

**Visualização de Informações**

*Fernando Rosa Vaz      Cedric Luiz de Carvalho*

Technical Report - INF\_003/04 - Relatório Técnico  
October - 2004 - Outubro

The contents of this document are the sole responsibility of the authors.  
O conteúdo do presente documento é de única responsabilidade dos autores.

**Instituto de Informática**  
**Universidade Federal de Goiás**  
*www.inf.ufg.br*

# Visualização de Informações

Fernando Rosa Vaz <sup>\*</sup>  
frosa@inf.ufg.br

Cedric Luiz de Carvalho <sup>†</sup>  
cedric@inf.ufg.br

**Abstract.** *The Internet is a tool used by millions of people that stores a huge amount of information. The search for a specific information, desired by the user, in this environment is not an easy task. The main challenge for the Information Visualization area is to handle the difficulties of manipulating and presenting this large amount of data. This area of Computer Science is introduced in this paper. It is discussed some technics for improving the visualization of a set of data. Besides, it is presented a short discussion about good technics to visualize the result of a search in the Internet.*

**Keywords:** Information Visualization, Data in Internet, Data Presentation.

**Resumo.** *A Internet é uma ferramenta utilizada por milhões de pessoas, que guarda uma enormidade de informações. A busca por uma informação específica, de interesse do usuário, nesta grande quantidade de dados não é uma tarefa fácil. A dificuldade em se manipular estes dados e, principalmente, em apresentá-los de modo a transmitirem o maior número de informações é o desafio da Visualização de Informações. Esta área da computação é introduzida neste texto. São abordadas algumas das diversas técnicas empregadas para melhor visualização de um conjunto de informações. Além disto, é apresentada uma breve discussão a respeito da escolha da melhor técnica para visualizar o resultado de uma busca feita na Internet.*

**Palavras-Chave:** Visualização de informações, Dados da Internet, Apresentação de Dados.

## 1 Introdução

O avanço tecnológico das comunicações, das máquinas e equipamentos para manipulação de dados, assim como o crescente desenvolvimento da Internet, tem disponibilizado um volume cada vez maior de informações. Os usuários, de posse destes equipamentos, são capazes de obter um número muito grande de dados. Entretanto, nem todos os dados obtidos pelos usuários em uma pesquisa, são efetivamente úteis para ele. Este excesso de informações é o principal problema encontrado para a análise e compreensão dos resultados obtidos.

Um modo de facilitar a análise de dados, obtidos em pesquisas na Internet, e compreender seu significado é utilizar técnicas de Visualização de Informações. Utilizando-se estas

---

<sup>\*</sup>GEApIS - Laboratório de Internet e Sistemas Distribuídos

<sup>†</sup>GEApIS - Laboratório de Internet e Sistemas Distribuídos

técnicas é possível reunir milhares de dados em uma imagem e revelar padrões eventualmente ocultos neles. Estes métodos permitem ao usuário uma visão mais compreensível, abstraindo os detalhes, sem prejuízo do entendimento do real significado dos dados.

Reunindo aspectos da computação gráfica, interfaces homem-computador e mineração de dados, a Visualização de Informações permite a apresentação de dados de forma gráfica, de modo que o usuário possa utilizar sua percepção visual para melhor analisar e compreender as informações. Quando esses dados correspondem a medidas associadas a objetos físicos, fenômenos ou posições num domínio espacial, costuma-se referir a esse conjunto de técnicas como Visualização de Dados Científicos, enquanto a Visualização de Informações trata dados abstratos como relacionamentos ou informações inferidas a partir de dados mensurados. Esta distinção é, de certa forma, irrelevante no estudo das técnicas de visualização em si.

A Visualização de Informações é caracterizada pela necessidade do projetista de criar uma forma para transformar dados em uma representação gráfica. Esta representação deve expressar importantes propriedades dos dados e expressar como diferentes itens estão relacionados entre si [15]. Portanto, todas estas técnicas procuram representar, em gráficos ou figuras, informações que tentam explorar ao máximo a capacidade de percepção humana, levando à interpretações mais concretas e compreensíveis.

Ao contrário da Visualização de Dados Científicos, que possui seu foco na representação visual de dados, obtidos a partir de simulações ou coleta de dados físicos, a Visualização de Informações tem como elemento principal dados abstratos, isto é, que não possuem informações intrínsecas de geometria. Desta forma um dos problemas investigados pela comunidade científica é a descoberta de metáforas visuais para representá-los e o estudo de quais mecanismos de análise estas representações suportam. As facilidades incorporadas nas interfaces dos sistemas de visualização também são objeto de estudo. Em contraste com a Visualização de Dados Científicos, que geralmente é utilizada por cientistas altamente treinados, as interfaces criadas para a manipulação de informações podem ser utilizadas por uma comunidade leiga e diversificada, que possui diferentes níveis de educação, capacidades e necessidades.

O objetivo principal na escolha e uso de uma das técnicas de Visualização de Informações é a busca pela maior compreensão do usuário. As informações devem estar o mais claro possível, porém simplificadas. A maior preocupação ao se aplicar tais técnicas é em não prejudicar ou mesmo falhar em uma interpretação. Em alguns casos, como na área financeira e científica, uma visualização mal feita pode levar a grandes prejuízos e/ou acidentes.

Embora haja uma grande variedade de modelos e técnicas de visualização de informação, cada aplicação requer um estudo particular para determinar se a técnica selecionada é útil e utilizável. Este estudo é normalmente guiado pelos tipos de dados que deveriam ser representados e as tarefas do usuário ou o processo de análise que a visualização deveria ajudar ou deveria apoiar [10]. Muitos destes processos empregados em diferentes tipos de implementação serão apresentados neste estudo, demonstrando onde e como devem ser utilizados de acordo com as necessidades determinadas pelo usuário.

Partindo-se desta breve introdução à Visualização de Informações, é fácil observar a importância do estudo desta área, além dos benefícios que tal pesquisa pode trazer. Cada seção e subseção deste texto discute aspectos relevantes para o estudo da Visualização de Informações. A seção 2 demonstra o motivo de se aplicar a visualização para representação de informações. São apresentados o conceito de Visualização de Informações, a origem e evolução desta área, além dos fatores que motivaram o desenvolvimento da mesma, na seção 3 e suas subseções. A caracterização e prática deste estudo são discutidos nas seções 4 e 5, respectivamente. Já a elaboração de visualizações efetivas, a interação do usuário com a visualização e algumas das técnicas empregadas para representação de informações, são dispostas nas seções 6, 7 e 8, res-

pectivamente. Posteriormente a estas seções é apresentada uma conclusão deste estudo inicial e as referências relacionadas ao texto.

## 2 Visualização

Visualizar é tornar algo visual ou visível, ver uma imagem mental ou figurá-la mentalmente. Já o conceito de visualização é definido como a transformação de conceitos abstratos em imagens reais ou mentalmente visíveis, ou inserido no contexto computacional, como a conversão de números ou dados para um formato gráfico que pode ser facilmente compreendido.[19].

A visualização em si é uma atividade cognitiva, ou seja, basicamente restrita à mente humana. Como ela é algo muito difícil de ser exteriorizado, é possível encontrar inúmeras dificuldades para representar uma idéia baseada neste conceito. Demonstrar algo abstrato e com uma enormidade de conhecimentos, em apenas um gráfico ou imagem que resume e restringe a informação, porém que não desaparece com a idéia essencial é o principal desafio dos projetistas inseridos na área de Visualização de Informações.

Por que escolher a visão como o sentido para a obtenção de conhecimento? O motivo é simples, pois visualizar é algo feito naturalmente, já que a visão humana é o sentido com maior capacidade de captação de informações por unidade de tempo; é rápido e paralelo, além de ser treinado para reconhecer padrões.

