

Emerging Trends in Business Information Visualization

Vince Kellen
Vice President, Information Services
Lecturer, CTI
DePaul University

Topics we will cover

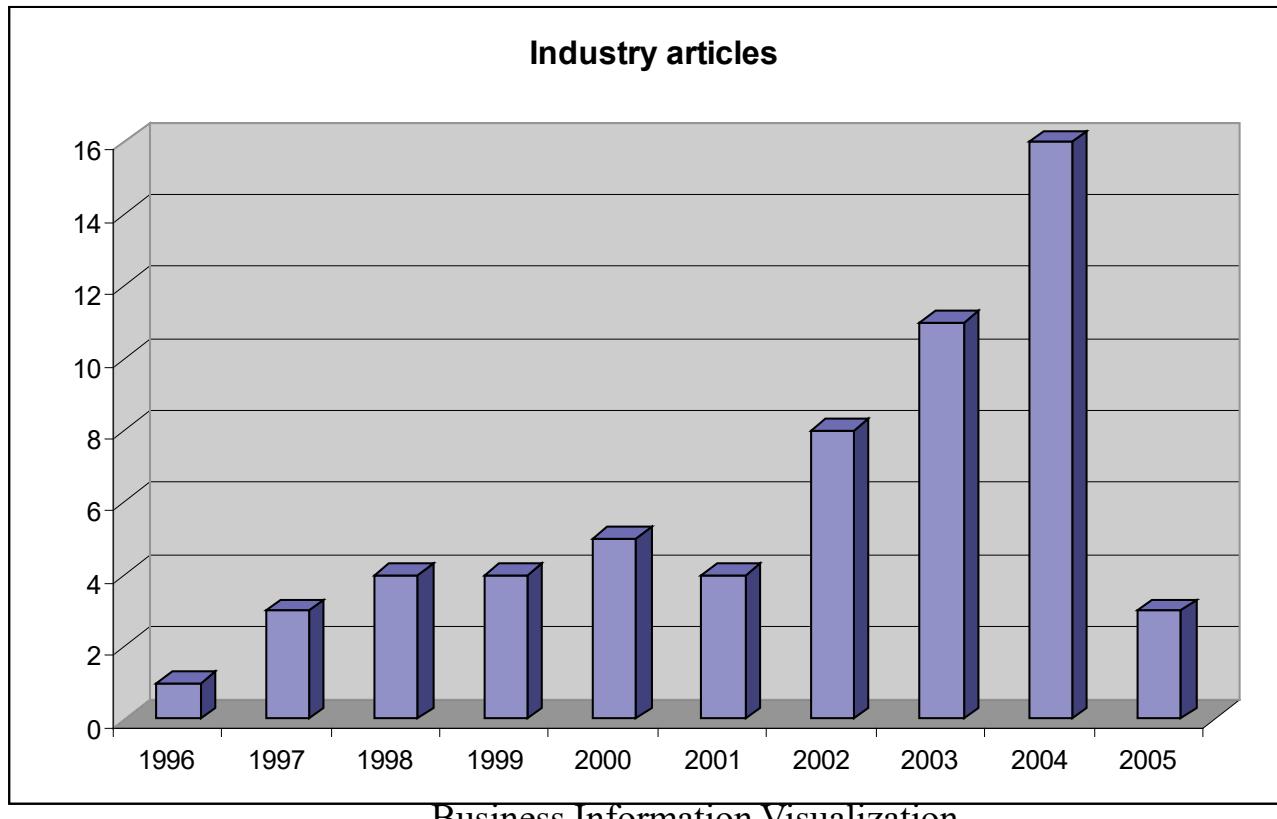
1. Introduction
2. Definitions
3. Where is information visualization useful?
4. How does information visualization help us think?
5. Key deployment challenge: user and task analysis
6. Possible adoption factors
7. Some visualization software examples
8. Visualization vendors
9. Strategic implications

Introduction

- A picture is worth, well, you know
 - <http://www.inxight.com/products/vizserver/>
 - <http://www.hivegroup.com/demos/amazon/3017951.html>
 - http://www.peets.com/selector_coffee/coffee_selector.asp?rdir=1&
 - <http://www.smartmoney.com/content/index.cfm?story=map&hpadref=1>

Visualization is emerging: industry

- Not a lot of industry press coverage compared with other software categories
 - ComputerWorld: only 7 articles with *visualization* in the title
 - CIO.com: only 8 articles with *visualization* in the body text
 - ZDNet.com: 45 articles with *visualization* in body text
- However, interest is growing



Visualization is continuing: academics

- Strong academic interest and history over the past 20 years or so
- Several books of note
 - *The Craft of Information Visualization: Readings and Reflections*, edited by Benjamin Bederson, Ben Schneiderman. Morgan Kaufmann, 2003.
 - *Readings in Information Visualization: Using Vision to Think*, edited by Stuart Card, Jock Macinlay, Ben Schneiderman. Morgan Kaufmann, 1999.
 - *Information Visualization*. Robert Spence. Addison-Wesley, 2001.
 - And of course, the three books by Edward Tufte:
 - *The Visual Display of Quantitative Information*
 - *Envisioning Information*
 - *Visual Explanations*
- Old and new interest in visualization and decision making
 - What impact does the format of information have on judgment (assessing, interpolating, extrapolating, calculating) and decision making (choosing)
 - Graphs versus tables research (accounting data, spreadsheets)
 - Heuristics and biases research (psychology, economics)

Software vendors are adapting

- Statistical software vendors
 - Visualization is a mainstay in this category
 - Various 2-D and 3-D visualizations
- Decision support vendors
 - Reporting and OLAP vendors are now adding more advanced visualization techniques that go beyond the catalog of various charts (pie, bar, line, etc.)
- Newer visualization vendors
 - Antarctica, Inxight, VoxVue
 - ThinkMap, SpotFire, Advanced Visual Systems, MindJet
 - The trend is towards server-side HTML solutions with downloadable object technology (Java typically)

Cognos Visualizer (we'll get back to this later)

COGNOS VISUALIZER: KEY FEATURES UP CLOSE

Integration with the other Cognos business intelligence components.

Report Viewer - Corporate Portfolio

The screenshot displays a dashboard titled "Business Portfolio". At the top, there is a navigation bar with tabs: "Overview", "Products", "Revenue", and "Discounts". Below the tabs, there are five circular gauges, each labeled "Revenue (x100,000)". The gauges are color-coded with red, yellow, and green segments. Below the gauges, the categories are labeled: "Home Theatre", "Consumer Electronics", "Entertainment Media", "Home Office", and "Lifestyle Products". To the right of the gauges, a callout points to the gauges with the text "Advanced charting for multi-metric displays". At the bottom of the dashboard is a map of the United States where states are colored according to their revenue performance. A callout from the map points to the text "Interactive maps deliver instant impact visuals".

Business Portfolio

Overview Products Revenue Discounts

Revenue (x100,000) Revenue (x100,000) Revenue (x100,000) Revenue (x100,000) Revenue (x100,000)

Home Theatre Consumer Electronics Entertainment Media Home Office Lifestyle Products

United States

Advanced charting for multi-metric displays.

Interactive maps deliver instant impact visuals.

Definitions

- Visualization
 - To form a clear mental image, vision or picture of something
 - A graphical representation of data or concepts
- Information representation
 - How information is displayed, depicted, conveyed
 - Three ways to convey information
 - Language symbols: numbers, words, icons
 - Visual symbols or marks: no direct mapping to language
 - Mixed displays using language and visual symbols
- External representations
 - Information conveyed using an artifact of some kind
- Internal representations
 - Information maintained in the human mind

Levels of analysis

- Researchers examine visualization at two levels
 - Low-level
 - How does the eye function; how does the brain manage visual information; how does visualization interact with consciousness? Which neural processes are parallel? Which are serial?
 - High level
 - What impact does visualization have on task speed, decision or judgment accuracy, user satisfaction, user confidence; how do differences in cognitive abilities affect the use of visual representations; what is the interaction between emotions and visual representations?

Low-level tidbits

- The eye and the visual cortex of the brain form a massively parallel processor that provides the highest bandwidth channel into human cognitive centers*
- Our brains process much (most?) visual information without being aware of it (preattentive processing)
- These preattentive processing centers interact with the slower, language/logic-centered, processes in subtle ways
- In other words, lots of stuff is going on making it difficult to build a low-level, much less a high level model of how visualization and cognition interact

*Information Visualization: Perception for Design, Colin Ware, Morgan Kaufmann, 2000.

