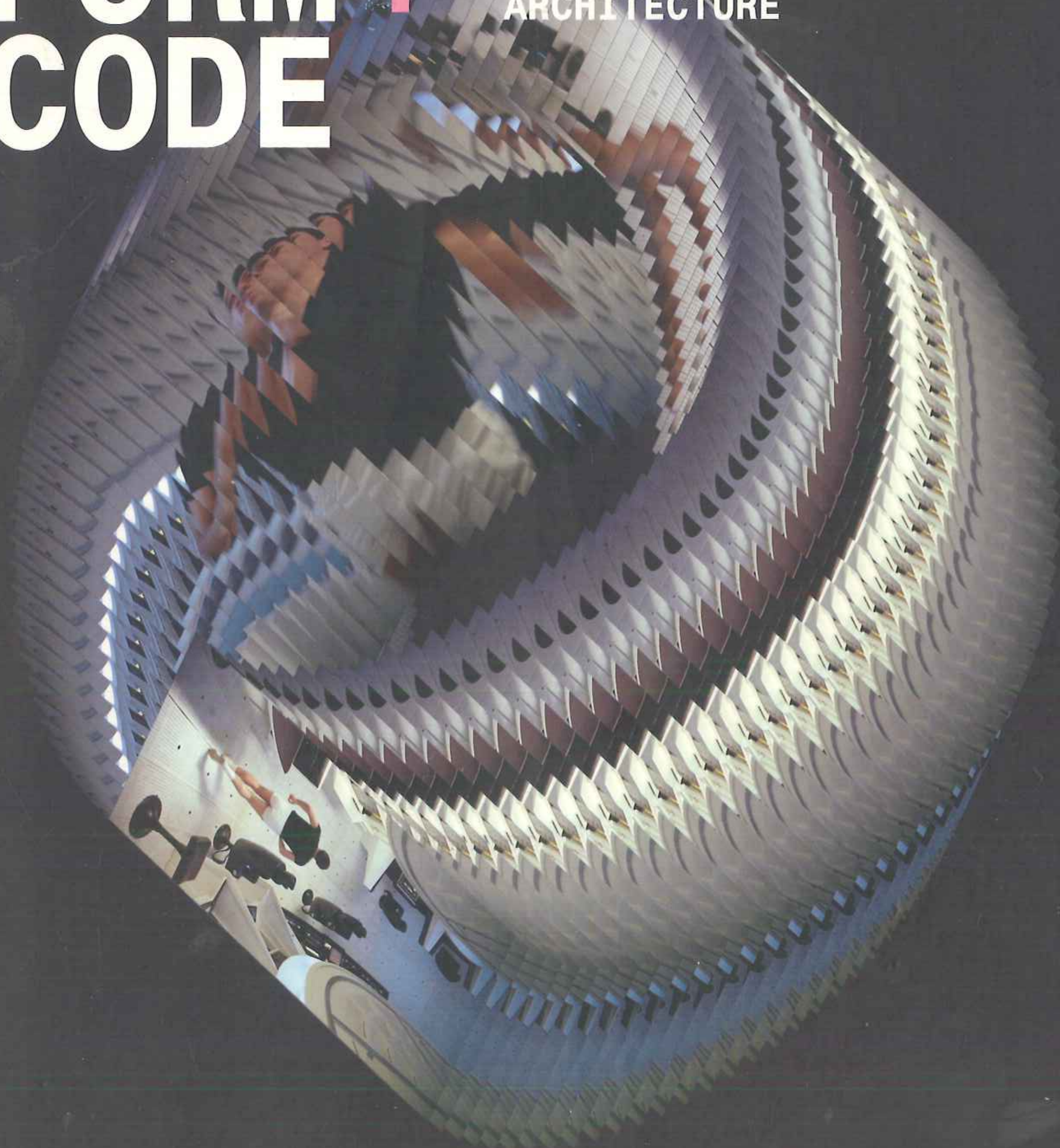