Os estudiosos, sabendo do grande poder de percepção na visualização, se dispuseram a criar e desenvolver a idéia de visualizar informações. Eis que surge uma nova área repleta de desafios, a Visualização de Informações, apresentada na seção seguinte.

## 3 Visualização de Informações

A Visualização de Informações pode ser definida como: “O uso de representações visuais, interativas e suportadas por computador, de dados abstratos para ampliar a cognição” [2]. O embasamento deste texto é totalmente ligado a este conceito. Esta idéia atende a inúmeras áreas, podendo representar informações não propriamente físicas, pois estas são dispostas na Visualização de Dados Científicos, e sim informações mais abstratas, tais como: dados financeiros, informações de comércio, categorias e classificações, e conceitos não muito bem definidos.

A Visualização de Informações é uma área que dia após dia vem crescendo em importância e relevância nas tarefas humanas. A facilidade em interpretar grande número de dados em uma representação concisa, tal como uma imagem.

O estado atual dos estudos na representação de informações passou por um longo processo de evolução. Inúmeras idéias foram surgindo, umas inovadoras, outras acrescentando conhecimento às já pré-existentes. Este caminho extenso percorrido pela Visualização de Informações é abrangido na subseção seguinte.

### 3.1 Origem e evolução da Visualização de Informações

O conteúdo desta subseção foi extraído de [2]. A partir deste estudo é possível observar que os primeiros trabalhos com dados gráficos utilizando propriedades visuais abstratas são datados da época de Willian Playfair (1786), inventor da maioria das formas gráfica conhecidas hoje: o gráfico de barras, o gráfico de linhas baseado em dados econômicos e o gráfico circular. A partir deste momento os clássicos métodos de plotar dados começaram a ser desenvolvidos. Em 1967, um cartógrafo francês chamado Bertin publicou uma teoria de gráficos que

identificava os elementos básicos dos diagramas e descrevia a armação destes. Em 1983, Tufte publicou uma teoria em que enfatizava a densidade de informações úteis. Ambas as teorias tornaram-se conhecidas e influenciaram vários estudiosos que levaram o desenvolvimento da Visualização de Informações como uma disciplina.

Apesar da comunidade científica da época estar interessada em desenvolver gráficos estatísticos, Edward Tukey (1977) iniciou um movimento que enfatizava a análise de dados através de figuras para alcançar um rápido entendimento dos dados estatísticos, ao invés de explorar ao máximo a qualidade dos gráficos.

Cleveland e McGill, em 1988, escreveram o livro *Dynamic Graphics for Statistics*, que explicava as novas visualizações de informações na área de estatística. Uma das grandes questões que eles enfrentaram foi como visualizar os dados fixos em muitas variáveis. O método de coordenadas paralelas de Inselberg e Dimsdale, em 1990, contribuiu muito para o desenvolvimento da área. O grupo de Eick trabalhou com técnicas para gráficos estatísticos de larga escala fixando informações associadas com importantes problemas em redes de telecomunicações e em grandes programas computacionais [6]. A ênfase dos estatísticos era na análise dos dados multidimensionais, multi-variáveis e nos seus diferentes tipos.

Em 1985, a NSF (*National Science Foundation*), uma agência independente do Governo Norte Americano, lançou uma nova iniciativa na Visualização de Dados Científicos. A primeira Conferência de Visualização do IEEE (*Institute of Electrical and Electronics Engineers*) foi em 1990. Este grupo legou vários recursos aos cientistas, físicos e cientistas da computação.

Enquanto isto, as comunidades de computação gráfica e inteligência artificial se interessavam pelo desenho automático de apresentações visuais de dados. Este esforço foi catalisado pelo sistema de Mackinlay, que gera automaticamente ótimas representações de dados na forma de quadros, barras, gráficos ou tabelas. Este sistema formalizou a teoria de Bertin, que analisa visualizações em um jogo de componentes primitivos e especifica procedimentos para combiná-los na criação de visualizações boas, adicionando informações psicofísicas, usando-as para gerar apresentações. Partindo de Tufte, Roth e Mattis, estudiosos da área gráfica que deram uma grande contribuição ao desenvolvimento da Visualização de Informações, desenvolveram um sistema para construir visualizações mais complexas. No entanto, a principal preocupação deste grupo de estudiosos não era dar qualidade aos gráficos; a intenção estava em comunicar através deles.

Finalmente, com o avanço crescente do hardware na área gráfica foi possível criar novas interfaces com o usuário. Estas interfaces focavam-se no uso da interação com uma grande quantidade de informações, como vários bancos de dados ou grupo de documentos. Os primeiros a usarem o termo Visualização de Informações foram Robertson, Card e Mackinlay, em 1989. Feiner e Beshers [7] apresentaram um método (*Worlds Within Worlds*) para mostrar seis dimensões de dados financeiros em realidade virtual. Shneiderman desenvolveu uma técnica chamada *Dinamic Queries* que interativamente seleciona subconjuntos de itens de dados em mapas de árvores. Um exemplo desta técnica pode ser observada na **Figura 1**. Card, Robertson e Mackinlay apresentaram um jeito de usar animação e distorção para interagir com muitos dados jogados em um sistema chamado *The Information Visualizer* [3]. A preocupação não era novamente a qualidade dos gráficos, as ferramentas mais importantes deste sistema eram a interação e a animação.

Atualmente, as visualizações são voltadas para uma grande variedade de profissionais e atividades (desde acompanhamento de compra/venda em supermercados, até investigações de fraudes), envolvem apresentações dinâmicas e estáticas e são suportadas por ferramentas interativas que permitem reorganizar e explorar as figuras. A Visualização de Informações, hoje apresentada, possui um alto grau de evolução. Esta ocasionada por diversos fatores, discutidos

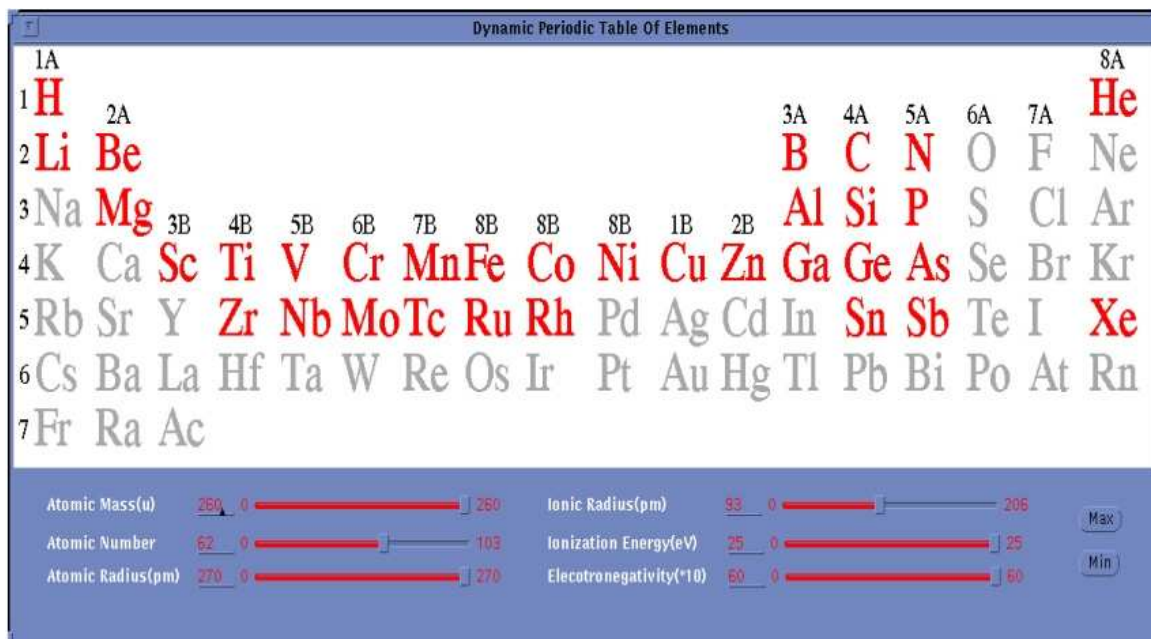


Figura 1: Tabela periódica usando a técnica de Shneiderman [18]

a seguir.

