

Explore

Discover

Understand



A New Way to Look at Search

Powering Search Results with Visualization April 2008

Marcelline Saunders
VP of Marketing
Groxis

Overview

This presentation discusses trends and benefits of the convergence of search and visualization tools, including brief historical information to provide context. It looks at the latest tools as well as reviews the early adoption and specific use cases in a number of verticals. Finally it discusses how powering search results with visualization will impact the market moving forward, including the impact on today's most popular business models.

Outline

- Background
- Evolution and Roots of Visualization
- Business Adoption
- Convergence of Search and Visualization
- Future

From Data to Pictures to Insight...

“Visualization” is the rendering of realistic or abstract images that represent an underlying data set. The data set being represented can be the product of physical measurements of the real world, or it can be the product of a computational model.

- IBM, alphaworks,
- <http://www.alphaworks.ibm.com/visualization/newto>

Early Visualization – Periodic Table

ОПЫТЪ СИСТЕМЫ ЭЛЕМЕНТОВЪ.
ОСНОВАННОЙ НА ИХЪ АТОМНОМЪ ВѢСѢ И ФИЗИЧЕСКОМЪ СХОДСТВѢ.

				Ti—50	Zr—90	γ—180.
				V—51	Nb—94	Ta—182.
				Cr—52	Mo—96	W—186.
				Mn—55	Rh—104,4	Pt—197,4.
				Fe—56	Ru—104,4	Ir—198.
				Ni—59	Pd—106,4	Os—199.
				Co—59	Pt—106,4	Os—199.
				Cu—63,4	Ag—108	Hg—200.
H—1				Zn—65,2	Cd—112	
	Be—9,4	Mg—24		U—116	Au—197?	
	B—11	Al—27,4	?—68	Sn—118		
	C—12	Si—28	?—70	Sb—122	Bi—210?	
	N—14	P—31	As—75	Te—128?		
	O—16	S—32	Se—78,4	I—127		
	F—19	Cl—35,4	Br—80	Cs—133	Tl—204.	
Li—7	Na—23	K—39	Rb—85,4	Ba—137	Pb—207.	
		Ca—40	Sr—87,4			
		?—45	Ce—92			
		?Er—56	La—94			
		?Yt—60	Di—92			
		?In—75,4	Th—118?			

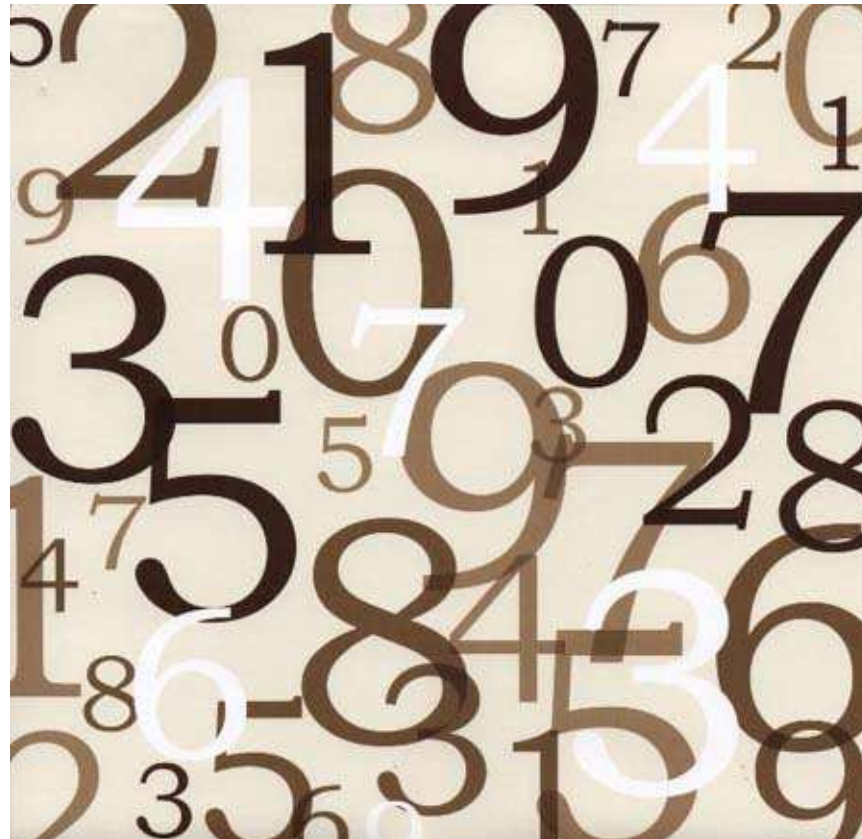
Mendeleev's Periodic Table of 1869¹

				Ti 50	Zr 90	? 100
				V 51	Nb 94	Ta 182
				Cr 52	Mo 96	W 186
				Mn 55	Rh 104.4	Pt 197.4
				Fe 56	Ru 104.4	Ir 198
				Ni, Co 59	Pd 106.6	Os 199
H 1				Cu 63.4	Ag 108	Hg 200
	Be 9.4	Mg 24		Zn 65.2	Cd 112	
	B 11	Al 27.4		? 68	U 116	Au 197?
	C 12	Si 28		? 70	Sn 118	
	N 14	P 31		As 75	Sb 122	Bi 210?
	O 16	S 32		Se 79.4	Te 128?	
	F 19	Cl 35.5		Br 80	I 127	
Li 7	Na 23	K 39		Rb 85.4	Cs 133	Tl 204
		Ca 40		Sr 87.6	Ba 137	Pb 207
		? 45		Ce 92		
		Er? 56		La 94		
		Yt? 60		Di 95		
		In 75.6?		Th 118?		

Periodic Table

- Overview of all known elements
- Relationships amongst the elements that previously had been studied in isolation
 - Properties correlate with position on the table
 - Atomic mass, density, electron affinity, physical property (gas, solid, etc.)
- Identified gaps
 - Mendeleev predicted the properties of the missing elements
 - Cornerstone of understanding radioactive decay and atomic theory
- Allowed for predictions on the nature of the future elements
 - Some gaps were for elements that are not available in nature
- Meets the needs of chemists (their tasks)