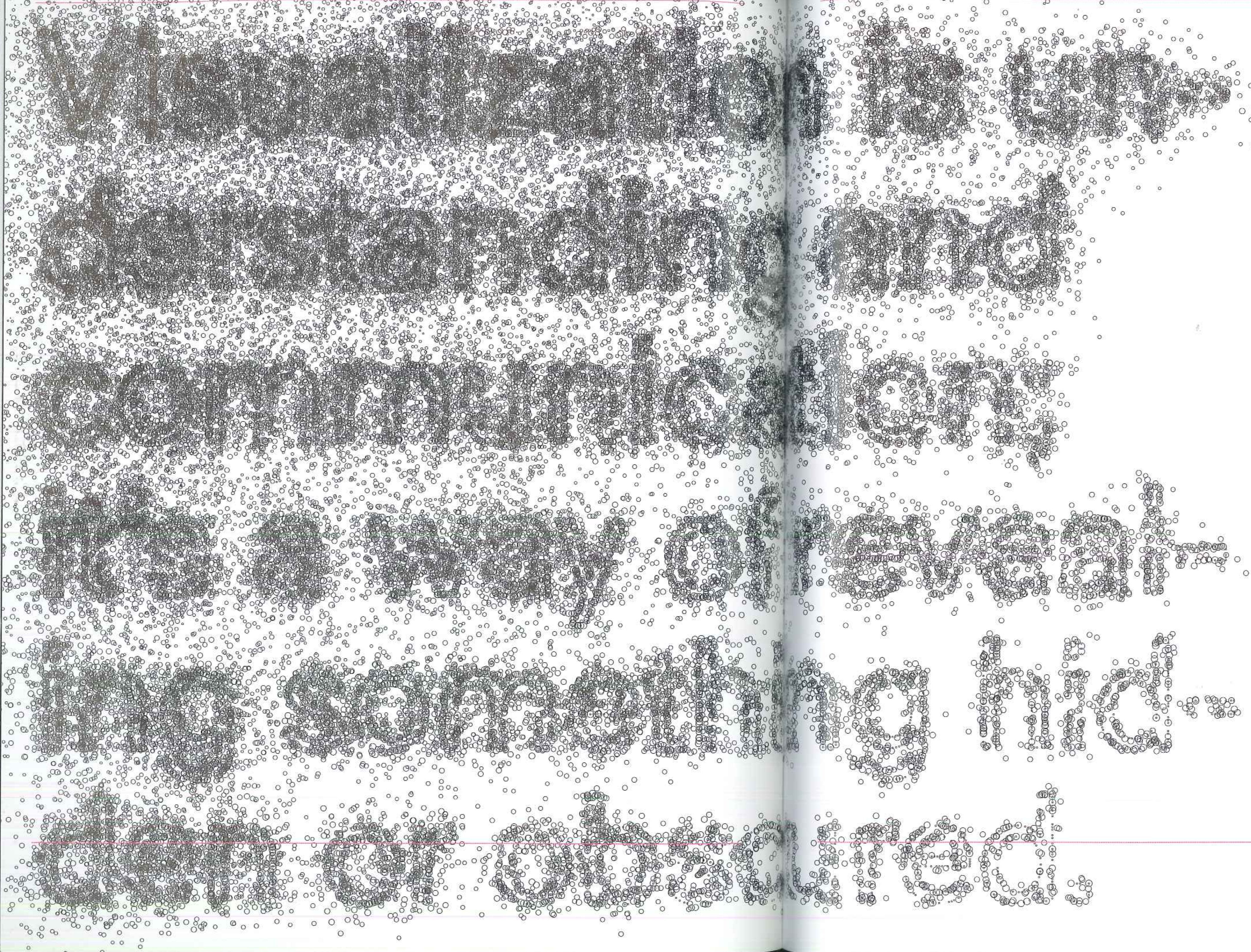
FORM + CODE