### 3.2 Fatores que motivaram o desenvolvimento da área de Visualização de Informações

O desenvolvimento da área de Visualização de Informações foi motivado por alguns fatores, todos eles relacionados ao desenvolvimento tecnológico, pois isto disponibilizou inúmeras ferramentas para manipulação e desenvolvimento de todas as técnicas atualmente apresentadas. Entre os elementos que tornaram a área de Visualização de Informações tão ampla recentemente, está a emergência de computadores pessoais poderosos e relativamente baratos, o aprimoramento de dispositivos eletrônicos dedicados à construção e visualização de gráficos de alta qualidade, por exemplo, placas de vídeo, monitores, impressoras, etc, a alta disponibilidade de memórias RAMs e de discos de baixo custo, além do desenvolvimento de técnicas e equipamentos para a interação com programas de computadores.

A seguir, na seção 4, se discute a importância da caracterização das informações na escolha de uma técnica de Visualização de Informações.

## 4 Caracterização das Informações

Informações descrevem fenômenos (ou processos) ou entidades que são objeto de estudo ou análise [15]. Desta forma, informações correspondem a atributos que podem ser caracterizados de acordo com diferentes critérios. A identificação destas características é a consideração inicial a ser feita na escolha de uma técnica de visualização para um dado domínio de aplicação.

Um primeiro critério para caracterizar atributos é o tipo de informação que eles representam [9]. Atributos podem enquadrar a entidade dentro de uma classe; um atributo deste tipo é considerado como característica, categoria, atributo nominal ou ordinal [21]. Atributos podem representar uma propriedade expressa através de uma grandeza: atributos com valores escalares, vetoriais, ou tensoriais, que assumem valores inteiros ou reais dentro de um certo intervalo.

Finalmente, atributos podem indicar a existência de um relacionamento (hierarquia ou ligação). Cabe aqui ressaltar que Ware [21], baseado em Bertin [1], argumenta que relacionamentos não são exatamente atributos de entidades, mas um tipo de dado diferente que representa as relações estabelecidas entre entidades. No trabalho de Ware, a opção escolhida foi em enquadrar um relacionamento ou ligação como um atributo.

Um segundo critério de caracterização dos atributos, não obstante dependente do primeiro, é quanto ao tipo de dado, no sentido clássico de tipo primitivo, ou seja, se o atributo pode assumir valores alfanuméricos, inteiros, reais ou simbólicos, representando este último a identificação de entidade ou fenômeno relacionado à entidade examinada no momento [9].

Finalmente, os dados podem ainda ser caracterizados de acordo com a dimensão e a natureza do domínio onde estão definidos. Dados podem estar associados a um domínio unidimensional, bidimensional, tridimensional ou  $n$ -dimensional. Este domínio pode ser contínuo, contínuo-discretizado ou discreto. Exemplos podem ser observados na **Tabela 1**, ressaltando-se que, para o caso  $n$ -dimensional, em geral, as entidades são consideradas como pontos num espaço  $n$ -dimensional de atributos.

Tabela 1: Classificação de informações (adaptada de Freitas e Wagner [9])

Critério	Classes	Significado	Exemplo
Classe de Informação	Características	Característica isolada, independente	Gênero
	Escalar	Grandeza escalar, amostrada de uma função	Temperatura
	Vetor	Grandeza vetorial, amostra de uma função	Grandeza física associada a um fluido
	Tensor	Grandeza tensorial, amostra de uma função	Grandeza física associada a um fluido
	Relacionamento	Ligação entre entidades	Link num hiperdocumento
Tipo dos valores	Alfanumérico	Valores de identificação	Gênero
	Númérico	Valores ordinais, discretos ou contínuos	Temperatura
	Símbolo	Sub-atributo	Link num hiperdocumento
Natureza do domínio	Discreto	Enumeração, conjunto finito ou infinito	Marcas de automóvel
	Contínuo	Todos os pontos no espaço 1D, 2D, 3D, $n$ -D	Superfície de um terreno
	Contínuo-Discretizado	Regiões no espaço 1D, 2D, 3D, $n$ -D	Anos(tempo discretizado)
Dimensão do domínio	1D	Dado definido no espaço 1D	Medida de uma grandeza no tempo
	2D	Valor associado ao espaço 2D	Superfície de um terreno
	3D	Valor associado à posição no espaço 3D	Volumes de dados médicos
	$n$ -D	Valores no espaço $n$ -dimensional	Dados de uma população

A classe de informação que um dado representa pode ser uma característica, escalar, vetor, tensor ou agregação. Um dado do tipo característica (ou categoria) corresponde a uma característica isolada de uma entidade onde não há uma função associada. Este é o caso da altura de indivíduos de uma população ou da flora predominante de alguma região. Um dado de natureza escalar denota um conjunto de valores amostrados de uma relação definida no espaço-domínio da entidade, como por exemplo, a temperatura e a pressão de um gás em um recipiente. Um dado do tipo vetor permite a representação de grandezas vetoriais, isto é, grandezas dotadas de

magnitude, direção e sentido, como é o caso da velocidade de fluidos. Atributos do tipo tensor são utilizados em áreas de aplicações como dinâmica de fluidos computacionais e análise de elementos finitos. Um tensor de segunda ordem em três dimensões, por exemplo, é representado por uma matriz de nove componentes [15].

A natureza ou tipo dos valores corresponde ao conceito de tipo de dado encontrado em linguagens de programação tradicionais. Um dado pode assumir valores alfanuméricos ou numéricos, dentre uma enumeração finita ou infinita ou dentro dos reais, ou, ainda, pode ser composto por valores simbólicos. O tipo dos valores está intimamente relacionado com a natureza das informações.

A natureza do domínio indica se um dado pode estar definido num domínio discreto, restrito a um conjunto finito ou infinito de elementos, num domínio contínuo ou num domínio contínuo-discretizado. Em um domínio contínuo, os valores podem estar definidos para todos os pontos (continuamente) ou para regiões. Considerando-se uma região geográfica como o objeto em estudo, o dado que designa a altura do terreno é definido para todos os pontos do domínio contínuo, enquanto outro que indica a densidade populacional em sub-áreas é definido por regiões, sobre o mesmo domínio contínuo. Uma amostragem no tempo (que é um domínio contínuo) feita mensalmente indica um domínio contínuo-discretizado.

A dimensão de um dado indica, na realidade, a dimensão do espaço (domínio) onde o objeto está definido. Um atributo pode estar definido no espaço unidimensional (1D), bidimensional (2D), tridimensional (3D) ou  $n$ -dimensional ( $nD$ ). Exemplos de domínio 1D são distâncias medidas de um ponto a outro ou alguma característica de uma entidade observada ao longo de um período de tempo. Informações clássicas 2D e 3D são valores observados para áreas geográficas num plano ou no espaço 3D, respectivamente. Exemplos de informações  $n$ -dimensionais são obtidos de aplicações que geram informações multivariadas, como por exemplo, dados de sensoriamento remoto e dados populacionais. Conforme a interpretação, a dimensão de um atributo permite representar informações espaciais, temporais, espectrais ou multidimensionais [15].

Informações que indicam relacionamentos em geral correspondem, na realidade, à especificação de um grafo representativo de uma dada relação entre entidades. No caso do relacionamento ser hierárquico, a estrutura subjacente pode ser uma árvore.

Além da classificação apresentada na **Tabela 1**, outra abordagem que pode ser utilizada é analisar os dados de acordo com a estrutura de dados adotada para representá-los e com o domínio.