Interpreting data



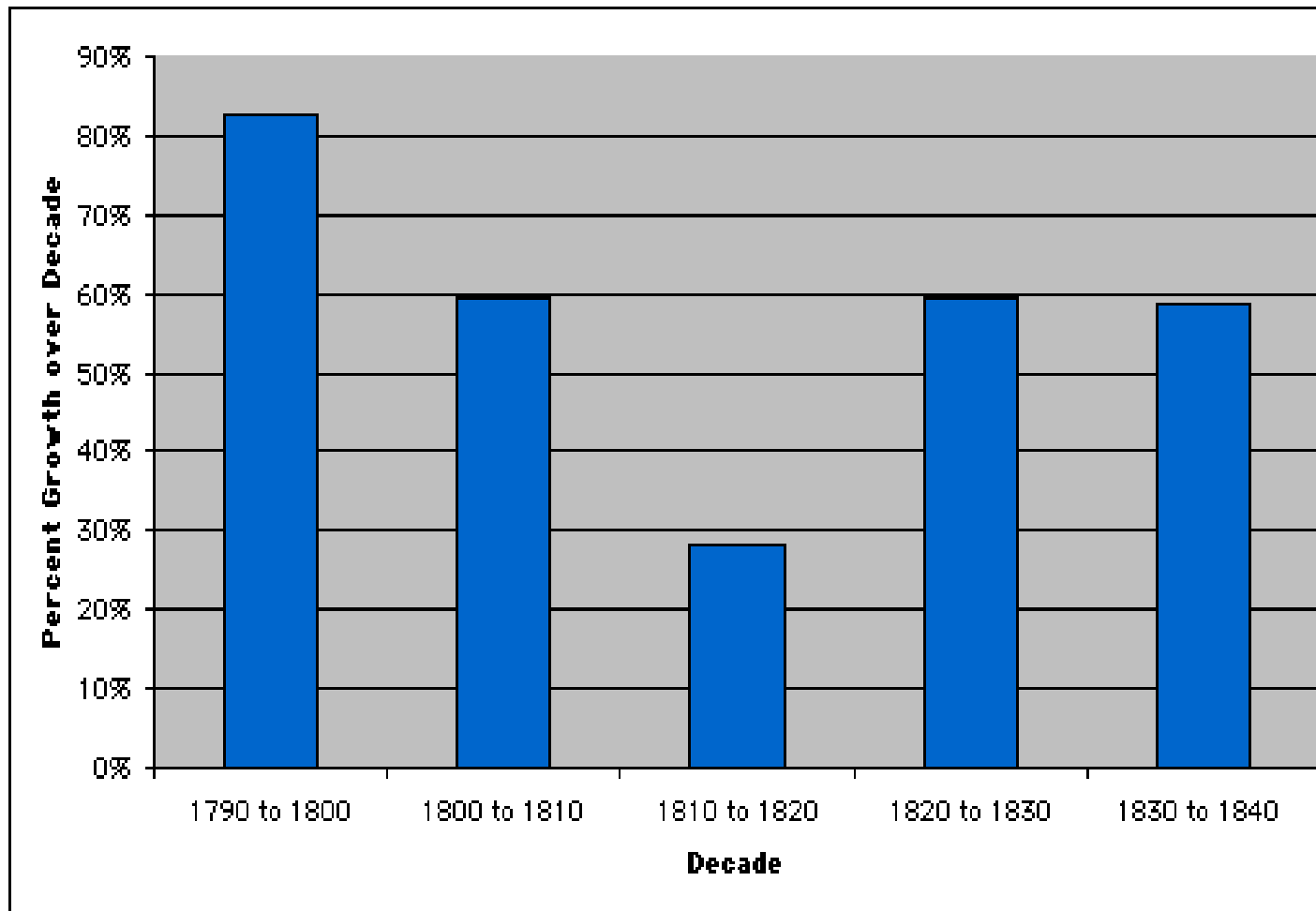
Organize

U.S Census, New York City

Year	Population
1790	33,131
1800	60,489
1810	96,373
1820	123,704
1830	197,112
1840	312,710

Understand

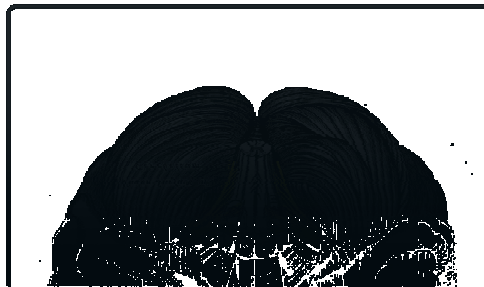
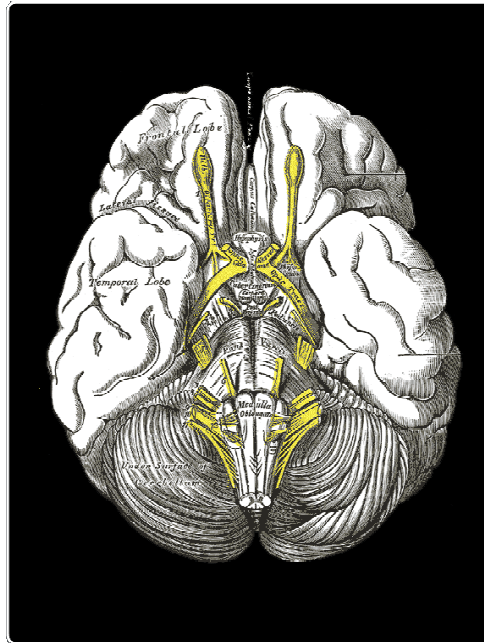
The Rate of Population Growth by Decade, New York City, 1790-1840



How information is processed

Left Hemisphere

- Analyzes
 - Break the whole into parts
 - Logic
- Sequential
- Language
 - Words
- Numbers



Right Hemisphere

- Synthesizes
- Apprehends the whole
- Process Context
 - Intonation
 - Body language
- Think in images
- Imagination

Auditory-Sequential vs Visual-Spatial Learner

• Auditory-Sequential Learner

- Drill and repetition
- Sequentially

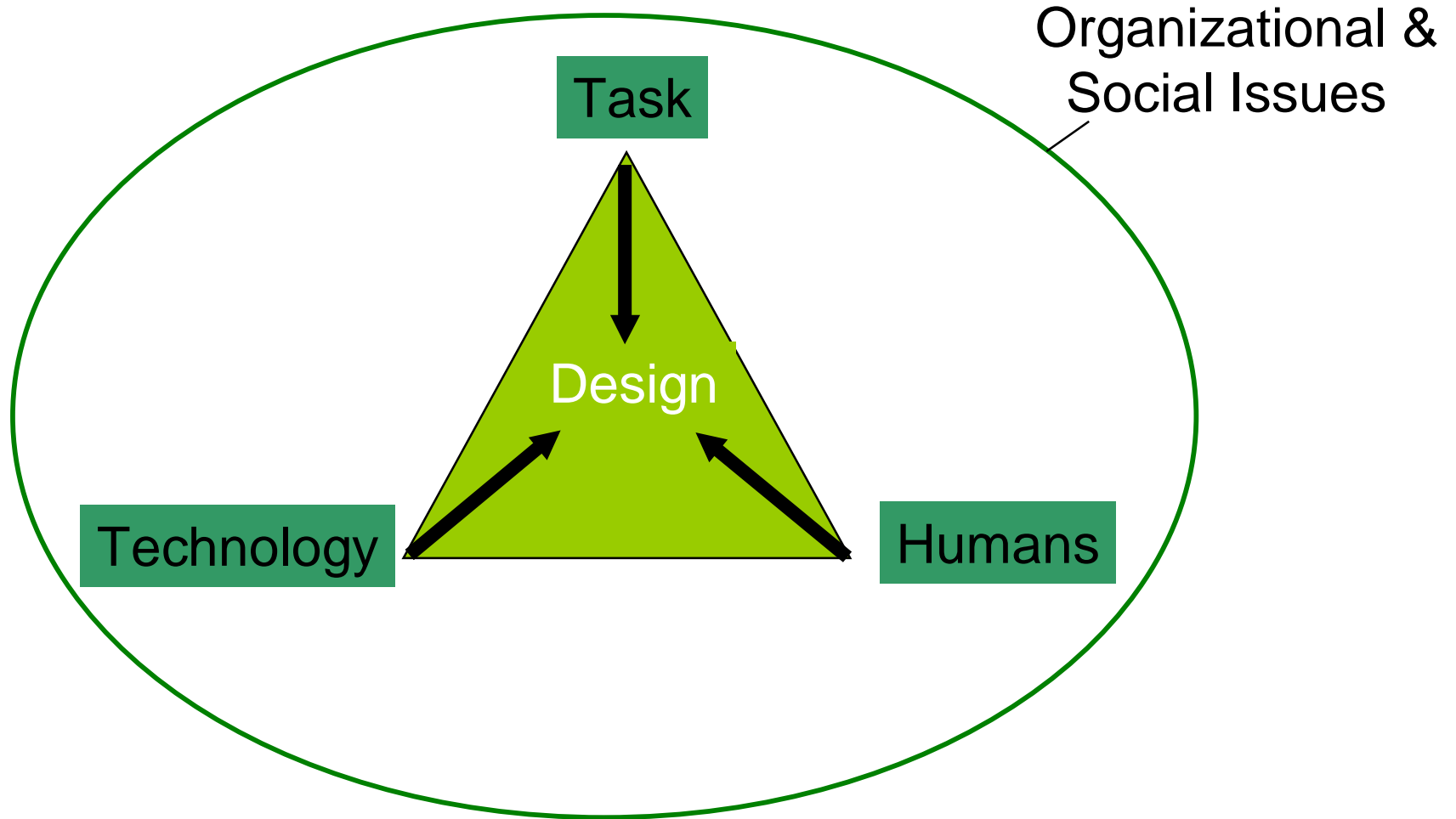
• Visual-Spatial learner

- Mechanical leaning
- Maps
 - Sees complex routes
- No language just images
- 1/3 of students are visual learners
 - Linda Silverman, *Upside-down Brilliance*, 2002