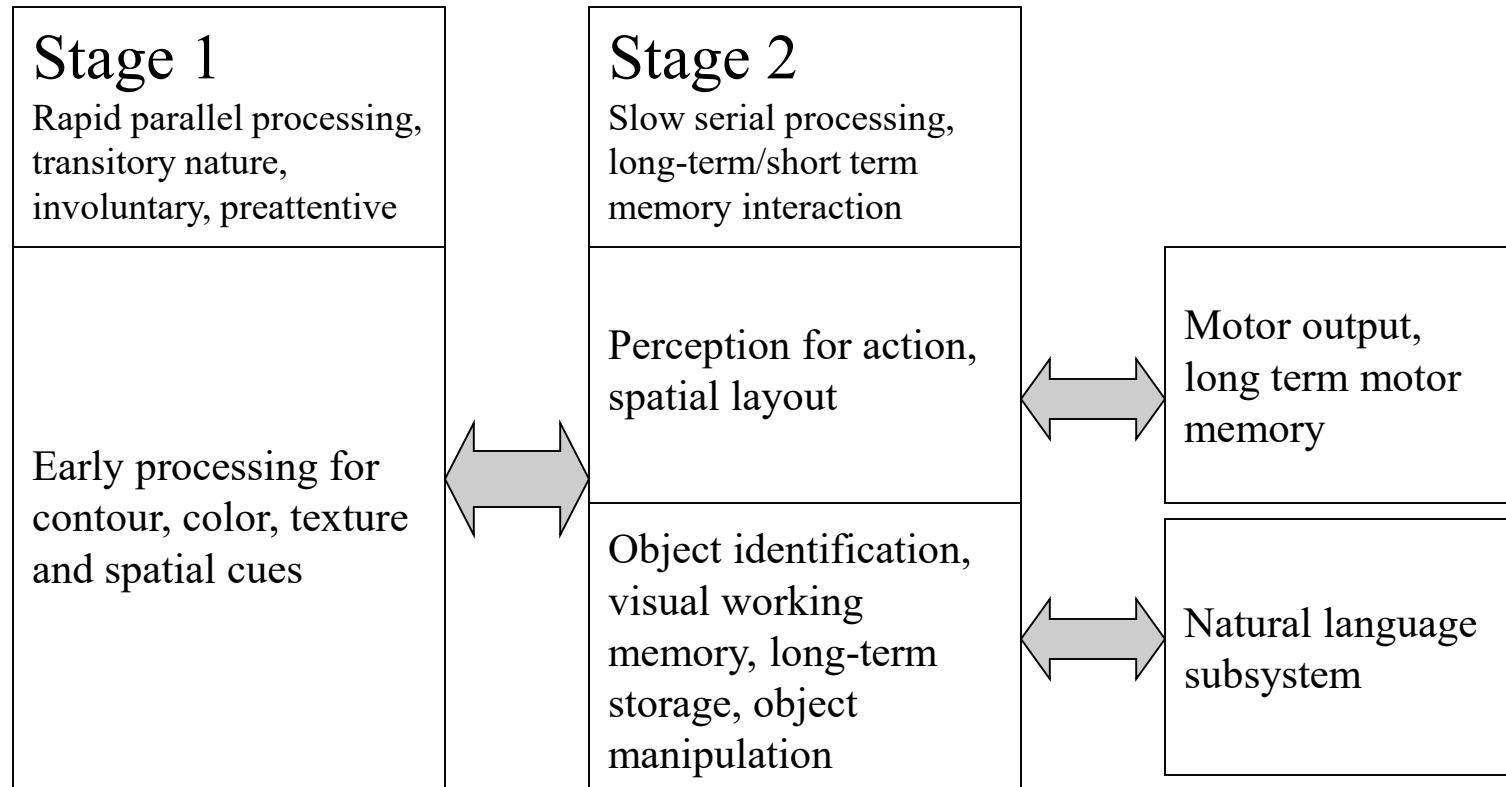
Categories of information visualization

- Naturalistic or physical data
 - Data that has a relatively direct mapping to a two or three dimensional real-life representation
 - Scientific data
 - E.g., molecules, DNA, biological or man-made structures
 - Geographical data
 - E.g., sales by region, weather patterns, location of objects in physical space
- Abstract data
 - Data that has no direct mapping to a two or three dimensional real-life representation, data that has an arbitrary ordering of attributes
 - Lots of business data is abstract
 - E.g., Sales by product, quarter, division, sales person

Sensory versus arbitrary visualizations

- Sensory symbols require no learning, are stable across cultures, time and individuals, are well matched to early stages of neural processing
 - E.g., a cave drawing of a hunt
- Arbitrary visualizations must be learned, are easy to forget, are embedded in culture and applications, can be formally powerful (e.g., visual design languages-UML), can evolve rapidly (relative to sensory symbols)
 - E.g., a business process diagram
- Most visualizations fall along the continuum between sensory and arbitrary
- Our brains are “hardwired” for certain visual patterns (e.g., faces) and these hardwired networks interact with learned visual conventions

Visual processing



Information Visualization: Perception for Design, Colin Ware, Morgan Kaufmann, 2000

See also: [Dual processing theory](#)

How visualization aids in processing data

Parallel perceptual processing	Some attributes of visualizations can be processed in parallel compared to text.
Offload work from cognitive to perceptual system	Some cognitive inferences done symbolically can be recoded into inferences done with simple perceptual operations.
Expanded working memory	Visualizations can expand the working memory available for solving a problem.
Expanded storage of information	Visualizations can be used to store massive amounts of information in a quickly accessible form (e.g., maps).
Locality of processing	Visualizations group information used together, reducing searching.
High data intensity	Visualizations can often present a large amount of data in a small space.
Spatially indexed addressing	By grouping data about an object, visualizations can avoid symbolic labels.
Recognition instead of recall	Recognizing information generated by a visualization is easier than recalling that information by that user.
Abstraction and aggregation	Visualizations simplify and organize information, supplying higher centers with aggregated forms of information through abstraction and selective omission.
Visual representations make some problem obvious	Visualizations can support a large number of perceptual inferences that are extremely easy for humans.
Perceptual monitoring	Visualizations can allow for the monitoring of a large number of potential events if the display is organized so that these stand out by appearance or motion.
Manipulable medium	Unlike static diagrams, visualizations can allow exploration of a space of parameter values and can amplify user operations.

Card, Stuart K., Mackinlay, Jock D. & Shneiderman, Ben. (1999). Readings in Information Visualization: Using Vision to Think. Academic Press.

Visualization and metaphor

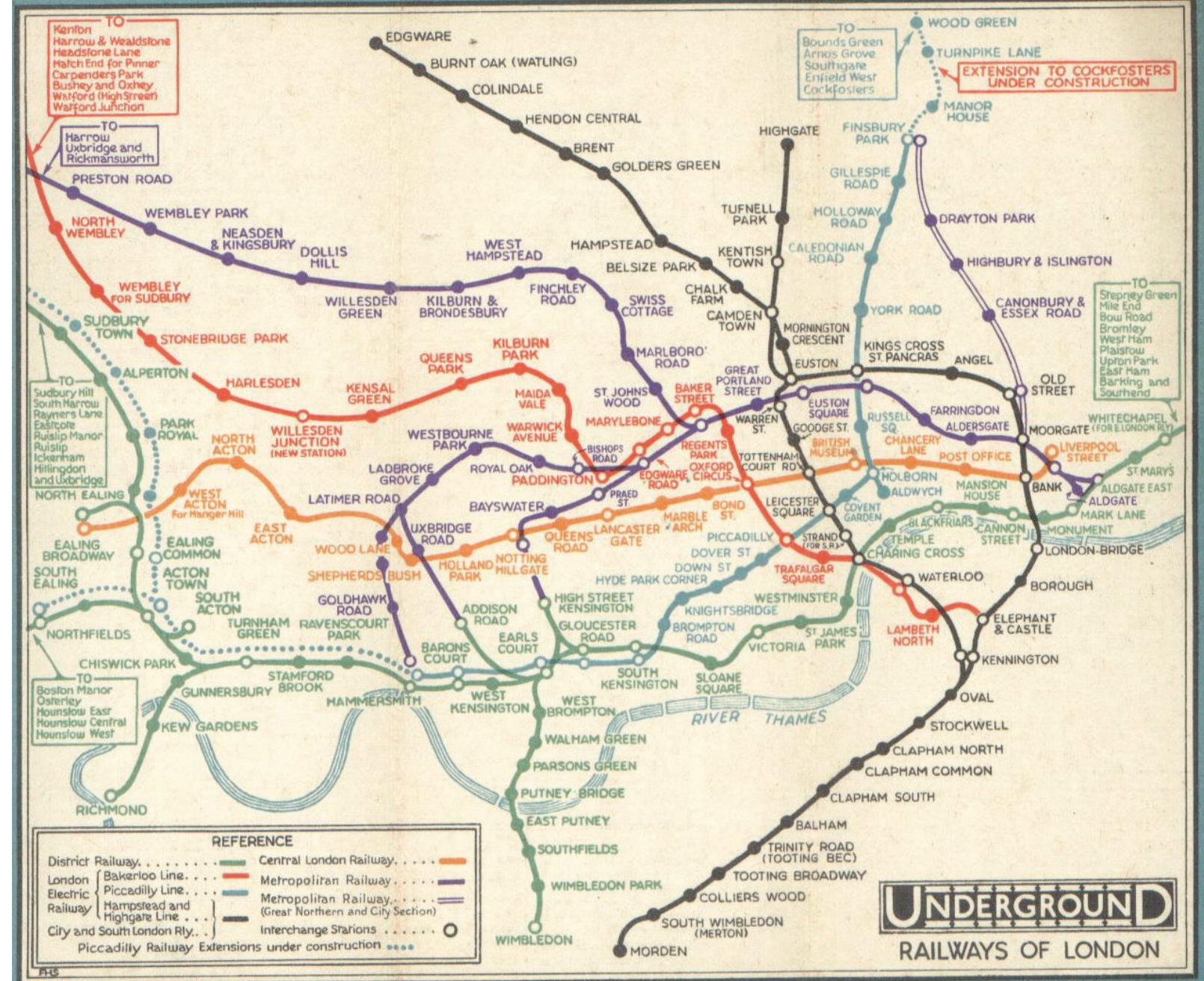
- The conversion of data into personal knowledge can be aided greatly using metaphor and visualization
- Is all cognition grounded in perceptual and performance (kinesthetic) schemas rather than arbitrary symbolic code?

<i>Metaphor</i>	<i>Abstract Schema</i>	<i>Concrete Schema</i>
Important is big	Significance in situation	Seen/felt size
Difficulties are burdens	Obstacle to intention	Felt weight
More is up	Quantity or degree	Seen or felt elevation
Categories are containers	Classification	Seen or felt containment
Similarity is closeness	Diagnostic/predictive similarity	Seen or felt proximity
Help is support	Assistance toward intention	Felt firmness underneath
Time is motion	Passage of time	Seen/felt movement
States are locations	Situational equivalence	Seen/felt place
Change is motion	Variation over time	Seen/felt movement
Action is self-propulsion	Autonomous activity	Intentional movement
Purposes are desired objects	Intention	Reinforcing object
Causes are physical forces	Causes and origins	Felt pressure and weight
Relationships are enclosures	Relational dependency	Seen/felt enclosure
Controlling is being above	Causal dependency	Vertical alignment
Seeing is understanding	Knowledge	Objects seen
Understanding is grasping	Knowledge and comprehension	Objects actively felt

Lakoff, Mark & Johnson, George. (1999) Philosophy in the Flesh: The Embodied Mind and its Challenges to Western Thought. Basic Books.