Quanto à estrutura dos dados, pode-se considerar que as informações e os dados a serem visualizados encontram-se representados e organizados em uma das seguintes formas:

- **Listas e tabelas:** Permitem representar conjuntos de dados que mantêm uma relação de ordem linear entre os componentes, cujos elementos podem conter um dado primitivo ou estruturado. Como exemplo, tem-se: textos, imagens, mapas, dados multi-dimensionais, relacionais e estatísticos;
- **Árvores:** São estruturas de dados que representam uma relação hierárquica ou de composição entre os dados, ou seja, um dado é subordinado a outro dado. Como exemplo, tem-se: a estrutura de diretório de arquivos, árvores genealógicas, diagramas organizacionais, manuais e catálogos de bibliotecas;
- **Grafos:** São representações de relações mais gerais entre os elementos dos dados. Podem ser descritos num espaço euclidiano de  $n$  dimensões como sendo um conjunto de vértices e um conjunto de curvas contínuas (arestas). Pode ser representado por um



conjunto de nodos conectados por linhas (arestas). Grafos surgem como representações naturais de classificações e mapas organizacionais. Como exemplo de grafos, tem-se: estruturas de sítios, relacionamentos entre entidades, históricos de navegação, ligações entre documentos, redes de computadores, sistemas orientados à objetos (navegadores de classes), diagramas de fluxo de dados, grafos de chamadas a sub-rotinas, diagramas entidade-relacionamentos (estruturas de banco de dados e UML), diagramas de representação de conhecimentos e redes semânticas, gerenciamento de projeto (diagramas PERT), programação em lógica (árvores SLD), VLSI (esquemas de circuitos) e sistemas de gerenciamento de documentos. Na química e na biologia, os dados e informações em forma de grafos são encontrados em árvores evolucionárias, árvores do estudo de raças (filogênicas), mapas moleculares e mapas genéticos [15].

Quanto ao domínio em que as informações estão definidas, pode-se considerar que elas podem:

- Estar associadas a posições no espaço (bidimensional ou tridimensional);
- Ser temporais, ou seja, relacionadas com uma variação no tempo;
- Estar associadas a entidades;
- Estar associadas aos relacionamentos entre entidades.

Identificando as características da informação, pode ser encontrada a melhor técnica de visualização para uma representação concisa e completa. Um exemplo simples é exposto na seção seguinte, demonstrando a prática da Visualização de Informações.

## 5 Prática da Visualização de Informações

A função principal, para qual a prática da Visualização de Informações foi criada, é para absorver informações o máximo possível de um grupo de dados. Logo, esta representação é disponibilizada de diversas formas: gráficos, tabelas, grafos, entre outros. Deste modo, o usuário é capaz de adquirir um conhecimento às vezes implícito nos dados.

Um exemplo será utilizado para demonstrar a utilidade e funcionalidade em se visualizar dados de diferentes formas. No exemplo seguinte, são examinadas duas formas de visualização. Nesta visualização podem ser reconhecidas as diferenças de interpretação, dependendo da maneira que os dados são manipulados.

Primeiramente, é disposta uma tabela em que estão presentes quatro alunos e suas respectivas notas, de acordo com os bimestres do ano letivo, esta representação é disponibilizada na **Tabela 2**. Os dados, dispostos desta maneira, tornam difícil qualquer tipo de interpretação, pois os dados são apenas disponibilizados sem um objetivo concreto, sendo assim, é difícil para o usuário retirar qualquer tipo de informação neste tipo de representação. Desta maneira, a imagem não atendeu à necessidade pela qual ela foi criada, pois as informações implícitas nos dados permanecem de difícil percepção.

Buscando-se contornar este problema de percepção de uma informação proveitosa dos dados, pode-se representá-los de maneira diferente, de modo a possibilitar uma maior absorção de informações implícitas nos dados.

Colocam-se os alunos representados por linhas. Cada linha com cor diferente, tornando a visualização melhor, com uma maior distinção dos alunos. As notas são inseridas no eixo Y,

Tabela 2: Notas dos alunos em cada bimestre

	1ºAluno	2ºAluno	3ºAluno	4ºAluno
1ºBimestre	5,0	7,0	6,0	8,5
2ºBimestre	5,5	8,0	7,0	7,0
3ºBimestre	6,0	6,5	8,5	6,5
4ºBimestre	7,0	4,0	9,5	5,5

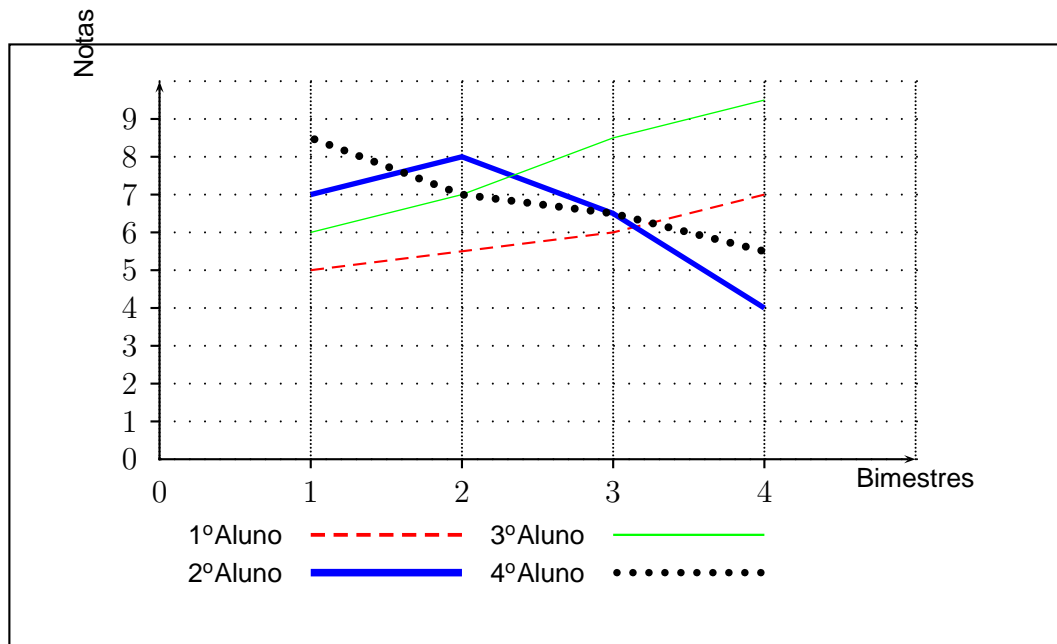


Figura 2: Notas dos alunos em cada bimestre

para uma análise da variação das notas de cada aluno. No eixo X, são colocados os bimestres do ano letivo, fazendo com que possa ser analisado o percurso pelo qual cada aluno passou, regredindo, progredindo ou apenas mantendo-se na média. Pode-se observar esta construção na **Figura 2**. O progresso ou regresso do aluno é facilmente identificado, sendo esta uma informação de grande utilidade, por exemplo, para os professores que, através da análise deste gráfico, podem identificar uma queda de rendimento de um determinado aluno, podendo assim auxiliá-lo.

Analisando o gráfico apresentado, é possível observar que o 3º Aluno apresenta-se em uma ascendência nas notas a cada bimestre. Já o 4º Aluno que no primeiro bimestre apresentava uma média alta (8,5), nos bimestres seguintes apresentou notas cada vez menores.

Após esta análise, é possível atestar o quão importante é a escolha de uma técnica de visualização na representação de dados. A partir desta escolha é que se atinge ou não o objetivo esperado, ou seja, uma absorção ideal das informações contidas nos dados.

Na primeira representação, verificou-se uma dificuldade em se extrair informações em um curto espaço de tempo. A percepção da verdadeira informação na qual os dados pretendem passar, não é tão natural na primeira visualização. Porém, na segunda apresentação, a percepção das informações implícitas se torna naturalmente clara, facilitando uma extração de conhecimento pelo usuário.