• Example: Art

- Art is conceived from images not words
- Once an image is conceived, it is the blue print

Visualization Drivers



Data Mining Techniques

• Clustering

- Discern groupings within data. Extract the categories from the data itself

• Classification

- Assigning categories to features or trends within the data. Not the same as clustering, as classification already exists or is created by the investigator

• Summarization

- Compact representation for data
- Example: statistics – mean and standard deviation
- Important in the visualization and interactive exploration of a data set

• Regression

- Function that approximates the mathematical relationship between two numerical attributes

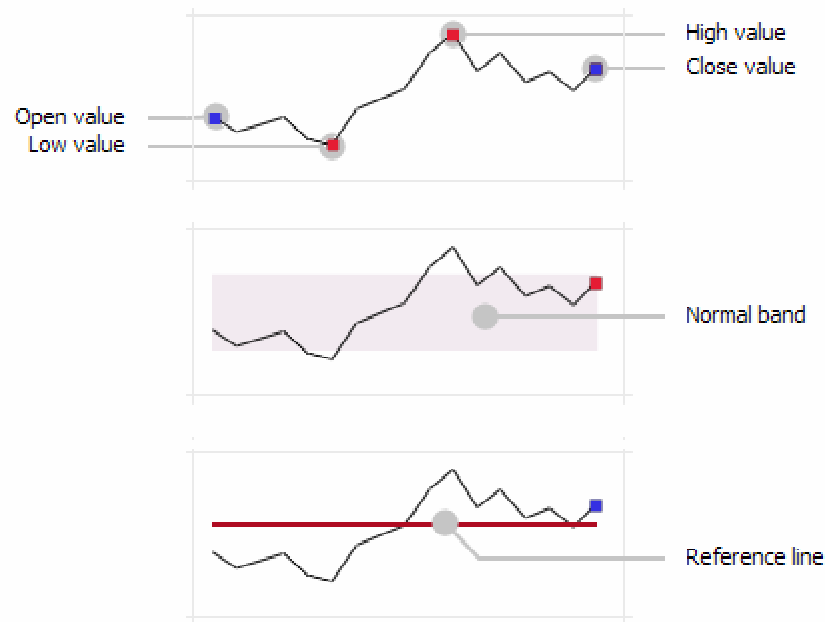
• Dependency modeling

• Change and deviation detection

- Significant changes in data from previous values or looking for data outside a “normal” range

Early Generation

- Simple rendering of every event in a visual way
- Static display
 - Snapshot in time
- 2 Dimensional
- Data representations
 - Pie and bar charts
 - Candlestick
 - Stock
 - Sparkline
- Landscapes
- Categorization



Technology Drivers

- Size and complexity of data structures
 - Visualize thousands of data points
 - Neural nets
- Real-time data feeds
- Computer processing power
- Algorithms and Computational Linguistics
 - Dynamic clustering
 - Bayesian
 - Sentiment analysis
- Animation
- 3D

Usability Impacting Visualization

- “Wholistic” view
- User interaction
 - Filter
 - Zoom & drill down
 - Details on demand
 - Relate
 - Export and save
- Fits the task at hand
- Apple influence
 - ipod/iphone browse of record covers
 - Vista
 - Sets user expectations



Strong Adoption in Business

❁ Tasks

- Decision Making
- Knowledge Management
- Business Performance Management
- Customer Relationship Management