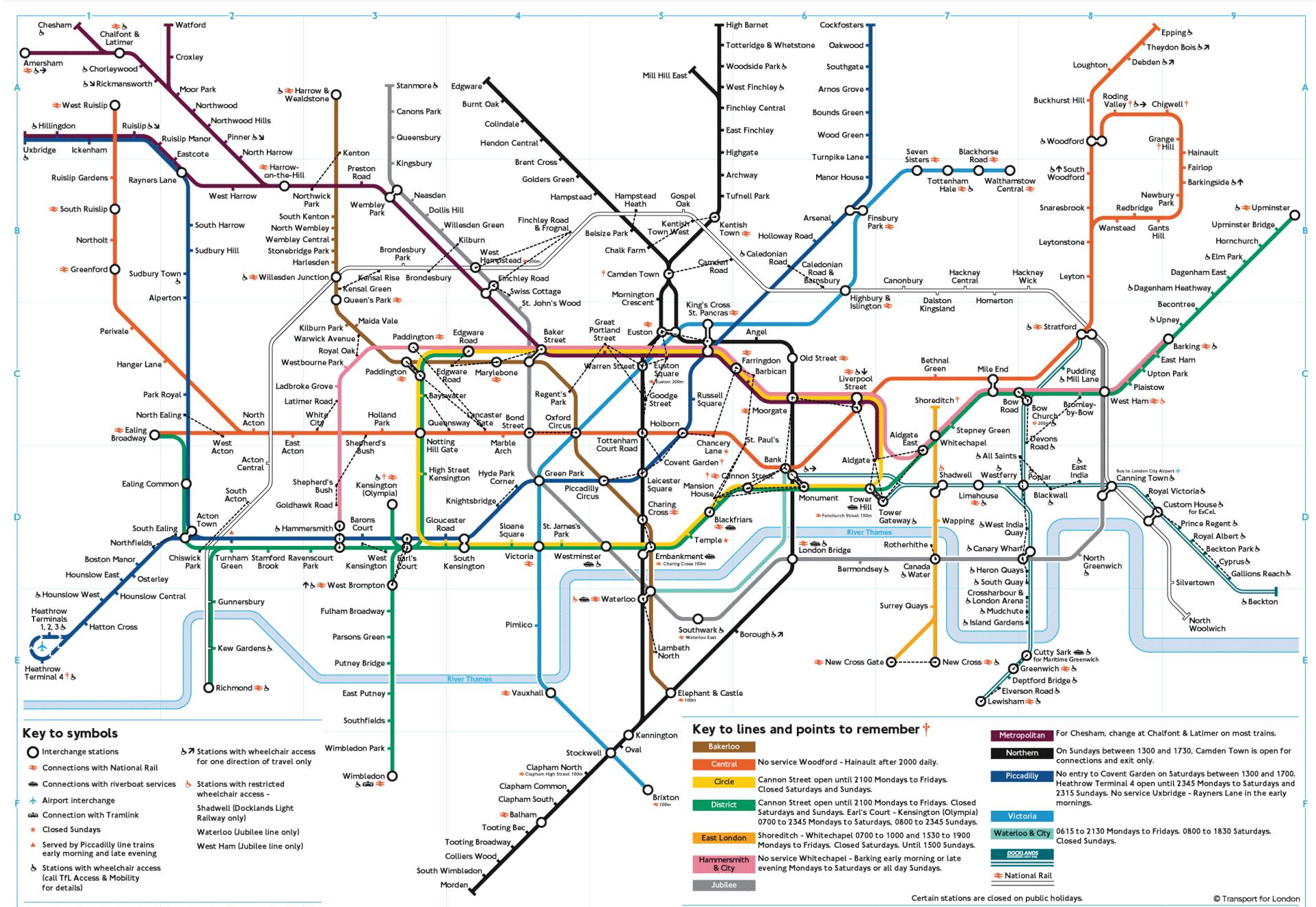
A geographical information example

1932 London Tube Map



Harry Beck's 1933 London Tube Map



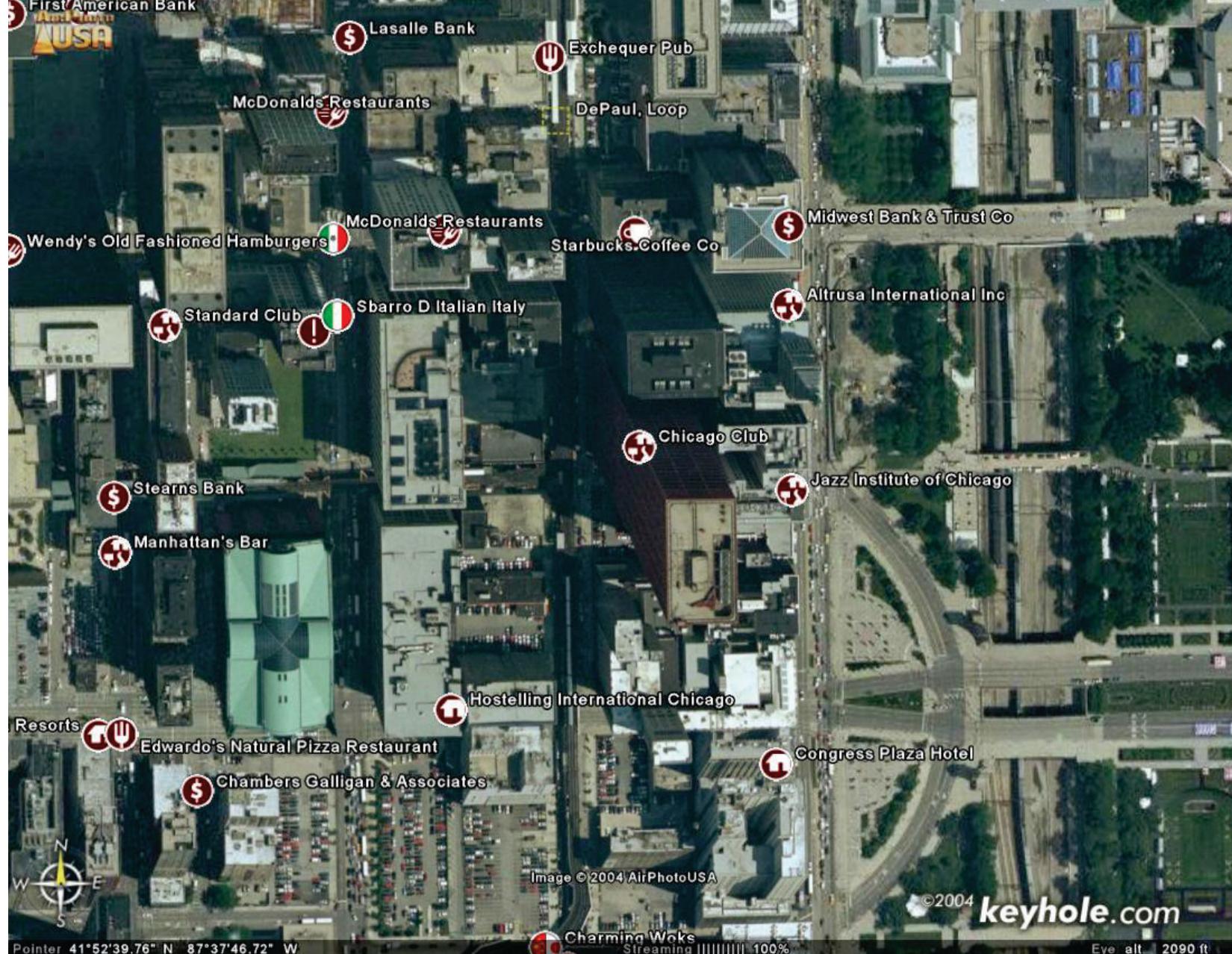


Where is IV useful?

- Searching for information
 - Save time by finding patterns faster
- Problem modeling
 - Find a suitable representation for a problem by manipulating the display
- Persuading
 - Explaining or convincing others
- Decision making
 - Combine information and render for quicker action
 - Reduce errors in judgment and decision making

Searching for information

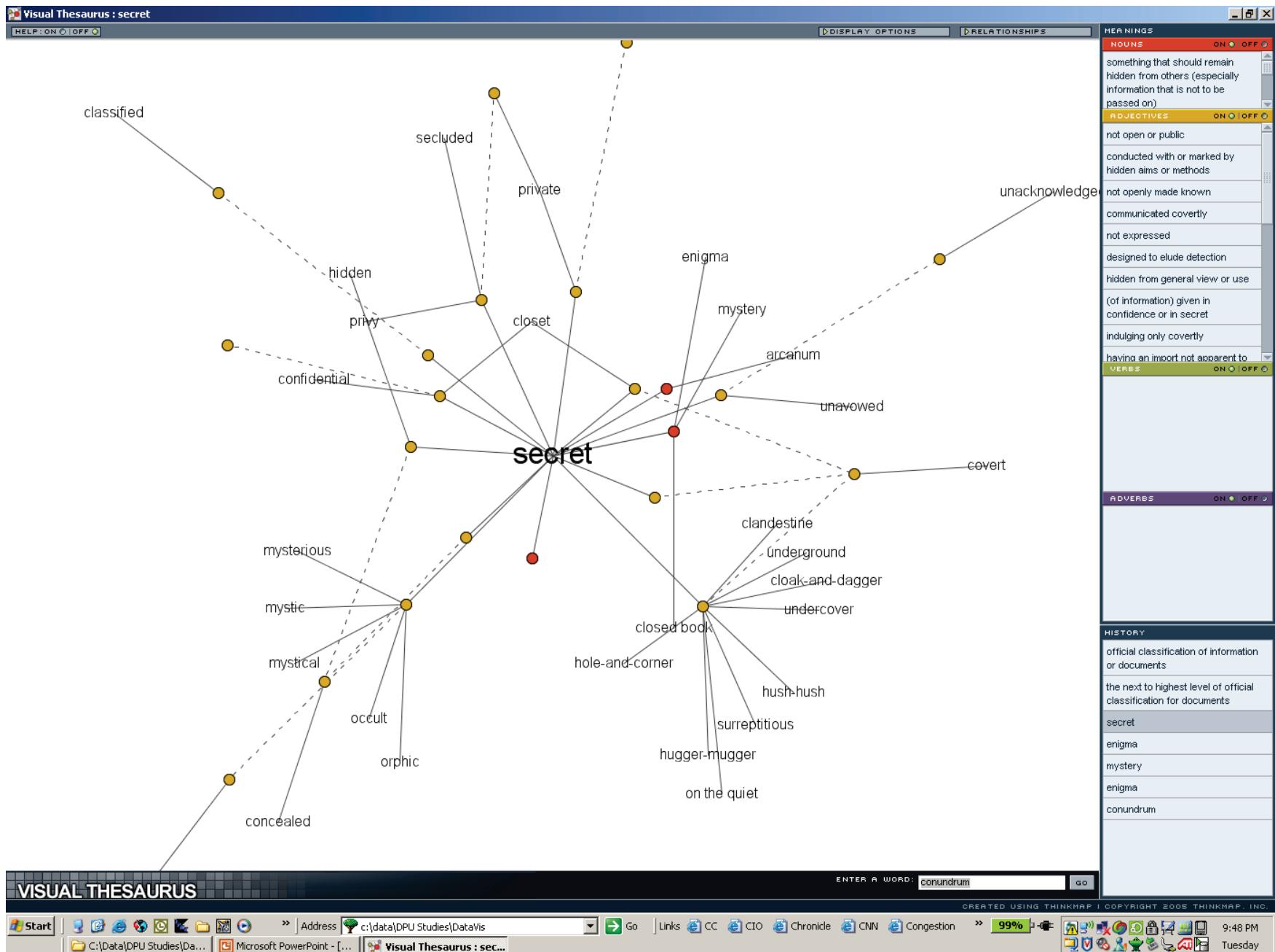
- Keyhole
- Largest database of commercially available satellite imagery
- Keyhole was recently bought by Google
- The software dynamically retrieves images from the Keyhole server and renders images incrementally on the client
- Some images are 3-D