Avaliar a efetividade desta visualização exige conhecer alguns critérios, os quais são discutidos na seção 6, a seguir. Após este estudo, a garantia que a visualização proposta seja a melhor, pode ficar comprometida.

## 6 Elaboração de Visualizações efetivas

Um modelo básico que se pode seguir para a construção de uma visualização efetiva segue o roteiro exposto na **Figura 3**.

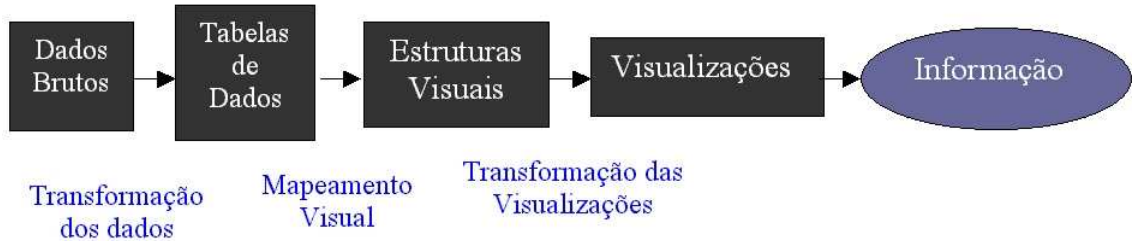


Figura 3: Processo para construção de uma visualização efetiva

Para o andamento do processo, apresentado na **Figura 3**, é necessário que os dados brutos sejam transformados em tabelas de dados através do processo de transformação de dados (subseção 6.1); posteriormente as tabelas de dados devem ser modificadas para estruturas visuais, através do método de mapeamento visual, (subseção 6.2); e finalmente para chegar à visualização é necessária a transformação das visualizações, modificando e estendendo as estruturas visuais (subseção 6.3).

### 6.1 Transformação de Dados

A transformação de dados consiste em modificar dados brutos, estes de difícil compreensão humana, em relações lógicas que são mais estruturadas e, portanto, mais fáceis de serem visualizadas. Este processo pode envolver a eliminação de dados redundantes, errados ou incompletos. Também pode ser feita a inclusão de novas informações, tais como resultados de análises estatísticas.

### 6.2 Mapeamento Visual

O mapeamento visual transforma dados presentes nas tabelas em estruturas visuais. Estruturas visuais são compostas de *substrato espacial*, *marcas*, e *propriedades gráficas das marcas*.

1. Substrato espacial: Substrato espacial é o espaço para a visualização, normalmente representado por eixos, tais como os eixos X e Y do plano cartesiano. Há 4 tipos elementares de eixos:
  - U = eixo não estruturado (ou sem eixo);
  - N = eixo nominal (a região é dividida em sub-regiões);
  - O = eixo ordenado (a região é dividida em sub-regiões), e a ordem das mesmas tem importância;
  - Q = eixo quantitativo (a região tem uma métrica).
2. Marcas Visuais: Marcas visuais são símbolos gráficos utilizados para representar os itens de dados. Tipos de marcas:
  - Pontos;

- Linhas;
  - Áreas;
  - Volumes;
  - Figuras.
3. Propriedades Gráficas: As propriedades Gráficas são atributos gráficos das marcas visuais associadas aos atributos dos itens de dados das tabelas. Algumas propriedades gráficas:
- Posição, x, y, z;
  - Tamanho, comprimento, área, volume;
  - Orientação, ângulo, inclinação;
  - Cor, brilho, textura;
  - Forma;
  - Animação, tempo.

### 6.3 Transformação das Visualizações

As transformações das visualizações possibilitam modificar e estender estruturas visuais interativamente. Há três transformações básicas de visualizações:

- **Testes de localização** - permitem obter informações adicionais sobre um item da tabela de dados.
- **Controles de ponto de vista** - para zoom e troca da imagem, de modo a oferecer diferentes visões. Uma técnica chamada visão geral+ detalhe (*overview+detail*) também é um tipo de controle de ponto de vista.
- **Distorção** - distorções da imagem visando criar visualizações do tipo foco+contexto.

### 6.4 Características de uma Visualização Efetiva

O principal objetivo em se visualizar dados é a obtenção de informações para uma melhor compreensão do problema. A expressividade e a efetividade são características a serem alcançadas na representação dos dados, tornando sua visualização mais clara e eficaz.

#### 6.4.1 Expressividade

Apenas os dados existentes na tabela, e nada mais, devem ser representados na visualização. Informações adicionais podem dificultar ou mesmo mudar a interpretação do gráfico ou figura.

#### 6.4.2 Efetividade

A visualização deve ser de rápida interpretação, deve possuir muitas informações distintas, ou deve induzir a uma quantidade menor de erros do que outras figuras.

### 6.4.3 Passos a serem seguidos para uma Visualização Efetiva

Alguns passos devem ser seguidos para a obtenção de uma visualização efetiva, eles são apresentados a seguir:

- Associar os atributos de maior importância dos dados a atributos visuais de maior clareza. O grau de importância de alguns atributos foram definidos por Cleveland e McGill [4], apresentado na **Figura 4**, e por Mackinlay [16], como mostra a **Figura 5**.

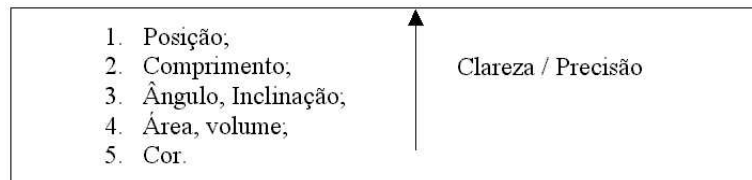


Figura 4: Grau de importância dos atributos (Esquema proposto por Cleveland e McGill)[4]

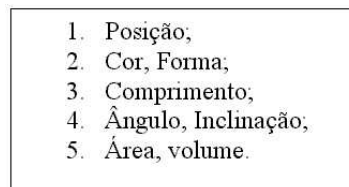


Figura 5: Dados das categorias da hipótese de Mackinlay [16]

- Retirar gráficos e textos desnecessários na visualização, elementos que não são dados devem ser descartados.
  - Procurar sempre a simplicidade (tendo duas formas para representação, deve-se fazer escolha da mais simples).
  - Aumentar a densidade de dados por centímetro quadrado.
  - Permitir diferentes formas de interação com a figura.

A criação de uma visualização efetiva exige, além dos passos apresentados anteriormente, uma interação com o usuário. A explicação e apresentação de algumas dessas técnicas são colocadas na seção 7.

## 7 Interação com Visualizações

Uma visualização de modo estático, por si só, na maioria dos casos, não possibilita uma avaliação concreta e eficiente de uma grande quantidade de dados. Normalmente este problema é contornado utilizando-se de ferramentas que possibilitem ao usuário explorar diversas ações em diferentes níveis da visualização. Estas ações caracterizam-se pela modificação da representação visual, possibilitando que novos aspectos dos dados sejam observados e interpretados [8].

A um nível mais básico estão as funções de navegação e reposicionamento do observador, quer pelo deslocamento horizontal/vertical de uma *scroll bar*, por exemplo, no caso de uma representação plana, como pelo deslocamento de uma câmera virtual no espaço (ou rotação do

conjunto de dados) no caso de uma representação 3D. Num segundo nível, estão funções de seleção de elementos de dados de interesse que podem provocar tanto um reposicionamento do conjunto de dados para uma melhor visualização como um detalhamento (*zooming* semântico) de uma parte desse conjunto e, eventualmente a supressão de outra parte. Operações de poda (*prunning*), agrupamento (*clustering*) e expansão são fundamentais no apoio ao processo de navegação e exploração em diagramas que representam grandes hierarquias e grafos, e são encontradas na maioria das técnicas [12]. Nestas operações podem ocorrer três situações:

1. A representação é alterada para mostrar apenas a região de interesse selecionada;
2. A região de interesse ocupa o campo de visão principal e o restante do conjunto de dados é mantido em área à parte;
3. A região de interesse e a visão geral são exibidas concomitantemente.