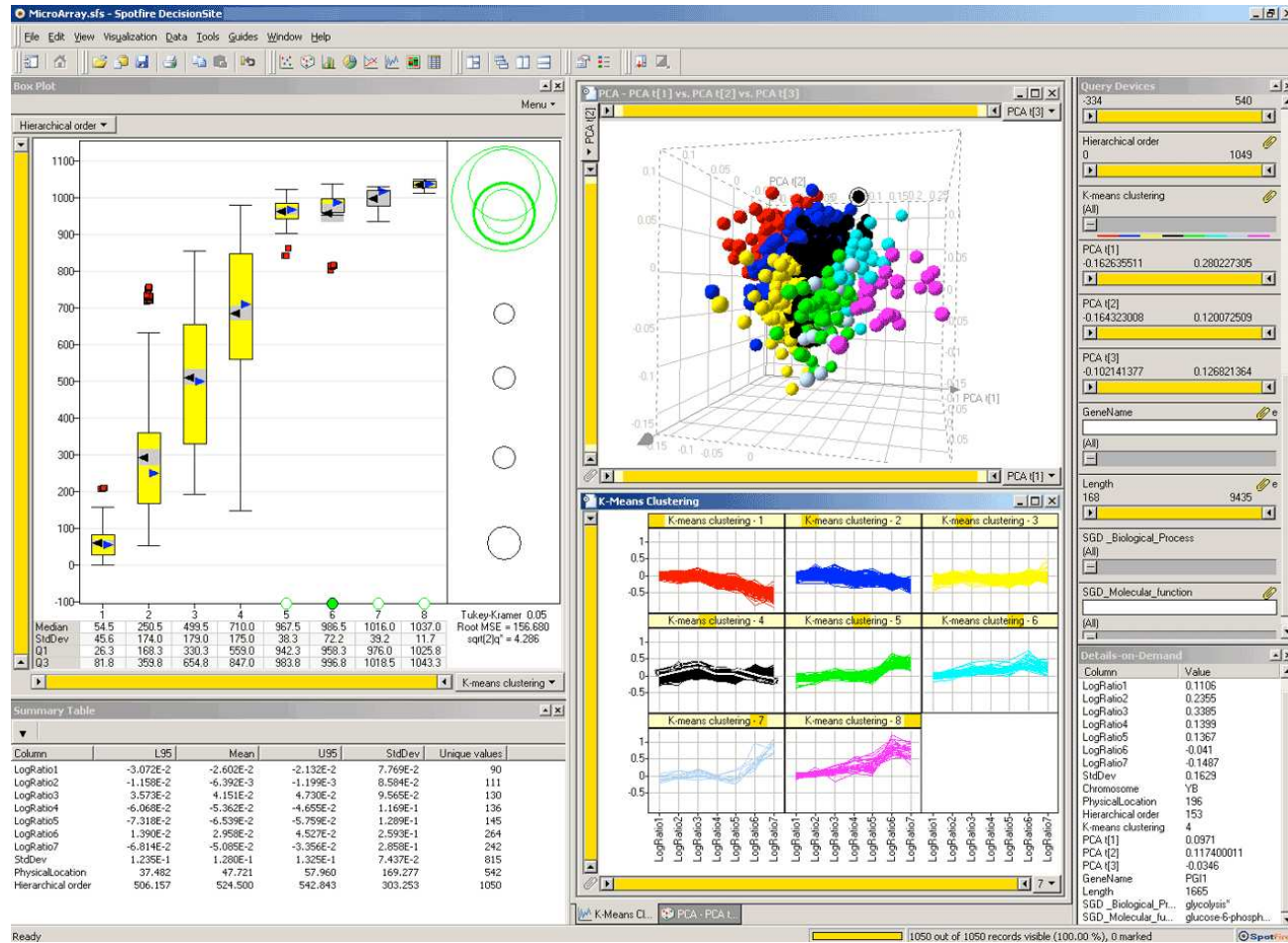
❁ Interaction allowed

- Dynamic charts to discover new views of information
- Live data connectivity

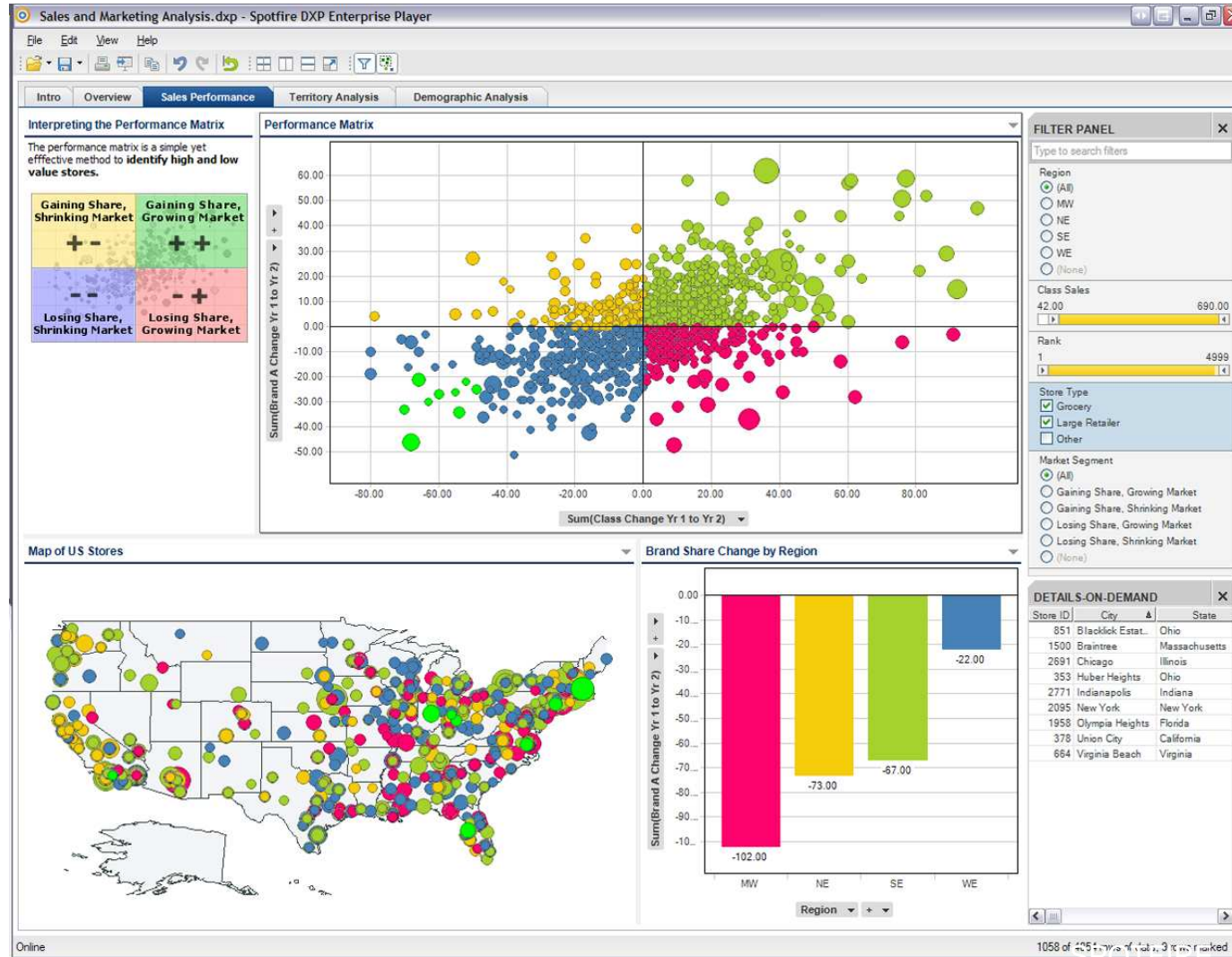
❁ Tools

- Enterprise charting systems (ECS)
- Dashboards
- Geographic, spatial
- Widgets - embedded in portals

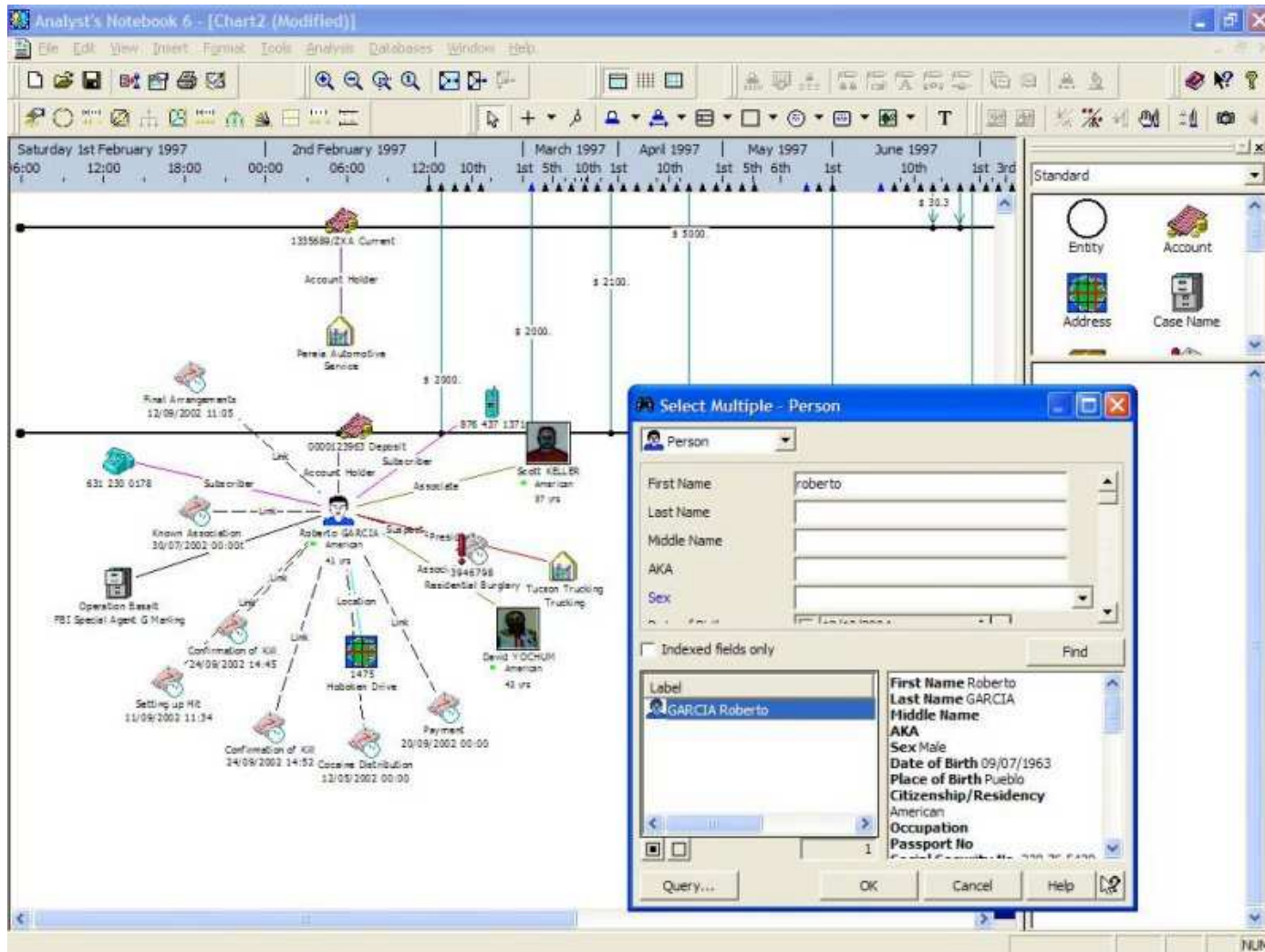
Drug Discovery (Spotfire)



Sales/Marketing Analysis (Spotfire)



Criminal Intelligence (i2)



Convergence of Search and Visualization

- Content is growing exponentially
 - Digital universe in 2007 – 281 billion gigabytes
 - 10% larger than previously thought
 - IDC 2008
- Search from one access point
 - Search across both data and text
 - Convergence of BI and Enterprise Search
- Information across many repositories
 - Subscription content
 - Enterprise repositories
 - Inconsistent meta-data
- Users don't know what information is available

Search today works if...





capital of canada

Search

[Advanced Search](#)
[Preferences](#)

Web [Books](#)

Results 1 - 10 of about 12,500,000 for capital of canada

Related searches: [canada](#) [canada's primate city](#) [cities of canada](#) [qantas](#)

Canada — Capital: Ottawa ✓

According to <https://www.cia.gov/library/publications/the-world-factbook/print/ca.html>

[Ottawa - Wikipedia, the free encyclopedia](#) ✓

There is no federal **capital** district in **Canada**. Ottawa is a municipality within the In addition to being the **capital of Canada**, Ottawa is politically ...