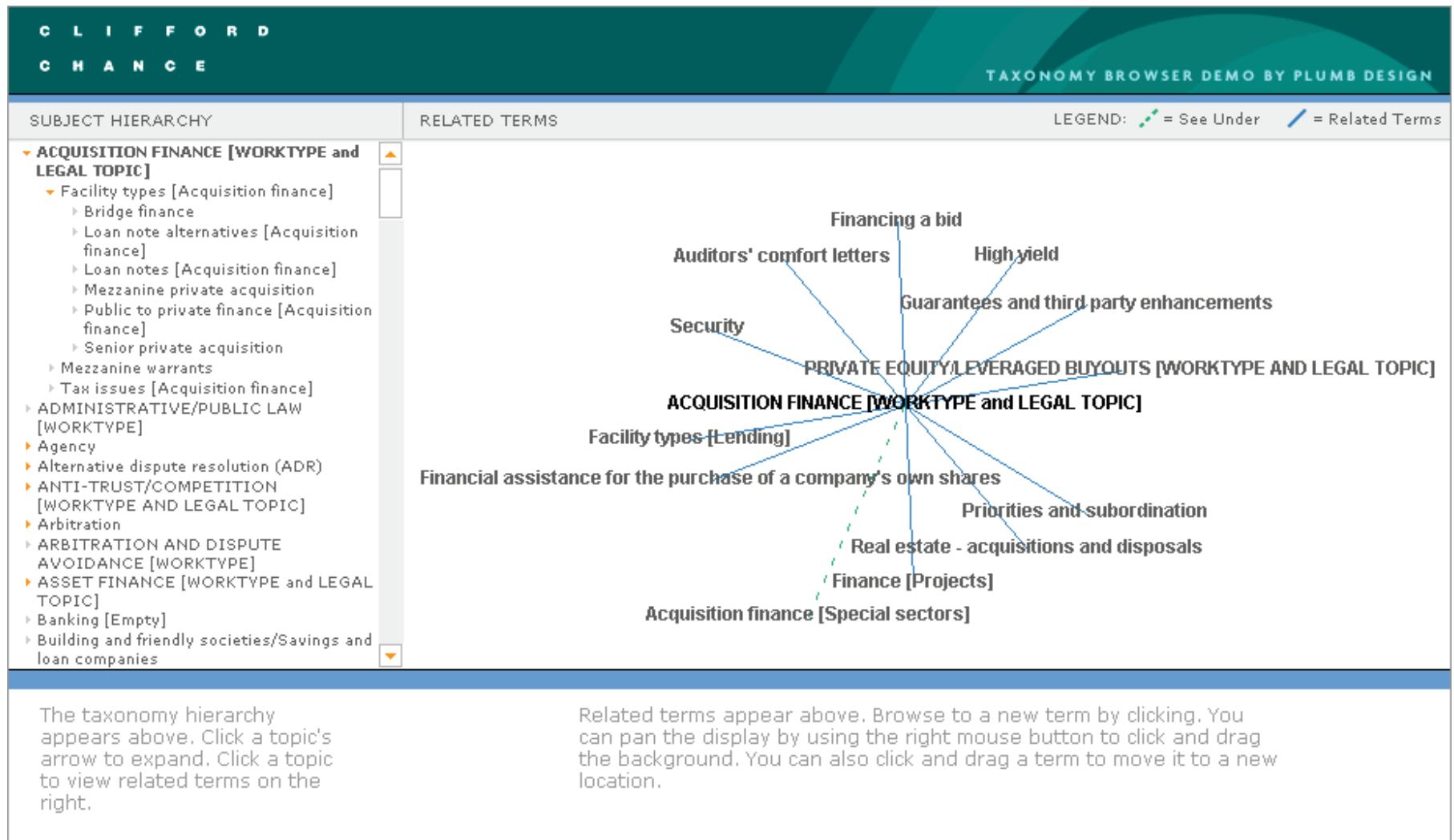
The Loop Campus, according to Keyhole

Visual Thesaurus

- Product offered by ThinkMap
- Find a similar or related word quickly
- The product is actually a framework for creating similar solutions
- Can be useful for documenting taxonomies, ontologies



