN, J. M. **Generation of Marketing Insights: Semantic Networks**. Disponível em: <<http://people.duke.edu/~mccann/mwb/15semnet.htm>>. Acesso em: 28 nov. 2015.

MEYER, Robert. Knowledge Visualization. In: **Media informatics advanced seminar on information visualization**, Munique, Alemanha, 2009.

MICROSOFT. **SQL Server**. Disponível em: <<http://www.microsoft.com/pt-br/server-cloud/products/sql-server/default.aspx>>. Acesso em: 09 jun. 2016.

MOISSA, Barbara et al. Uma ferramenta de Visualização da Informação para analisar o comportamento do aluno em um ambiente e-learning e sua trajetória de aprendizagem. In: **Revista Brasileira de Design da Informação**, São Paulo, v. 11, n. 3, p.337-351, 2014.

OLIVEIRA, W. C. C. **SENTIMENTALL: Ferramenta para análise de sentimentos em português**. 2015. Trabalho de Conclusão de Curso (Bacharelado em Sistemas de Informação). Centro Universitário Luterano de Palmas, Palmas, Tocantins, 2015. Disponível em: <<http://ulbra-to.br/bibliotecadigital/publico/home/documento/151>>. Acesso em: 15 jun. 2016

PHP.NET. **O que é o PHP?** Disponível em: <[http://php.net/manual/pt\\_BR/intro-what-is.php](http://php.net/manual/pt_BR/intro-what-is.php)>. Acesso em: 09 jun. 2016.

SENSIOLABS.ORG. **Silex, Micro-framework PHP**. Disponível em: <<http://silex.sensiolabs.org/>> Acesso em: 09 jun. 2016.

PROTOVIS. **Parallel Coordinates**. Disponível em: <<http://mbostock.github.io/protovis/ex/cars.html>>. Acesso em: 28 nov. 2015.

ROUSE, M. **Timeline (Internet timeline, history of the Internet)**. 2005. Disponível em: <<http://whatis.techtarget.com/definition/timeline-Internet-timeline-history-of-the-Internet>>. Acesso em: 28 nov. 2015.

Schumann, H., & Müller, W. **Informationsvisualisierung: Methoden und Perspektiven**. it - Information Technology, 46(3), 135-141. 2004

SHNEIDERMAN, B.; PLAISANT, C. **Treemaps for space-constrained visualization of hierarchies**, 2009.

TIBCO SPOTFIRE. **Help, What is a Bar Chart?**, 2012. Disponível em: <[http://stn.spotfire.com/spotfire\\_client\\_help/index.htm](http://stn.spotfire.com/spotfire_client_help/index.htm)>. Acesso em 30 de Novembro de 2015

TERGAN, S.-O; KELLER, T.; BURKARD, R. A. **Integrating knowledge and information: digital concept maps as a bridging technology**. *Information Visualization*, 5(3):167–174, 2006.

WIKIPÉDIA, **Big Data**, 2015. Disponível em: < [https://pt.wikipedia.org/wiki/Big\\_data](https://pt.wikipedia.org/wiki/Big_data) > Acesso em: 10 dez. 2015.

YI, J.; S., Kang, Y. A., Stasko, J., & Jacko, J. A. (2007). **Toward a Deeper Understanding of the Role of Interaction in Information Visualization**. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 13(6). Presented in InfoVis 2007, Sacramento, California, October 28 - November 1, 1224-1231.

ZUVA, K.; ZUVA, T. Evaluation of information retrieval systems. In: **International Journal Of Computer Science & Information Technology (ijcsit)**. p. 35-43. jun. 2012.

## APÊNDICES

## APÊNDICE A – Estrutura do banco de dados