<en.wikipedia.org/wiki/Ottawa> - 206k - [Cached](#) - [Similar pages](#) - [Note this](#)

[Capital Cities of Canada - National Capital and Provinces Capital ...](#) ✓

The **capital** cities of **Canada** - the national **capital** and **capital** cities of the provinces and territories of **Canada**.

<canadaonline.about.com/library/bl/blcapitalcities.htm> - 25k -

[Cached](#) - [Similar pages](#) - [Note this](#)

[The City of Ottawa](#) ✓

The City of Ottawa, **Canada's Capital** · La Ville d'Ottawa, La Capitale du **Canada** · Welcome · Bienvenue.

www.ottawa.ca/ - 4k - [Cached](#) - [Similar pages](#) - [Note this](#)

[GE Capital Solutions - Homepage](#) ✓

GE **Capital** Solutions **Canada** offers commercial financing solutions for corporate, asset, franchise and vendor finance: lending products, growth **capital**, ...

www.gecapitalolutions.ca/ - 10k - [Cached](#) - [Similar pages](#) - [Note this](#)

[Welcome to Ottawa - Index](#) ✓

Ottawa is not just the **capital of Canada** - its fourth-largest city and an intriguing four-season

What if.....

Google hypertension Search [Advanced Search](#) [Preferences](#)

Web News Results 1 - 10 of about 20,600,000 for hypertension [Definition] (0.1)

Refine results for **hypertension**:

[Treatment](#) [Tests/diagnosis](#) [For patients](#) [From medical authorities](#)
[Symptoms](#) [Causes/risk factors](#) [For health professionals](#) [Alternative medicine](#)

Hypertension - Wikipedia, the free encyclopedia ✓
Hypertension, most commonly referred to as "high blood pressure", HTN or HPN, is a medical condition in which the blood pressure is chronically elevated. ...
en.wikipedia.org/wiki/Hypertension - 130k - [Cached](#) - [Similar pages](#) - [Note this](#)

Hypertension ✓
AHA Scientific Statement - Resistant **Hypertension**: Diagnosis, Evaluation, and Treatment. A Scientific Statement From the American Heart Association ...
hyper.ahajournals.org/ - [Similar pages](#) - [Note this](#)

WebMD **Hypertension**/High Blood Pressure Center - Find high blood ... ✓
High blood pressure affects 1 in 4 American adults. Find in-depth high blood pressure and **hypertension** information including its causes, symptoms, ...
www.webmd.com/hypertension-high-blood-pressure/ - 122k - [Cached](#) - [Similar pages](#) - [Note this](#)

What is **hypertension** - WebMD ✓
Get the basics about **hypertension** -- what normal blood pressure is, and what causes high blood pressure.
www.webmd.com/hypertension-high-blood-pressure/guide/blood-pressure-basics - 122k - [Cached](#) - [Similar pages](#) - [Note this](#)

MedlinePlus Medical Encyclopedia: High blood pressure (**Hypertension**) ✓

Sponsored Links

I Had **Hypertension** ✓
Now my BP is down to 120/75.
Find out how I did it without drugs
Resperate.com

Blood Pressure Too High? ✓
A Powerful Combination Treatment Proven To Significantly Lower BP.
www.EXFORGE.com

I Cured My **Hypertension** ✓
Naturally, w/o Drugs. Learn how I dropped from 170/110 to 120/75!
www.BloodPressureNormalized.com
[Google Checkout](#)

Nephrology | **Hypertension** ✓
Learn about kidney disease
Free online blood pressure log
www.ukidney.com

Hypertension ✓
Learn About High Blood Pressure Causes & Prevention at AOL® Body.
Body.AOL.com

Exploratory Search

- ❁ Needs are not clearly defined
 - Undergo a process of refinement
- ❁ Iterative Process
 - Formulate the query, examine results, view documents
 - Berry pick relevant information
- ❁ Topic is refined
 - Information needs become narrower
- ❁ Learn, discover, and understand complex topics
- ❁ Keyword searching itself is not effective
 - Great recall, low precision
 - Rows of results
- ❁ Users need to see the *context* of their requests

Visualization in Search

- Good for exploring large result sets
 - See the big picture
 - Discover patterns
 - Interact with results
- Non-linear results
 - Discovering trends/anomalies
 - Provides more entry points
- Strength is when the use of space is clear and meaningful to the user
 - Lets users see and explore the *contexts* of their search
 - Related information
- Strategies and tactics that searchers use depend on:
 - Capabilities provided by the search interface
 - User experience and task
- Needs to fit information seeking tasks

Visualization Techniques

• Results

- Birds-eye view
- Categorization
- Dynamic Clustering
 - Good for federated search

• Navigation

- Zoom
- Spatially traverse results
- Details on demand

• Explore and Refine

- Filter
- Facets
- Entity extraction

Treemaps (Hive Group)

Arrange the cameras by: **Brand** Size of the boxes represents: **Mega-pixels** Color of the boxes represents: **Lowest price** Lowest Highest No Data

1

2

3

4

5

Indicators
Highlight cameras that:

- Mac compatible
- Movie Mode
- USB output
- Video Output

Search
Show me only items with the following text:

Highlight cameras with the following phrases:

Filters
Show me only cameras that:
Are in the price range:
 All Prices

Within this mega-pixel range:
 All cameras

Sony Cybershot DSC-P1
Lowest price - \$599.50

Weight: 0.55 lbs.
Max image resolution: 2048 x 1536
Maximum number of pictures: 49
Megapixels: 3.34
Dimensions (inches): 4.5x2.1x1.7

Click mouse for options.

HIVE GROUP

Treemaps (EBSCO Visual Search)

The screenshot displays the EBSCO Visual Search interface. At the top, the EBSCO Research Databases logo is visible, along with navigation links for Sign In, Folder, Preferences, New Features!, Ask-A-Librarian, and Help. Below the logo are buttons for Basic Search, Advanced Search, Visual Search, and Choose Databases. The search bar contains the text "Find: global warming" and has Search and Clear buttons. A Language dropdown menu is also present.

The main area shows a treemap visualization of search results. The treemap is organized into three main categories: GLOBAL warming, CLIMATIC change, and UNITED States. Each category contains a grid of search results, with columns representing different search criteria and rows representing individual results. A mouse cursor is hovering over the "Blocks" option in the Display Style menu, which is currently set to "Columns".

On the right side, there is a Summary panel for the selected article. It includes the following information:

- Title:** The Security Implications of Climate Change...
- Date:** Jan 2008
- Journal:** Washington Quarterly
- Author:** Podeste, John
- Abstract:** The article discusses the possible implications of climate change for the nation's national security and for global security. It...

Treemaps (Hive Group)