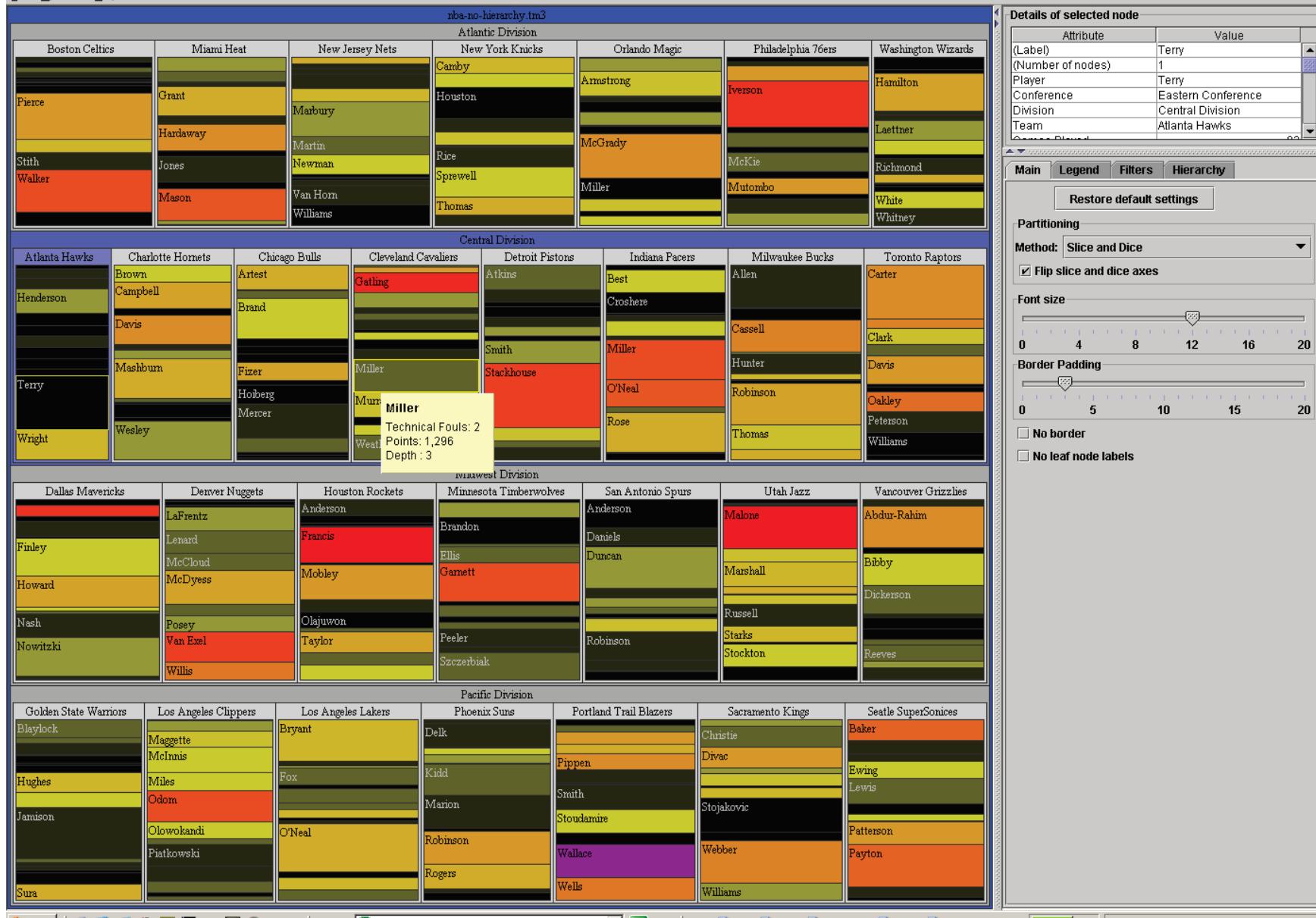
ThinkMap SDK Example



http://www.thinkmap.com/casestudies_clifford_chance.jsp

TreeMaps

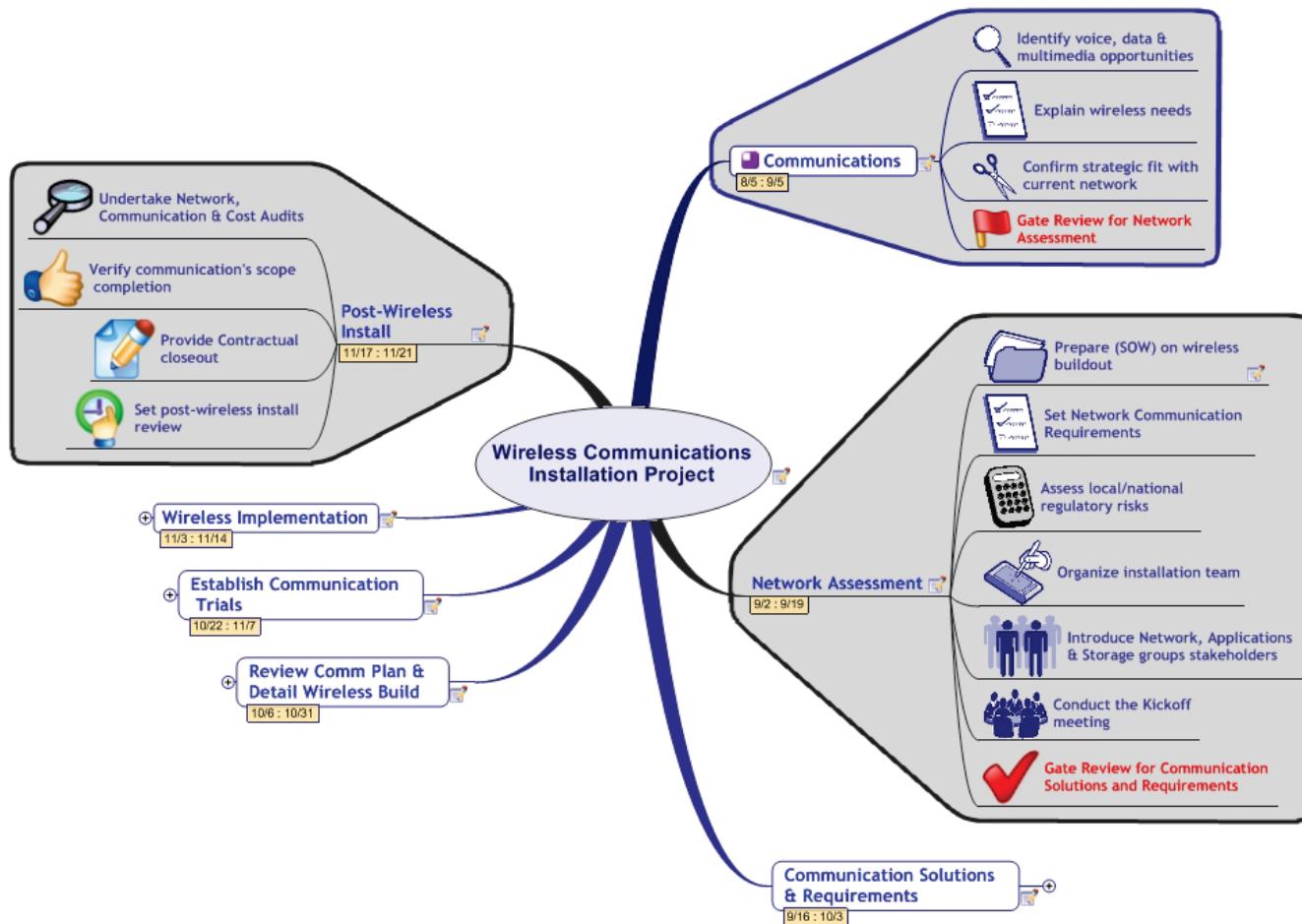
- TreeMaps let users browse through hierarchies of data
- Filtering, brushing
- Drill down, up
- TreeMap demo



Problem modeling

- MindManager
 - Free form tool for linking concepts together, annotating concepts with text and images
 - Useful for quick and easy modeling
 - You can get complex and intricate if needed
 - Not a formal modeling tool
 - Snap-in for Groove (for collaborative modeling)
 - FYI: Groove was recently bought by Microsoft. Watch this.
- MindManager Example

Project concept according to MindManager



© 2003 Mindjet

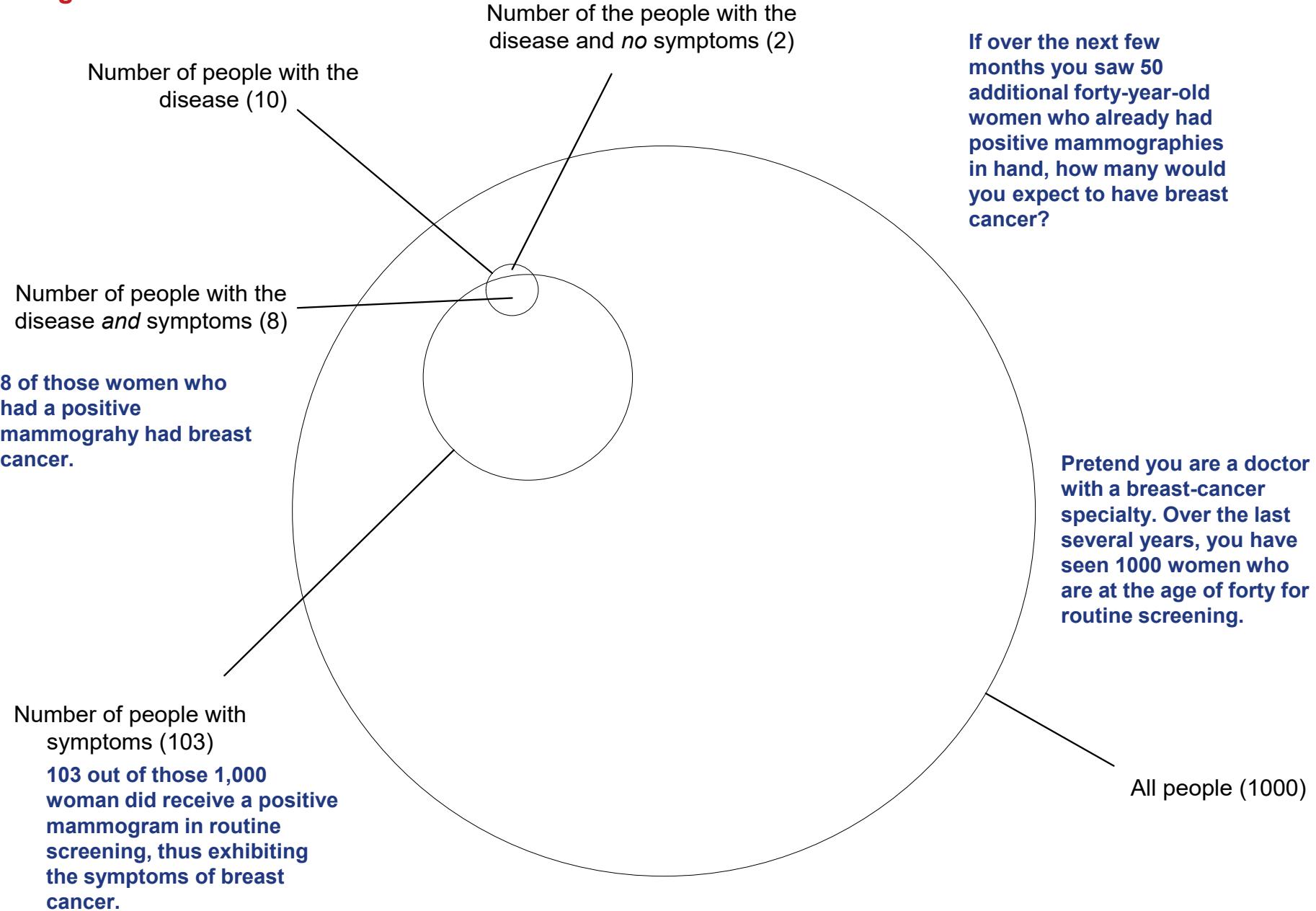
Judgment and decision making

- People make errors in judgment and choice
 - People consistently exhibit biases
 - Heuristics we use results in biases; we are cognitively limited
 - Ambiguity also results in biases; we often latch onto the wrong “rules”
- Reducing judgment and decision bias
 - Visualization crops up as one of several debiasing techniques

Pop quiz: Bayesian inference problem

The probability of breast cancer is 1% for women at age forty who participate in routine screening. If a woman has breast cancer, the probability is 80% that she will get a positive mammography. If a woman does not have breast cancer, the probability is 9.6% that she will also get a positive mammography. A woman in this age group had a positive mammography in a routine screening. What is the probability that she actually has breast cancer?

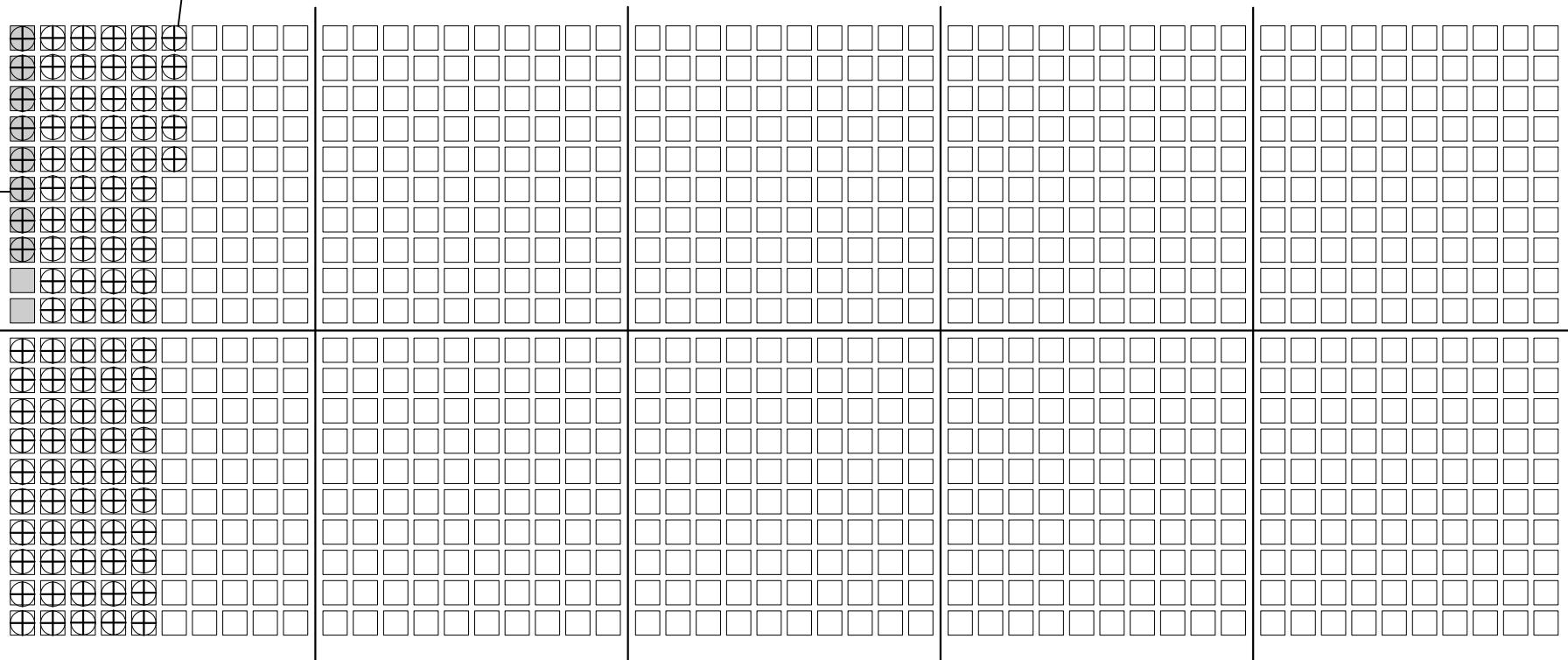
Begin



103 out of those 1,000 woman did receive a positive mammogram in routine screening, thus exhibiting the symptoms of breast cancer

Pretend you are a doctor with a breast-cancer specialty. Over the last several years, you have seen 1000 women who are at the age of forty for routine screening.