A alternativa 2 provê o que se costuma chamar de visão geral+detalhe, enquanto as técnicas que adotaram a alternativa 2 são denominadas de foco+contexto [2].

Observando em um nível superior, estariam funções que permitiriam representar visualmente apenas partes do conjunto dos dados, dependendo da satisfação de certos critérios. Tais funções podem simplesmente estabelecer filtros com base nos valores de atributos ou corresponder a consultas dinâmicas apoiadas em processos de mineração de dados.

As interações das visualizações com os usuários são de grande importância para a obtenção de informações. Os objetivos das interações (subseção 7.1) e o modelo para empregá-las (subseção 7.2) são expostos a seguir.

## 7.1 Objetivos das Interações

Permitir ao usuário explorar diferentes “cenários” de forma a melhor entender os dados que estão sendo visualizados.

Direcionar ao computador o esforço de explorar os dados, enquanto o usuário apenas observa o que acontece quando a visualização muda.

## 7.2 Modelo para interação

O modelo básico para construção de uma visualização efetiva e com interação com o usuário, é representado na Figura 6:

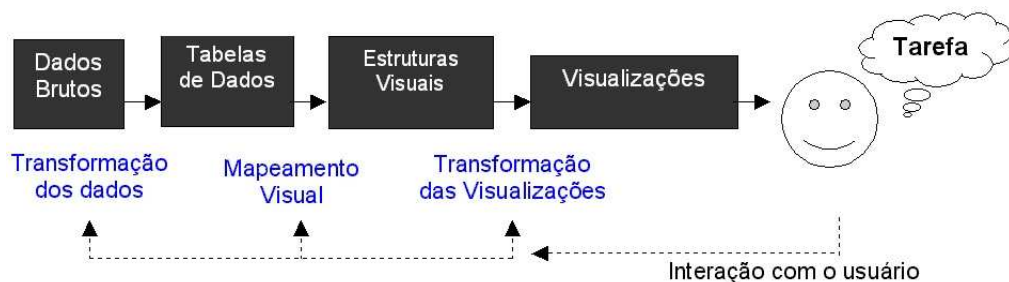


Figura 6: Modelo de referência de visualização [2]

Na subseção seguinte são exploradas as técnicas utilizadas para se obter uma melhor interação do usuário com a Visualização.

### 7.3 Técnicas de Interação

A seguir são descritas algumas técnicas de interação, que são diferenciadas de acordo com o modelo seguido por cada uma:

- **Brushing** - é um exemplo de interação com transformação de dados. Possui um método interativo que permite selecionar na tela dados pontuais específicos ou subconjuntos de dados, possibilitando a identificação de suas características comuns ou a investigação de seus efeitos sobre as relações entre variáveis consideradas relevantes.
- **Pivot Tables** - é uma interação com mapeamentos visuais. Esta técnica está presente em planilhas eletrônicas modernas que permite ao usuário manipular o mapeamento de dados para linhas e colunas.
- **Direct Selection** - é uma interação com transformações de visualizações, que permite selecionar e destacar objetos gráficos individualmente ou em grupos.

Na seção seguinte algumas das várias técnicas para Visualização de Informações são discutidas, mostrando sua utilidade ou não para uma determinada representação de um conjunto de dados.

## 8 Técnicas de Visualização de Informações

As técnicas de Visualização de Informações utilizam representações ou metáforas visuais para exibir graficamente dados que geralmente não possuem representação direta, óbvia e natural [15]. Nas várias técnicas, frequentemente, os autores buscaram inspiração em objetos do mundo real (ou geométricos), para mapear o conjunto de informações. As técnicas podem utilizar representações visuais 1D, 2D ou 3D, não necessariamente de acordo com a dimensão do espaço de informação.

Algumas técnicas são apresentadas a seguir, demonstrando uma representação eficiente de um conjunto variado de dados. Sabendo-se que a Internet é um imenso repositório de informações, estas técnicas também são utilizadas em aplicações para *Web*.

### 8.1 Bifocal Display

Na técnica *Bifocal Display* [20], itens de informação (documentos, figuras, gráficos, etc.) são apresentados em três áreas distintas, sendo a central aquela que contém a informação em foco, em destaque, e as outras informações do contexto geral são apresentados nas laterais da região focal. A parte central com o foco ocupa uma área maior que as laterais preenchidas pela informação contextual que, por isso, é exibida de forma distorcida. Esta distorção é basicamente na direção X podendo ser aplicada também na direção Y (**Figura 7**).

A técnica *Bifocal Display* possui uma vantagem na representação de dados, porque ela permite acesso a grande quantidade de informações. Entretanto, não é facilmente integrada a detalhes e ao contexto por ser bidimensional.

### 8.2 Flip Zooming

A técnica *Flip Zooming* [13] é uma técnica de visualização foco+contexto para conjuntos hierárquicos de informações. A informação visualizada é representada por um número de objetos distintos em uma ordem seqüencial (**Figura 8**). Cada objeto é apresentado como uma área

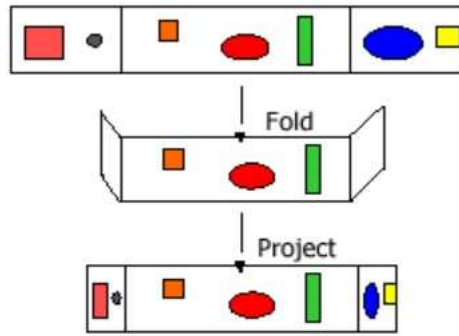


Figura 7: Representação visual da técnica *Bifocal Display*

retangular, um “ladrilho” (*tile*), com um deles como o foco (**Figura 9**). Este foco é colocado no centro da área de exibição com outros ladrilhos colocados ao redor do foco, que recebe uma área com maior destaque na tela (**Figura 9**). O *Flip Zooming* divide a informação em porções adequadas que são apresentadas dentro da área reservada a um ladrilho. No caso da informação consistir de objetos distintos, mas sem uma ordem seqüencial definida, um dos ladrilhos deve ser tomado como a informação inicial. Um exemplo disto é uma coleção não-relacionada de arquivos de imagens que podem receber uma ordem seqüencial pela ordenação alfanumérica dos nomes dos arquivos. *Flip Zooming* usa uma forma de distorção através de uma função de escala que apresenta os ladrilhos no contexto geral e no foco de forma que ambos estejam visíveis frontalmente. Isto torna o reconhecimento de um dado objeto mais fácil, visto que apenas a localização e o tamanho do objeto pode mudar.

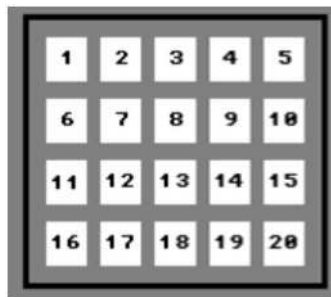


Figura 8: Representação de uma estrutura seqüencial de informações [14]

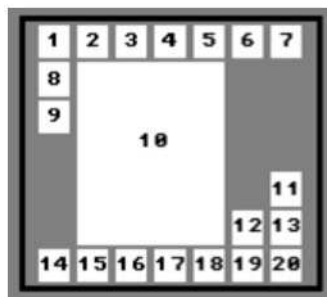


Figura 9: Representação utilizando *Flip Zooming* com foco no objeto 10 [14]

A técnica *Flip Zooming* pode dar ao usuário uma visão geral do documento. Isto facilita o acesso a qualquer parte de um documento. As vantagens de se usar esta técnica são inúmeras:



- Não há deformação do texto em foco;
- Tempo de performance é melhor;
- Possibilidade de incluir outros tipos de informações que não estão em foco;
- Implementação mais fácil.