Shop Amazon.com

Back

Amazon Launch Page

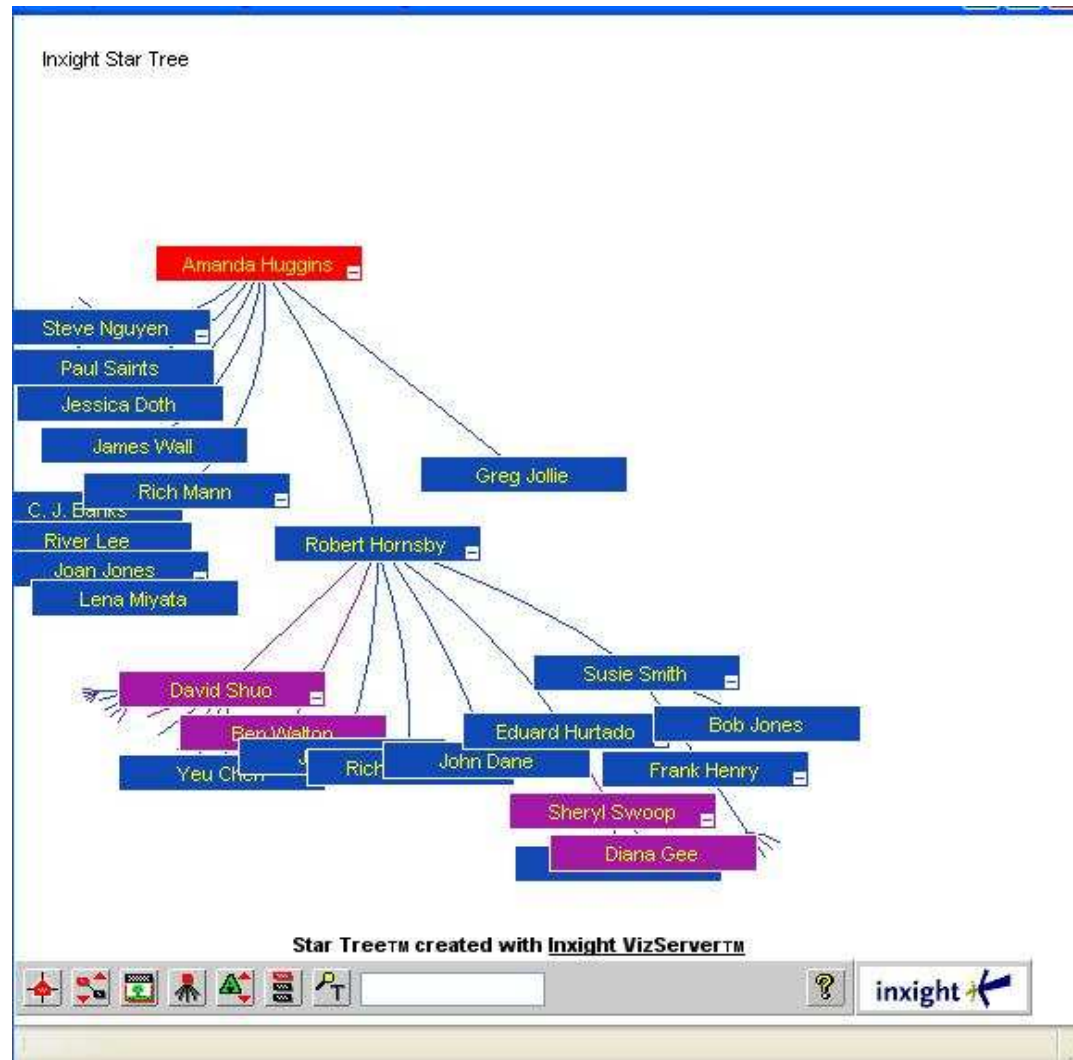
Share

Digital Camera Accessories: Digital Camera Lenses



Amazon prices and availability generated on Mon Mar 10 12:20:25 2008

Hyperbolic Trees (Inxight - StarTree)



Hyperbolic Trees (TouchGraph)

Keywords or URL: data visualization

Label Shows: Title URL

Text Over Icon

Zoom: Spacing:

Filter

Name	URL	Sim#
The R Project f...	r-proj...	2
Data Visualizati...	rese...	2
OLIVE: On-line ...	otal.u...	1
Applied Regres...	stat.u...	1
StatSoft Electro...	stats...	1
Exploring Data	explo...	1
HyperStat Onlin...	david...	1
The Data and S...	lib.st...	1
Data Visualizati...	sma...	0
GGobi data vis...	ggob...	12
Welcome to Bi...	bioco...	1
gnuplot homep...	gnup...	1
Weka 3 - Data ...	cs.w...	1
Cell Phone Fin...	myrat...	1
Statistics and ...	insig...	1
The Comprehe...	cran....	1
HCIL - Products	cs.u...	1

TouchGraph interface showing a hyperbolic tree visualization of data visualization resources. The interface includes a menu bar (Edit, View, Tools, Display), a search bar with 'data visualization' entered, and a list of results on the left. The main area displays a complex network of nodes and edges, with nodes of varying sizes and colors representing different weblogs and resources.

TouchGraph – Industry Uses

Newforth Partners (M&A Firm)

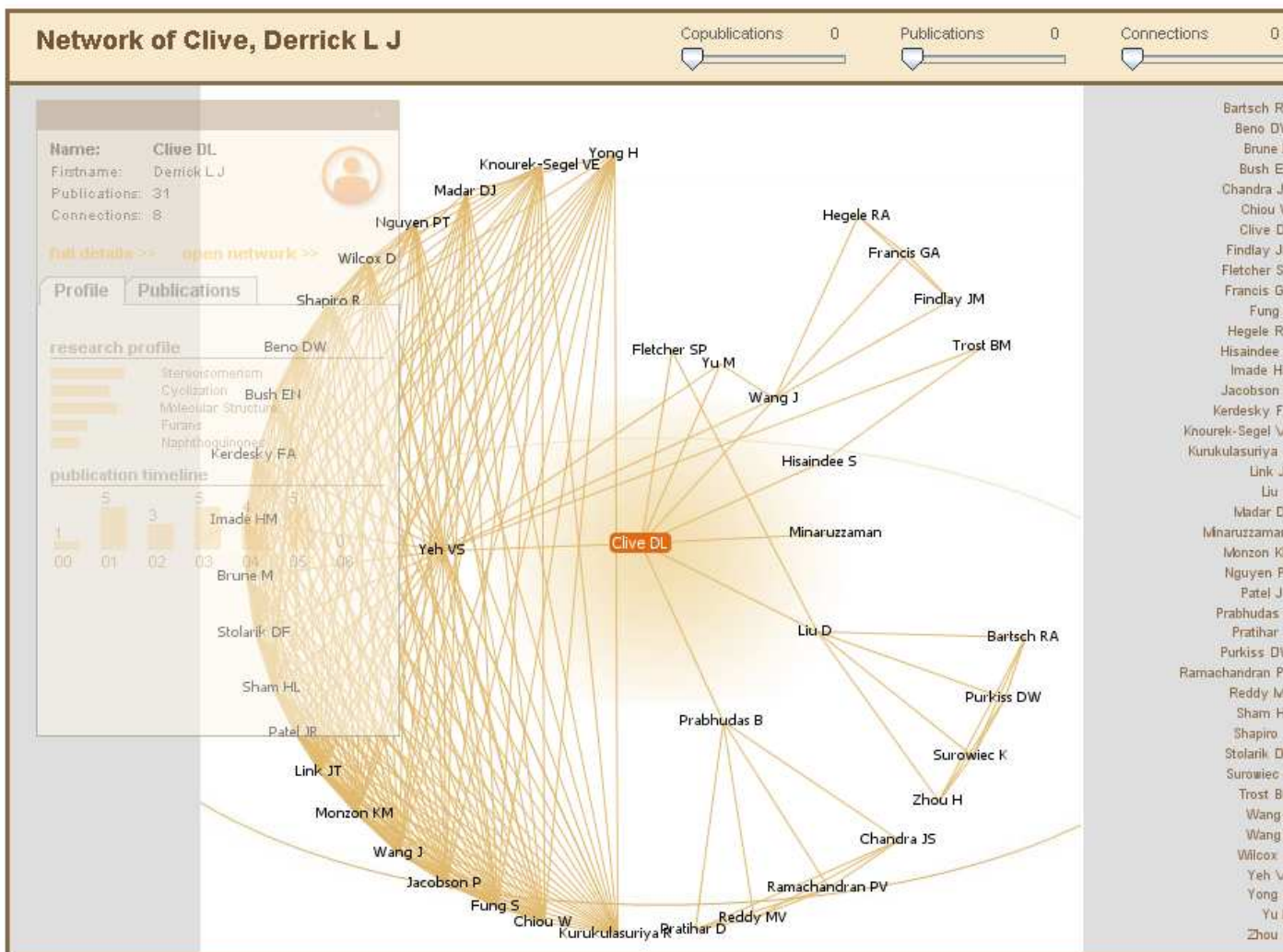
Uses TouchGraph to enhance their M&A advisory capabilities by displaying networks of merger and acquisition adjacencies formed by investment pattern and industry membership

Inter Public Group (Global Provider of Advertising & Mktg Services)