Begin



8 of those women who had a positive mammography had breast cancer.

Question:

If you over the next few months you saw 50 additional forty-year-old women who had positive mammographies in hand, how many would you expect to have breast cancer?

Breast cancer problem

- What are the categories?
 - All women at age 40 who get screen regularly
 - Women with a positive test result
 - Women with cancer
 - Women with cancer *and* a positive test result
 - Women with cancer *without* a positive test result
- What is the relationship between categories?
 - Mostly hierarchical, with two sets that overlap

What did the visualization do?

- Basic marks and symbols were processed preattentively (circles, lines)
- You probably examined the text and the visuals to build a mapping (mental model, internal representation)
- In building this mental model, preattentive processing probably informed the symbolic/language centers and *constrained* the development of multiple solutions. In other words, the visualization suggested a strategy or model which you may have nonconsciously used, thus suppressing the generation of alternative strategies
- The rule processing part of you probably reviewed and tried to validate the result, looking for incongruities
- Lack of familiarity with rules for solving this may have left you with little confidence in your answer
 - Like filling out a 1040 form!

Motivation and visualization

- Motivation plays a role in problems like this. Where there is extremely high motivation, subjects will usually leverage their symbolic/rule processing cognitive skills, reducing the importance and impact of the visuals
 - Give me the right answer or I'll reduce your pay by 50%
- When motivation is low, subjects will usually not expend the effort to solve the problem and will most likely accept the visual representation at hand
- Conclusion
 - Visualization will be more likely to correct or create bias for lower-motivation problems. Typical day-to-day, faster, apparently lower-impact decisions can be affected
 - Strategic decisions that have time, involve groups, oversight, devil's advocates, significant gain or loss will be less influenced by visualizations
 - Visualizations can aid in modeling a problem and bringing staff up to speed on the problem
 - Visualization may actually prevent the generation of alternatives, which can be good or bad

Individual attributes and visualization

- Some people have a preference for and greater ability to process visual information. It is not clear how and when this will affect problem solving
- Some people with a strong visual preference or skill solve the visual problem more easily (less time), some over-search the problem and do not
- Research on visual or spatial cognitive abilities and use of visual aids is mixed, probably because different cognitive capabilities are intertwined
- Conclusion
 - For this and purely economic reasons, visualization systems need to be stable across individuals or should be used with non-visual tools so that the right tool can be paired with the right person

Dual processing theory

System 1	System 2
Associative	Rule-based
Holistic	Analytic
Automatic	Controlled
Relatively undemanding of cognitive capacity	Demanding of cognitive capacity
Relatively fast	Relatively slow
Highly contextualized task construal	Decontextualized task construal
Conversational and socialized task construal	Asocial task construal

Stanovich, K. E. & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*. 23, 645-726.

Dual processing theory and visual cognition

- Some similarities exist
 - One system is serial and slow in processing
 - One system is fast
- Dual processing explanations often focus on difference between the hemispheres of the brain
- Implications
 - The interplay between fast, contextual processing and slower, rule-based processing is complex!
 - Visualizations will engage system 1
 - Well structured visualizations should speed problem solving, facilitate more accurate solutions

Visualization and decision making

- From a managerial perspective, things get interesting when you consider visualization and decision making
- How many managers forgo information search due to its sometimes unnecessarily high effort?
- How many wrong or suboptimal decisions do managers make because of bad information representation?
- How many managers deliberately bias or spin information representation with clever visualizations?
- Can you save time or generate better decisions using visualization?

Problems with abstract data

- Abstract data has no specific order in its attributes
- Attributes are only arbitrarily mapped to a physical 2-D or 3-D coordinate system
- Multiple relationships between pieces of data may exist
 - Multiple hierarchies
 - Different set relationships
- Abstract data mark or symbol has limited target attributes
 - 6 basic attributes: Length, width, color, texture, position, movement
 - The first three are most critical: size and color

Abstract data

- Division *Which attribute should drive length/width?*
- General ledger code
- General ledger sub code *Which one should drive color?*
- Accounting period
- Fund code *Which one should drive position?*
- Project code
- Budgeted
- Actual
- Remaining *How would you “naturalize” this?*

Process for dealing with abstract data

1. What is the benefit?

- How will the tool affect search time? Decision speed? Decision quality?
How will improving any of these lead to business benefit?

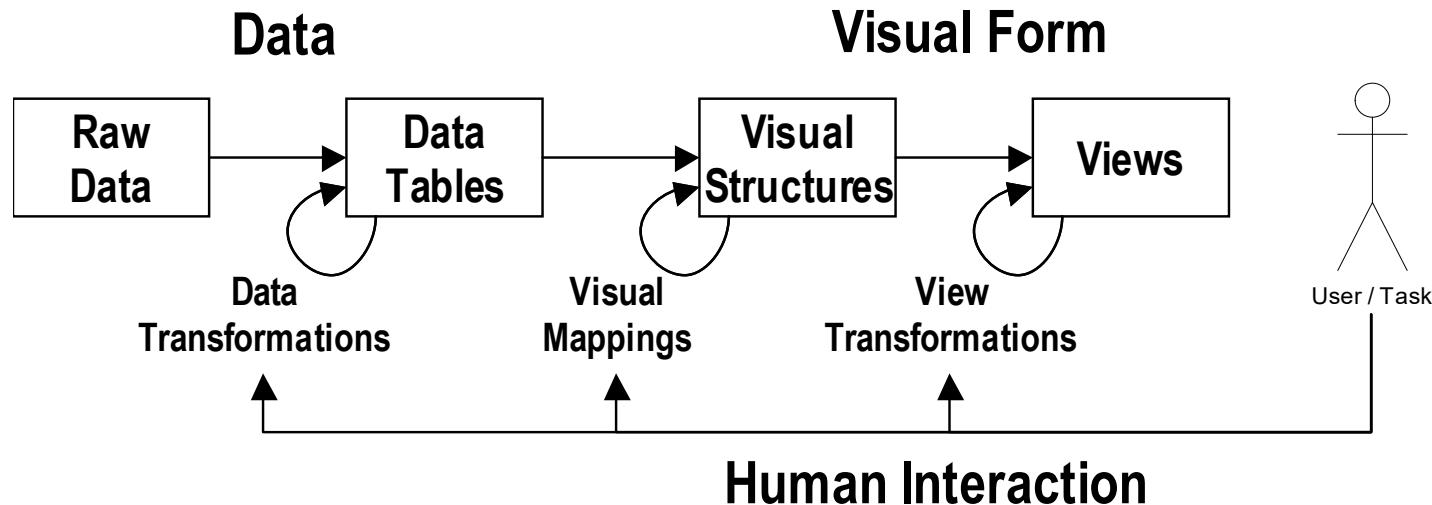
2. What is the specific task?

- How should the task be mapped to a visualization?
 - Understand the concepts in the problem
 - Understand the relationships between the concepts (hierarchical, network, overlapping sets, disjoint sets, etc.)
 - Assess the match between a tool and the problem
 - Select the top the attributes to drive the visualization mark/symbols
 - Identify the other attributes/symbol pairings

3. Iterate until done

- Rinse, repeat until the visualization gains user satisfaction or improves problem performance (or both)

A process for visualization?



Raw data: idiosyncratic formats

Data tables: relations (cases by variables) + metadata

Visual structures: spatial substrates + marks + graphical properties

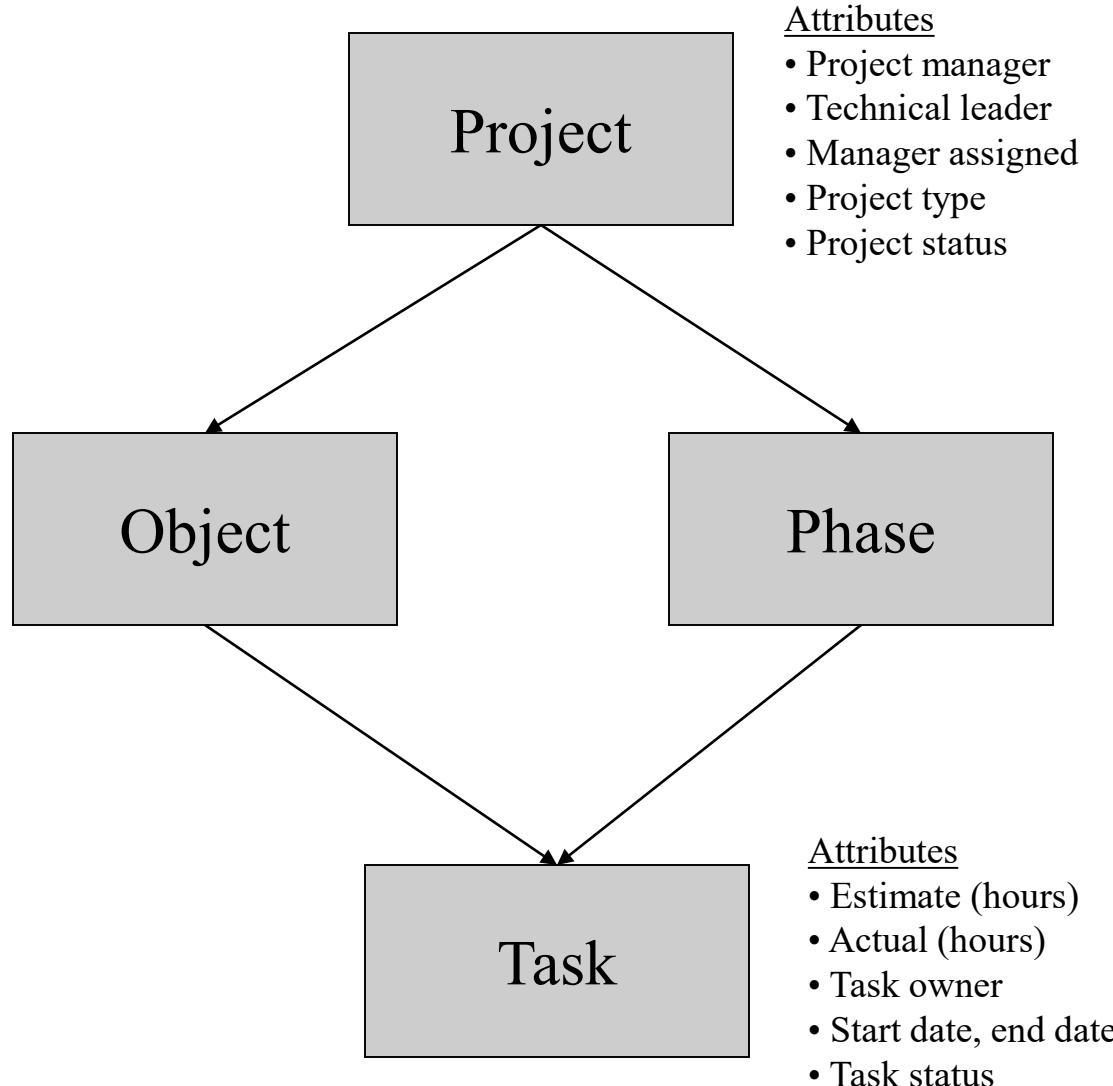
Views: graphical parameters (position, scaling, clipping, ...)