A técnica permite mover o foco seqüencialmente através dos ladrilhos e modificar o foco acessando de forma randômica. Entretanto, a separação entre eles no contexto pode ficar comprometida quando o número de peças é muito grande, pois sua área pode ficar muito reduzida para uma boa visualização. Nesse caso, o problema pode ser contornado definindo-se estruturas hierárquicas, em que um modo mais simples é representado, mas que posteriormente poderá expandir para algo mais completo.

### 8.3 *Perspective Wall*

A técnica *Perspective Wall* [17] utiliza visões detalhada e contextual para suportar a visualização linear de espaços de informações estruturadas. Isto é feito pela transformação do layout 2D para 3D, que integra uma região para visualização de detalhes com regiões em perspectiva para visualização do contexto (**Figura 10**). Essa distorção intuitiva do *layout* fornece uma utilização eficiente do espaço e permite transição suave das visões. Desta forma uma seção detalhada pode ser visualizada em primeiro plano enquanto o resto da estrutura aparece em segundo plano.

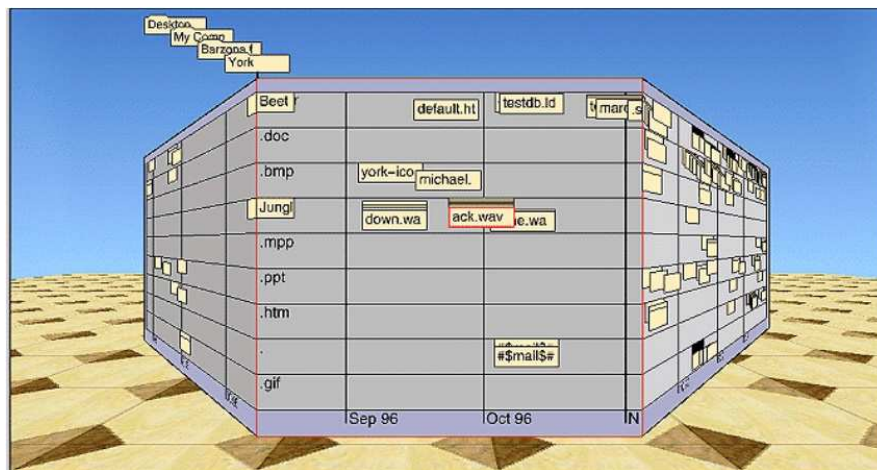


Figura 10: *Perspective Wall*: Histórico temporal de dados lineares [17]

O *Perspective Wall*, que usa o mesmo conceito da técnica *Bifocal Display*, foi proposto para visualizar vários tipos de estruturas temporárias de informações, incluindo sistemas de arquivos e documentos. A técnica suporta animações interativas 3D e efeitos visuais como brilho (*highlighting*) e sombreamento (*shading*) que auxiliam a percepção visual 3D e a localização do usuário na estrutura linear. Tem como vantagem manter sempre o contexto, fazendo com que o usuário não perca a localização das informações.

A técnica *Perspective Wall*, assim como toda técnica de foco-contexto, apesar das inúmeras vantagens, quando o problema exige uma exibição de uma enormidade de dados na tela, essa técnica torna-se inviável, por restringir a aparição das informações. Esta ressalva, não diminui a usabilidade desta visualização.

## 8.4 Hyperbolic Tree

Esta técnica representa hierarquias através de um *layout* radial definido no plano hiperbólico e depois mapeado para um disco 2D. Apresenta aspectos de construção como o efeito *fisheye* [11] aliados a mecanismo simples de navegação pela indicação de um nodo de interesse, que é exibido no centro da representação em detalhe e o contexto é mantido pela exibição do restante do diagrama com nodos diminuindo de tamanho até serem suprimidos na borda do círculo. A técnica *Hyperbolic Tree* é empregada atualmente no *browser* comercializado pela *Inxight Co. (Star Tree Studio)*[5].

Este tipo de implementação exibe inicialmente a raiz da árvore como ponto central da estrutura, a qual pode ser modificada pela seleção de um novo foco ou através do arrasto do novo ponto de interesse até o centro do disco (**Figura 11** e **Figura 12**).

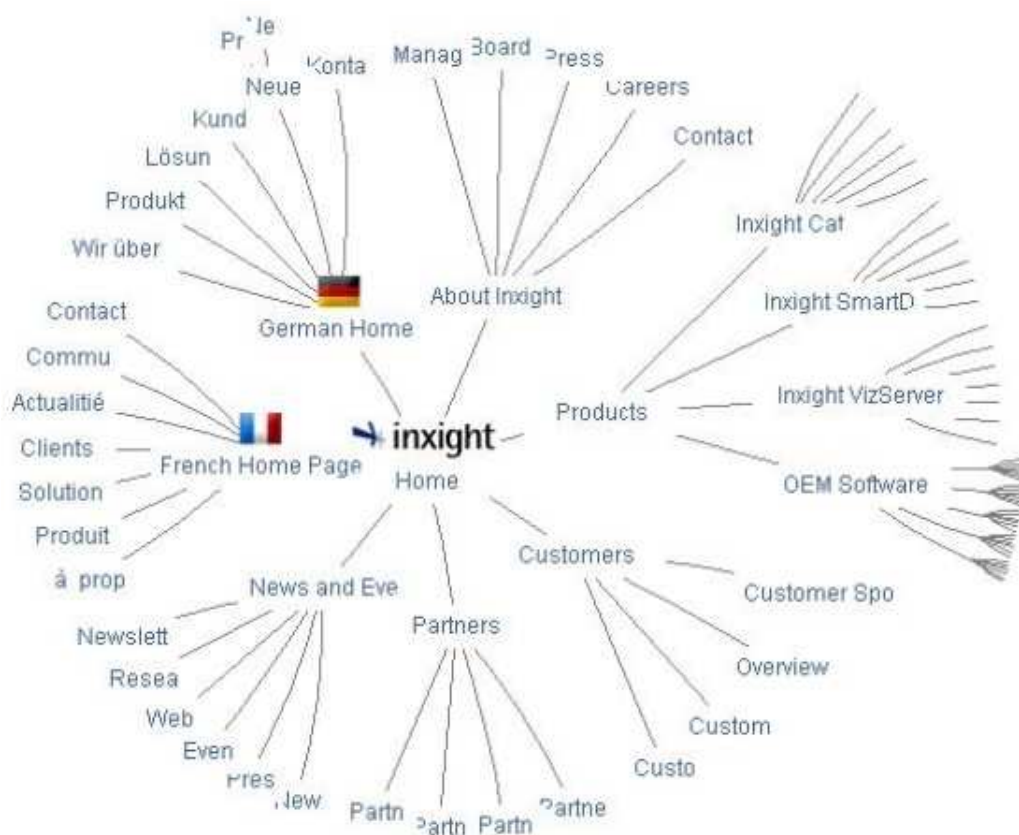


Figura 11: Mapa do sítio da *Inxight Software, Inc.*

As figuras demonstram o mapa do sítio da *Inxight Co.*, tendo como centro da raiz a página principal do sítio (**Figura 11**). Com o cursor é possível arrastar a árvore até *Partners* (**Figura 12**), por exemplo, ou a qualquer outro *link* demonstrado na figura, posicionando o elemento escolhido no centro da estrutura.

Na seção seguinte, são feitas algumas considerações finais a respeito do estudo apresentado neste texto, abrangendo as possíveis evoluções pela qual a área de Visualização de Informações deverá passar no decorrer dos próximos anos.