Uses TouchGraph to analyze Nielsen data (TV program popularity) by seeing relationships and clustering of shows based on similar audience preferences



Hyperbolic Trees



Kartoo

The screenshot displays the Kartoo search engine interface. At the top, a navigation bar includes tabs for 'Web', 'Images', 'Videos', and 'Wikipedia'. A search bar contains the text 'search visualization' and a 'Search' button. To the right of the search bar, it indicates '1 170 000 Found results 1 - 1'. Below the search bar, a large blue area contains a search visualization. This visualization is a network of nodes and edges. Nodes are represented by small images of search engine results pages (SERPs) and are labeled with terms like 'www.bloglines.com', 'www.programmableweb.com', 'infosthetics.com', 'www.ifs.univie.ac.at', 'www.become.com', 'www.Dealtime.com', 'www.llewellyn.com', 'www.memletics.com', 'www.keznews.com', 'dev.live.com', 'www.msdn.com', 'www.bloglines.com', 'altsearchengines.com', 'www.blogs.msdn.com', 'www.Dealtime.com', and 'www.llewellyn.com'. Edges are labeled with terms like 'guide', 'query', 'engine', 'blog', 'information', 'live brings user', 'save', 'book', 'quick research swahili released', and 'save'. On the left side, there is a sidebar with a 'Topics' list and a 'Sponsor' section. The 'Topics' list includes: tafiti search visualizati, quickapp released, brings a new user, blog, tafiti, quick, research, swahili, released, book, live, brings, user, information, save, guide, and engine. The 'Sponsor' section features three advertisements: 'Visualization for Learning' from Memletics, 'Visualizations' from Dealtime.com, and 'Visualizations' from become.com.

Topics :

- tafiti search visualizati
- quickapp released
- brings a new user
- blog
- tafiti
- quick
- research
- swahili
- released
- book
- live
- brings
- user
- information
- save
- guide
- engine

Sponsor

Visualization for Learning Learn how to use visualization for learning, and 80 more memory techniques. Memletics, a new learning system, helps you learn

Visualizations Shop & Save on Health Care Products. Store Ratings. Consumer Reviews. <http://www.Dealtime.com>

Visualizations Save Money. Find All That You Need. Save on What You Are Looking For. <http://www.become.com>

Euler Diagram (Grokker)



Selected Sources [6 of 8] [Add/Remove](#)

Annual Reviews, EMBASE, MEDLINE, Current Contents, BIOSIS, Google

[Search Options](#)

Refine search

by keyword

Exclude

by date

by source

<All sources>

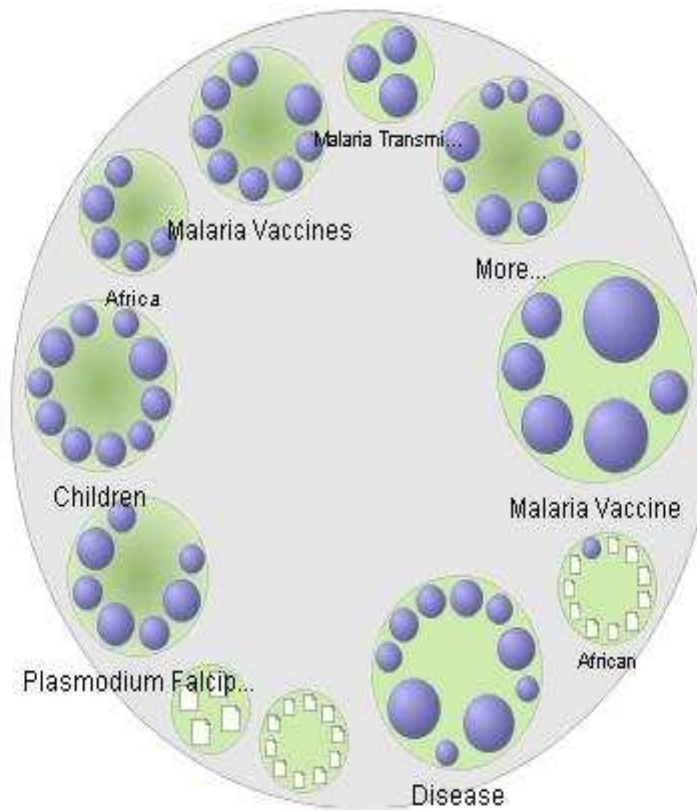
by domain

<All domains>

Layout Color

506 total results

[Expand View](#)



Results: [1-10 of 506] [«](#)
Detail: [Less](#) [Medium](#) [More](#)

[A review of human vaccin](#)

[Add to Working List](#) | [Post to c](#)
<http://gateway.ovid.com/ovidw>
Feb 1, 2007
Source: EMBASE

[Clinical diagnosis of malaria in rural health centres in E](#)

[Add to Working List](#) | [Post to c](#)
<http://gateway.ovid.com/ovidw>
Mar 1, 2008
Source: Current Contents

[Effects of environmental c Brazil](#)

[Add to Working List](#) | [Post to c](#)
<http://gateway.ovid.com/ovidw>
Jan 1, 2005
Source: BIOSIS

[Malaria - Wikipedia, the fr](#)