Card, Stuart K., Mackinlay, Jock D. & Shneiderman, Ben. (1999). Readings in Information Visualization: Using Vision to Think. Academic Press.

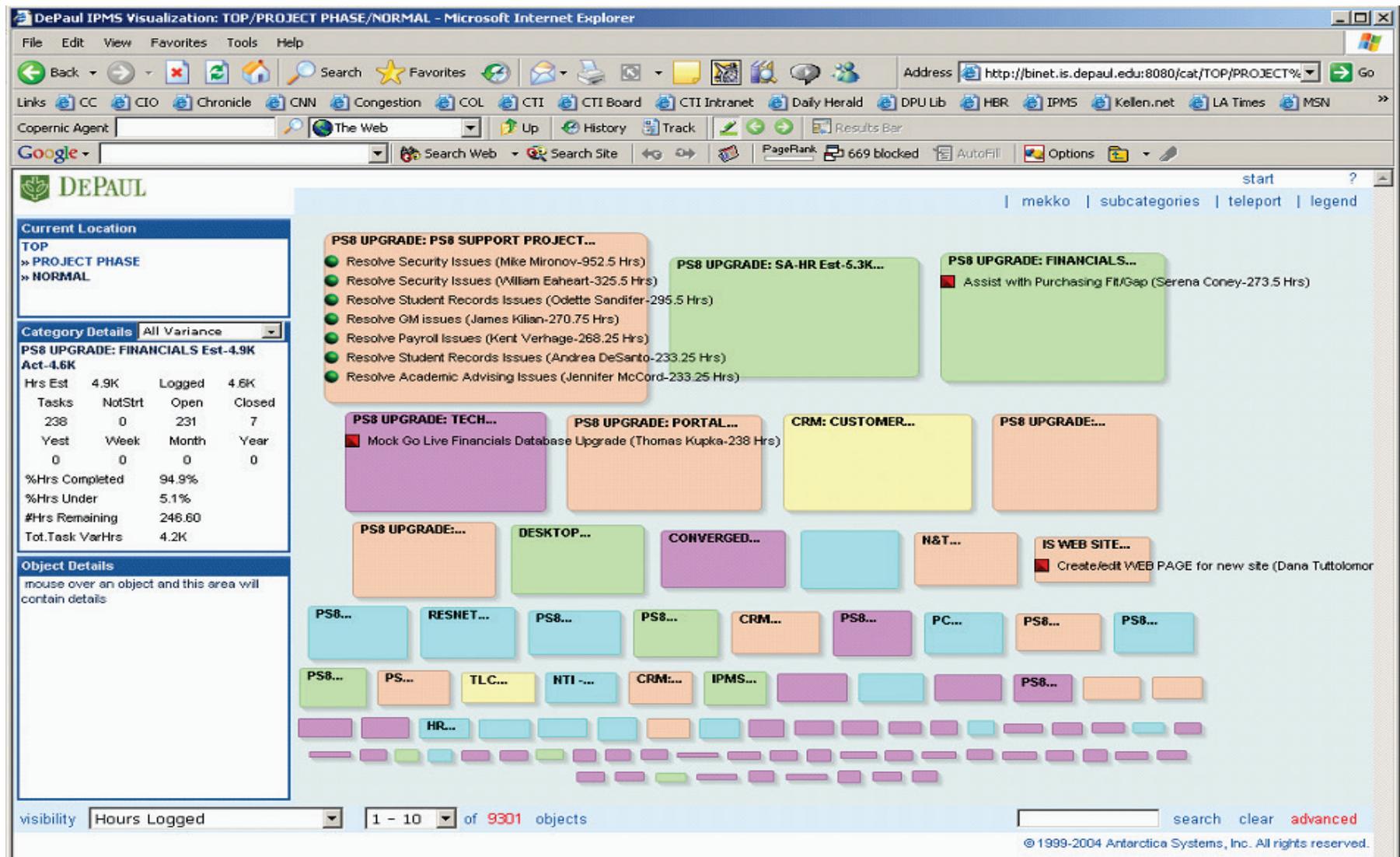
Visualizing project management data

- We applied a visualization tool (VisualNet from Antarctica) to the project management problem
- A different form of project management
 - We use a statistical model, not a causal model
 - Dependencies are ignored, or we use phases for high level dependencies
 - Not all resources are assigned before hand to tasks, no resource leveling
 - Tasks are estimated by a task owner, anyone can record time against a task, tasks must be less than 20 hours (optimal is 8-12 hours in size), can be open, closed or not started, must be estimated such that we are as likely to go over as we are to go under on the estimate (50/50 estimate)
 - Projects are made up of objects. Objects are tangible and can be evaluated by three independent people and assigned a score of simple, average or complex
 - Tasks roll up into phases
 - Tasks also roll up into objects
- No Gantt charts, etc.

Project data model



VisualNet



Other visualization issues

- Adding emotional content can help
 - Color, wording, social context
- Animation may have its place
 - The eye is very sensitive to movement
- Narration, story telling and “movie making” have been shown to be helpful
 - Powerful way to teach/learn

Are you seeing a pattern?

All these techniques exploit System 1, the context-sensitive, more rapid cognitive system much more than the slower, serial, rules-processing cognitive system

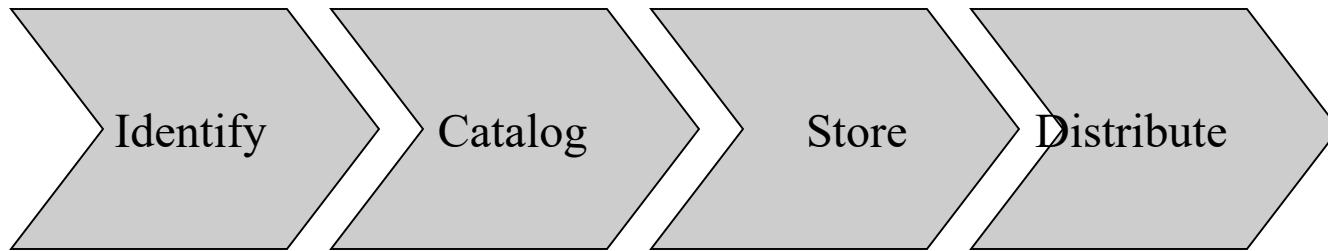
Business implications

- Peter Drucker contends that one of key management challenges for the 21st century is how to increase the productivity of knowledge workers
 - Literally, this means learning faster
- For businesses, this means saving time in learning and improving the outcomes from learning (decisions and results)
- Technology (and visualization) plays a role, but human culture is the more important component
- How do you get individuals (and teams) to convert information to actionable knowledge faster?

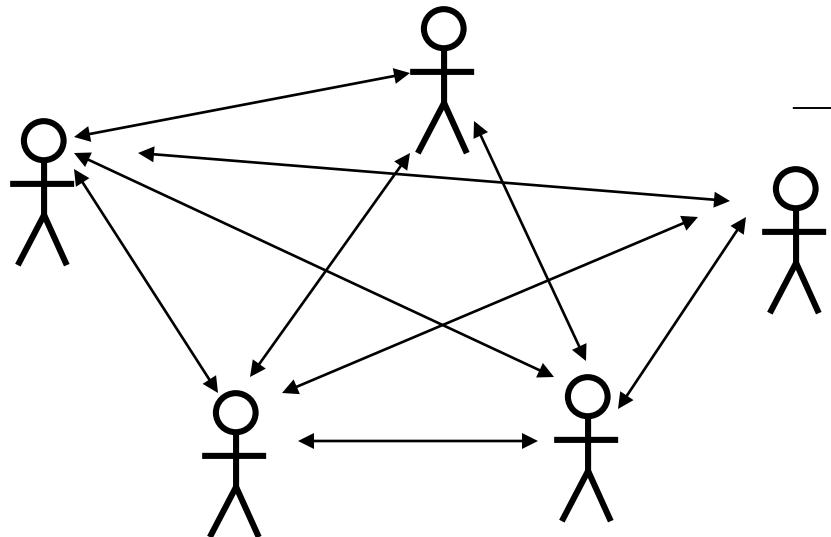
Why focus on culture?

- Visualizations is just one way to convey knowledge
- Knowledge has a social component
 - Knowledge is conveyed through human interactions which are powerful
 - Language, gesture, symbol all play a role
 - The interaction is ongoing and evolves, often unpredictably
 - Blogs, Podcasts, E-mail, SMS, cell phones

Knowledge chain or knowledge network?



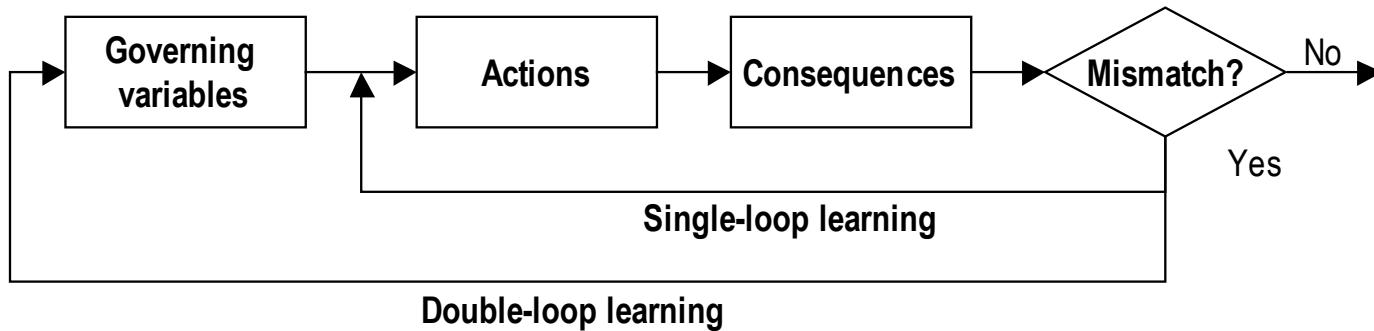
Or



What is the precise sequence of interactions that produces the knowledge needed by the organization? What comprises the interactions? Can you repeat the sequence each time? What rules govern the sequence?