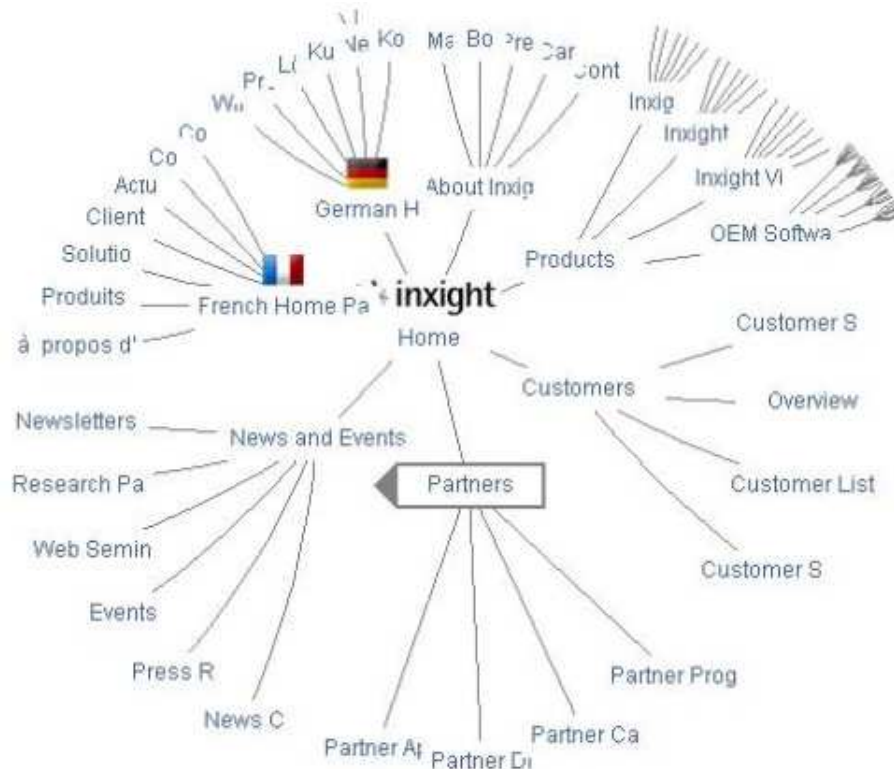


Figura 12: Mapa do site da *Inxight Software, Inc.*

## 9 Conclusão

Nos últimos anos o volume crescente de informações fornecido por diversas aplicações, diferentes instrumentos e, principalmente, a *Web* acarretou o desenvolvimento de técnicas para seleção de uma parte dos dados do subconjunto de informações que são relevantes para uma necessidade ou objetivo particular. A pesquisa em sistemas de consultas visuais, mineração de dados e técnicas de visualização interativas resultaram em uma grande variedade de apresentações visuais e técnicas de interação que podem ser aplicadas em diferentes situações. Esta estreita relação deverá fazer surgir técnicas mais poderosas, combinando algoritmos estatísticos típicos de mineração de dados e testes de hipóteses de um lado com técnicas de visualização de dados e análise exploratória visual de outro.

Apesar das técnicas de visualização serem inúmeras, cada aplicação requer um estudo particular a fim de determinar a eficiência ou não de um determinado modelo proposto. Este estudo baseia-se no problema em si, no tipo de informação a ser representada e as tarefas de usuários ou processos de análise que a visualização deveria auxiliar ou suportar.

No contexto atual, em que os dados são muitos, de diferentes fontes e características, juntar essa “confusão” de informações em uma visualização completa e intuitiva, não é uma tarefa fácil. A partir deste princípio, pode-se imaginar o quanto a tarefa de escolher e avaliar qual o método melhor para ser aplicado, é difícil.

O contexto, estudado e discutido, abrange um conjunto de fatores que se auto-relacionam; os aspectos visuais e a interface dos mecanismos de interação, no qual possibilitam ao usuário navegar e consultar um conjunto de dados através de uma representação visual.

## 10 Agradecimento

Ao Prof. Dr. Hugo Alexandre Dantas do Nascimento pela avaliação do presente trabalho e pelas sugestões feitas, as quais muito contribuíram para a melhoria do texto original.

## Referências

- [1] BERTIN, J. **Graphics and graphic information processing**. p. 62–65, 1999.
- [2] CARD, S. K; MACKINLAY, J. D; SHNEIDERMAN, B. **Readings in Information Visualization: Using Vision to Think**. Morgan Kaufmann Series in Interactive Technologies. Morgan Kaufmann Publishers Inc., 1999.
- [3] CARD, S. K; ROBERTSON, G. G; MACKINLAY, J. D. **The information visualizer, an information workspace**. In: PROCEEDINGS OF THE SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, p. 181–186. ACM Press, 1991.
- [4] CLEVELAND, W. S; MCGILL, R. **An experiment in graphical perception**. Int. J. Man-Mach. Stud., 25(5):491–500.
- [5] CO, I. **Star Tree Studio Browser**. <ftp://www.inxight.com/products/sdks/st/>, 2004.
- [6] EICK, S. G; STEFFEN, J. L; SUMNER, JR., E. E. **Seesoft - a tool for visualizing line oriented software statistics**. p. 419–430, 1999.
- [7] FEINER, S. K; BESHES, C. **Worlds within worlds: metaphors for exploring n-dimensional virtual worlds**. In: PROCEEDINGS OF THE 3RD ANNUAL ACM SIGGRAPH SYMPOSIUM ON USER INTERFACE SOFTWARE AND TECHNOLOGY, p. 76–83, 1990.
- [8] FREITAS, C. M. D. S; CHUBACHI, O. M; LUZZARDI, P. R. G; CAVA, R. A. **Introdução à Visualização de Informações**. Revista de Informatica Teórica e Aplicada, 8(2):143–158, 2001.
- [9] FREITAS, C. M. D. S; WAGNER, F. R. **Ferramentas de suporte às tarefas da análise exploratória visual**. Revista de Informática Teórica e Aplicada, 2(1):5–36, 1995.
- [10] FREITAS, C. M. D. S; WINCKLER, M. A; LUZZARDI, P. R. G; PIMENTA, M; NEDEL, L. P. **Data usability Issues in Information Visualization Applications**. In: WORKSHOP ON SPATIAL DATA USABILITY, volume 1, p. 1–1, Wageningen, Holanda, 2001. Center for Geo-Information.
- [11] FURNAS, G. W. **The FISHEYE view: a new look at structured files**. p. 312–330, 1999.
- [12] HERMAN, I; MELANÇON, G; MARSHALL, M. **Graph Visualization and Navigation in Information Visualization: A Survey**. IEEE Transactions on Visualization and Computer Graphics, 6(1):24–42, janeiro/março 2000.
- [13] HOLMQUIST, L. E. **Focus+Context Visualization with Flip Zooming and the Zoom Browser**. In: EXTENDED ABSTRACT OF CONFERENCE ON HUMAN FACTORS IN COMPUTER SYSTEMS, p. 263–264. ACM Press, 1997.

- [14] HOLMQUIST, L. E. **The Zoom Browser: Showing Simultaneous Detail and Overview in Large Documents.** In: HUMAN IT, volume 2, p. 333–340, Sweden, 1999. ITH.
- [15] LUZZARDI, P. R. G. **Critérios de Avaliação de Técnicas de Visualização de Informações Hierárquicas.** Tese de doutorado, UFRGS, Porto Alegre, 2003.
- [16] MACKINLAY, J. **Automating the design of graphical presentation of relational information.** ACM Transactions on Graphics, 5(2):110–141.
- [17] MACKINLAY, J. D; ROBERTSON, G. G; CARD, S. K. **The perspective wall: detail and context smoothly integrated.** In: PROCEEDINGS OF THE SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, p. 173–176. ACM Press, 1991.
- [18] SHNEIDERMAN, B. **Dynamic queries for visual information seeking.** p. 236–243, 1999.
- [19] SPENCE, R. **Information Visualization.** ACM Press, 2001.
- [20] SPENCE, R; APPERLEY, M. **Data base navigation: an office environment for the professional.** p. 333–340, 1999.
- [21] WARE, C. **Information visualization: perception for design.** Morgan Kaufmann Publishers Inc., 2000.