[Add to Working List](#) | [Post to c](#)
Hyperlinked encyclopedia artic

Grokker – Use Cases

Exploratory and Federated Search

❁ Libraries

- Federated search for patrons

❁ Publishers

- Provides research tool for subscription customers

❁ Engineering

- Federated search and visualization on proprietary content

❁ Health Insurance

- Marketing research

❁ Government

- Exploring document corpus for anomalies

Self-Organizing Maps

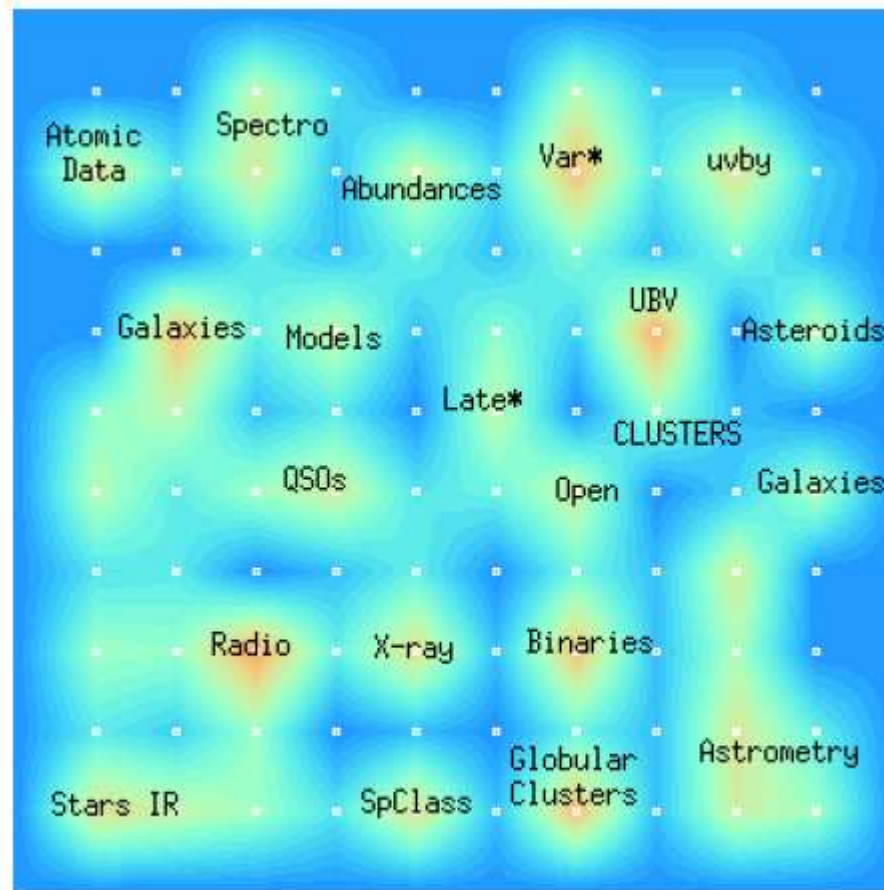
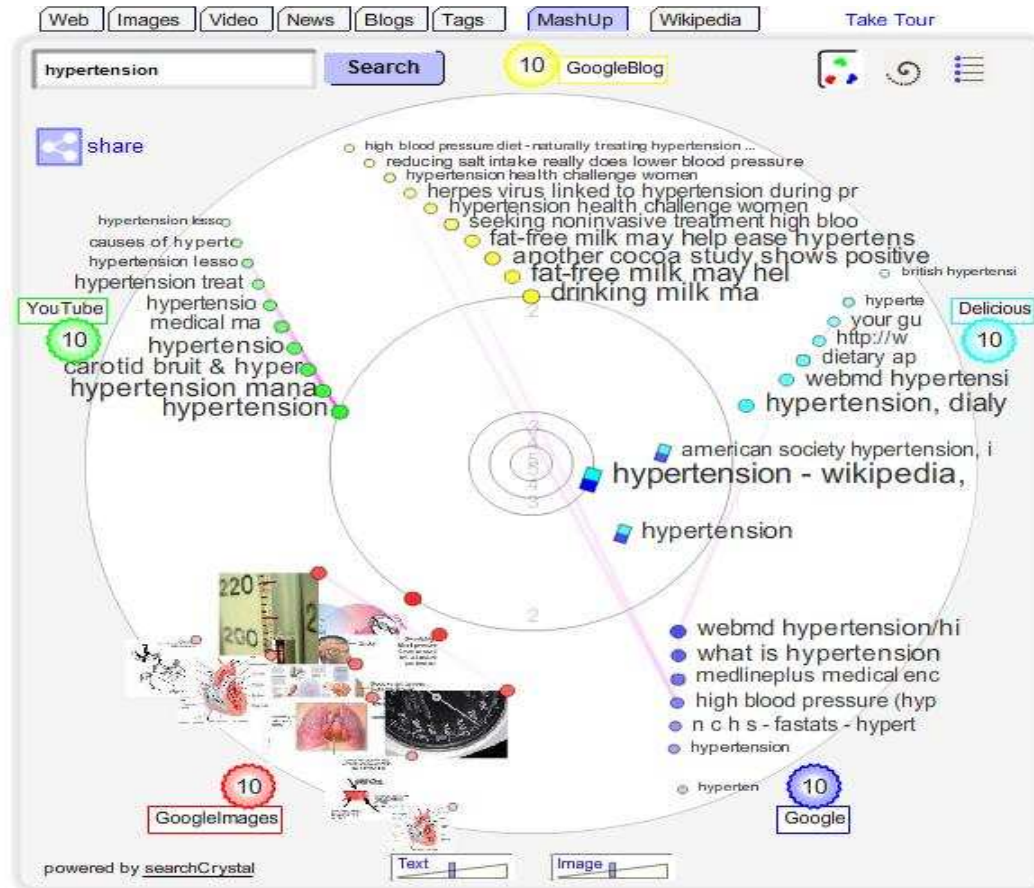


Figure 18: A SOM of search results, image taken from <http://vizier.u-strasbg.fr>

Visualization is still in its infancy...

- Visually appealing...but are they useful?
- Needs to fit the task
- Usability is critical
- Depends on strategy used for searching

SearchCrystal (meta search)



© 2007 searchCrystal | [privacy](#) | [terms of use](#)

SearchCrystal Widget

SearchMe

searchme: global warming

global warming ecology weather democratic party politicians

The screenshot shows a Wikipedia article titled "Global warming". The article text discusses the increase in average temperature of the Earth's near-surface air and oceans since the mid-twentieth century and its projected continuation. It mentions that the global average air temperature near the Earth's surface rose 0.74 ± 0.18 (± 0.22 ± 0.32 °F) during the twentieth century, with a 2005 IPCC working group report stating that the observed increase in global average temperature since the mid-twentieth century is very likely due to the observed increase in anthropogenic greenhouse gas concentrations.^[4] The article also notes that natural phenomena such as solar variation combined with volcanic activity probably had a small warming effect from pre-natural gas to 1950 and a small cooling effect from 1950 onward.^[4] These basic conclusions have been endorsed by at least fifty scientific societies and academies of science, 41 including all of the national academies of science of the major industrialized countries (NICs) where industrial scientists have worked (disagreement with some leading edge of the IPCC^[6] the overwhelming majority of scientists working on climate change agree with the scientific conclusions^{[10][11]} could mean precipitation is represented by the IPCC indicate that average global surface temperatures will likely rise a further 1.1 to 6.4 °C (2.0 to 11.5 °F) during the twenty-first century.^[12] The range of values results from the use of differing scenarios of future greenhouse gas emissions as well as models with differing climate sensitivities. Although most studies focus on the period up to 2100, warming and sea level rise are expected to continue for more than a thousand years even if greenhouse gas levels are stabilized. The delay in reaching equilibrium is a result of the large heat capacity of the oceans.^[13] Increasing global temperature will cause sea level to rise, and is expected to increase the intensity of extreme weather events and to change the amount and pattern of precipitation. Other effects of global warming include changes in agricultural yields, trade routes, glacier retreat, species extinctions and increases in the number of disease vectors. Remaining scientific uncertainties include the amount of warming expected in the future, and how warming and related changes will vary from region to region around the globe. Most national governments have signed and ratified the Kyoto Protocol aimed at reducing greenhouse gas emissions, but there is ongoing political and public debate worldwide regarding what, if any, action should be taken to reduce or reverse future warming, or to adjust to its expected consequences.

Contents (hide)

- 1 Terminology
- 2 Causes

Spacetime



oSkope

Help About oSkope Feedback

oSkope visual search

amazon

amazon.com

Choose ...

data visualization



amazon.com item

Data Visualization
2000: Proceedings of
the Joint ...

W. de Leeuw, R. van Liere
Springer, 2000
Paperback, 296 pages
Price: \$82.00 (\$89.95)
Used & new available from:
\$76.11

Click on image

1 - 100

And many more...

The screenshot displays a Digg Labs interface with a central circular visualization. The visualization is a donut chart divided into 18 segments, each representing a different user's learning style. The central text reads "Global Warming Explained, in Various Learning Styles" with "18 diggs | 0 comments" below it. The categories "SCIENCE" and "Environment" are listed above the title. The segments are labeled with usernames: BWB, Norman Brown, Tony_42, Erich100, foolfortove, asherchang, diggboss, rthead9, antdude, skoubyc, syhawk68, coffeetjnk, bakedpajamas, timmy3k, Mikhail101, theonlyjoker, Saloboy, and Chuck. The interface includes a navigation bar with "digg labs", "stack", "swarm", "bigspy", "arc", "pics", and "back to digg.com". A download link for "Arc as a screensaver" is also present, along with a note "(Adobe® Flash Player 9 required)".

Navigation: digg labs | stack | swarm | bigspy | arc | pics | back to digg.com

Download Arc as a screensaver (Adobe® Flash Player 9 required)

Global Warming Explained, in Various Learning Styles
18 diggs | 0 comments

SCIENCE
Environment

Users: BWB, Norman Brown, Tony_42, Erich100, foolfortove, asherchang, diggboss, rthead9, antdude, skoubyc, syhawk68, coffeetjnk, bakedpajamas, timmy3k, Mikhail101, theonlyjoker, Saloboy, Chuck

Visualizing the Future

- Web 2.0 influence
 - Social networks
- Visualization for many internet tools
 - Zillow (real estate)
 - Kayak (air fares)
- Video Games
 - Joysticks and other navigation techniques
 - 3D Realism
- Other Media
 - Music
 - Film

Visualizing the Future

- Tools will continue to evolve and become more intuitive
 - Focus on verticals
 - API for user interface

- More meta-data for unstructured information
 - Sentiment analysis
 - Concept extraction
 - Entity extraction

- Ease of use and implementation
 - Widgets
 - SaaS
 - Mashups

- Search to provide access to all types of information



Q & A



Explore

Discover

Understand



Thank you

Marcelline Saunders
marcelline@groxis.com

www.grokker.com