Deterministic, probabilistic reasoning

- Management relies too much on deterministic causality
- In an environment of uncertainty and change, probabilistic reasoning is needed.
 - Probabilistic causality implies seeking of disconfirming evidence, which is hard for people to do, hence the attachment to deterministic reasoning, which needs no disconfirming evidence (because the deterministic reasoning says there is none!).
- Discovering error is the first step toward learning



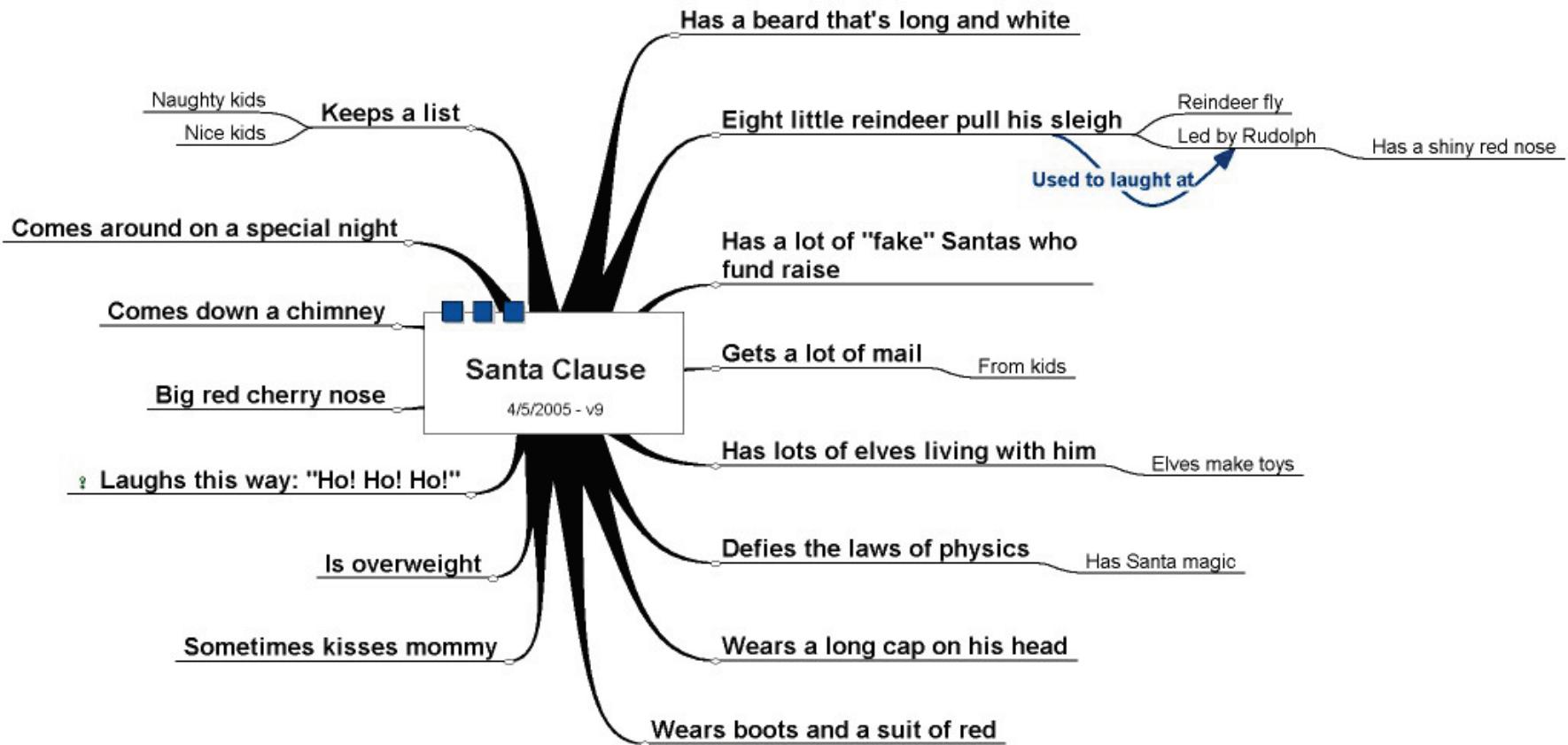
Visualization and ontologies

- Building visualization systems requires establishing an ontology. You need to map concepts to data to visual marks
- An ontology:
 - An explicit formal specification of how to represent the objects, concepts and other entities that are assumed to exist in some area of interest and the relationships that hold among them¹
 - The hierarchical structuring of knowledge about things by subcategorizing them according to their essential (or at least relevant and/or cognitive) qualities¹
 - Ontologies are simply hierachal description of the important concepts in a domain, coupled with a description of each of these concepts. Ontologies consist of various concepts that include: class, subclass, class hierarchy, instance, slot, value, defaults value, facet, type, cardinality, inheritance, variable and relation. The word ontology first appeared in Aristotle's philosophical essays, where it used to describe the nature and organization of being.²

¹ <http://www.dictionary.com>

² “Ontologies for Supply Chain Management,” Ahmad, A., Mollaghazemi, M., Rabelo, L.,
http://www.isye.gatech.edu/faculty/Leon_McGinnis/8851/Sources/Ontology/Ontologies.pdf

The Santa Clause Ontology, informal



Visualization and meaning

- Visualizations should be considered as part of the way an organization describes and categorizes things, not just a pretty pictures. Work the ontological assumptions
- The detail required to set up these visual interfaces usually triggers lots of conversations regarding the “meaning of things”
- As visual interfaces get used and reused, the metaphors and concepts within them get folded into the culture
- There is some measure of profundity in this!

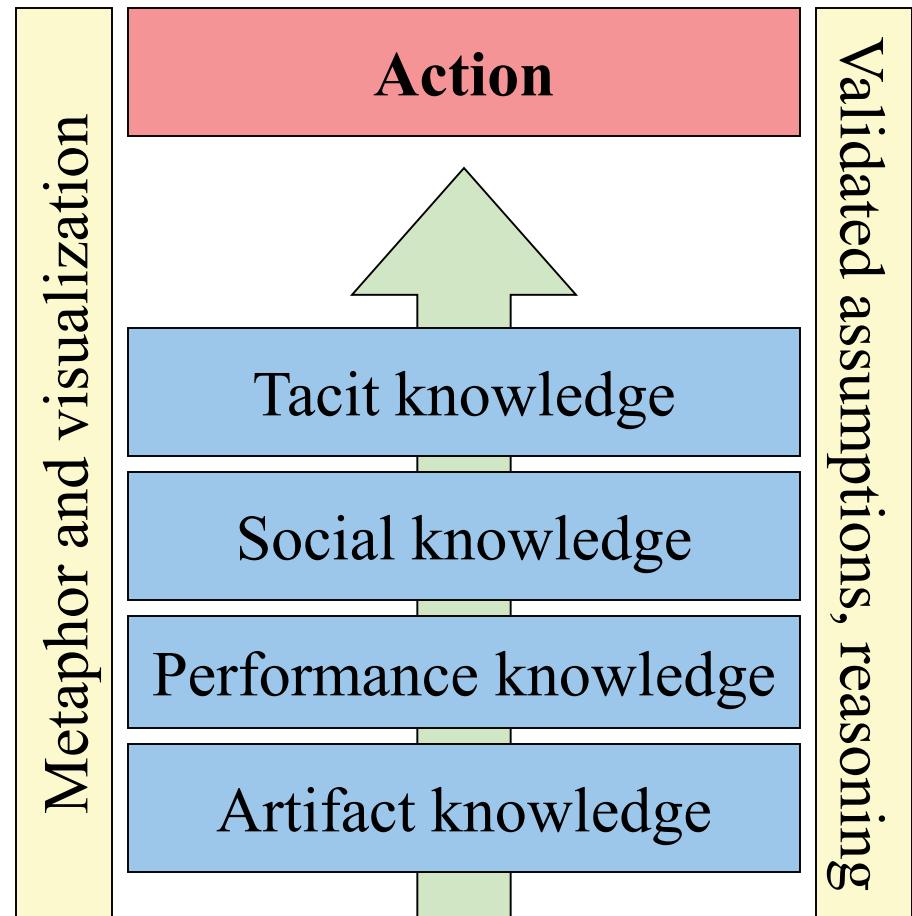
Knowledge layers

Artifact knowledge consists of all the myriads of documents and data that populate company computers and file cabinets.

Performance knowledge consists of the computer systems and documents that convey specific business performance information. This information is often tied to employee compensation and review programs.

Social knowledge consists of all the methods people use communicate outside of artifacts including gestures, symbols, language, culture, group norms. Social knowledge can involve unplanned activities and nonconscious learning.

Tacit knowledge consists of knowledge that resides private to individuals, not codified in an artifact and often not directly communicated as social knowledge.



Are you ready for adoption?

- What is the analytic maturity of the organization?
 - Beginners
 - Basic or no analytic projects; ad-hoc process for analysis
 - Intermediate
 - Just starting to build models, use statistics, uses a process
 - Advanced
 - Uses statistics, OLAP/DSS, scorecards, etc., enhances the process for knowledge discovery
 - Expert
 - Uses basic and advanced statistics and data mining techniques, reporting and analysis is pushed down the organization structure, operations and business process management skills are high
- Visualization can't replace advanced or expert users
 - It can augment and extend them through the development of reusable, operational analytic frameworks, ontologies and metaphors

Common business uses

- Business performance management
 - Scorecards, metrics, process management
- Security and information assurance
- Statistics
- Biological and physical sciences
- Data query and search tools
- Modeling business strategy
- Operational management and decision making
- Wherever one needs to:
 - Speed information search and decision making
 - Improve accuracy of judgments and decisions

Good questions for vendors

- What empirical evidence do you have that supports your claims?
 - Case studies? Take them with a grain of salt
 - Experiments? Is the study relevant to your situation?
- How would your tool work for our specific problem?
 - Can they develop a test case or pilot?
 - Can they speak intelligently about fitting the tool to the task?
- What level of configurability does your tool allow?
 - Can I change the mappings between data and attributes dynamically? Can I filter flexibly? Can I see detailed data?
- Do you have an example of a business case?
 - Is the ROI driven by time savings and/or decision opportunity cost? How is the business case monitored?

Are the issues clear?