IN DESIGN, ART, AND ARCHITECTURE



Casey Reas, Chandler McWilliams, LUST

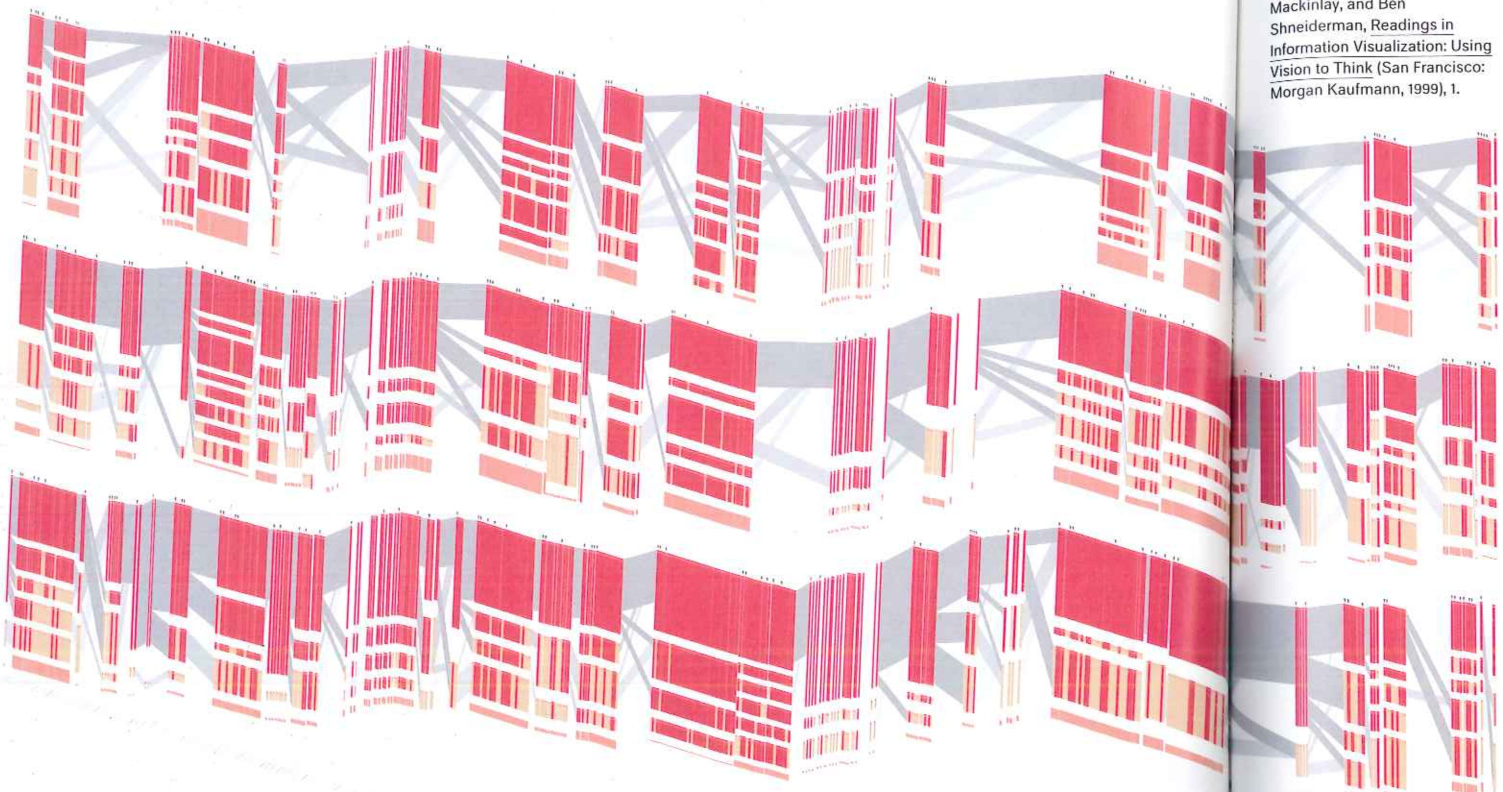
A GUIDE TO COMPUTATIONAL AESTHETICS



Consider New York City's elaborate subway system. The New York City subway map organizes the complexity of this transit system to help passengers navigate from one location to another. The visualization removes unnecessary geographic information and adds information related to train schedules and transfers. The system is still difficult to traverse, but the map's visual clarity makes it manageable. Maps are an early form of visualization—maps of the stars were created before recorded history—but they are only one among a myriad of techniques available to designers. Visualization also helps communicate abstract information and complex processes.

VISUALIZE

These letters were created using a particle system. The particles are attracted to a unique position within a set of interrelated points and, over time, they move toward a single point. This software demonstrates the potential for data visualization to bring clarity to an otherwise chaotic system.



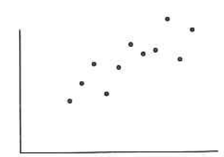
isometricblocks, by Ben Fry, 2003
Fry worked closely with researchers at the Broad Institute of MIT and Harvard to create this visualization of human genome data. It visualizes single letter changes (SNPs) of the genome data for approximately 100 people. It transitions fluidly between common representations for viewing the same data, therefore revealing the relationship between each technique.

DATA INTO FORM

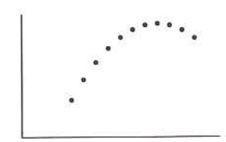
People have a remarkable ability to understand data when it's presented as an image. As researcher Stuart K. Card says, "To understand something is called 'seeing' it. We try to make our ideas 'clear,' to bring them into 'focus,' to 'arrange' our thoughts."[1] Like written words, visual language is composed to construct meaning. Our brains are wired to make sense of visual images. In contrast, it can take years of education to develop the ability to read even the simplest article in a newspaper. The fundamentals of visual understanding, originally pursued by Gestalt psychologists in the early twentieth century, are now researched at a deeper level within the field of cognitive psychology. The findings of this research have been communicated within the visual arts by educators including György Kepes, Donis A. Dondis, and Rudolf Arnheim, as well as through the work of visualization pioneers such as William Playfair, John Tukey, and Jacques Bertin. Data presentation techniques that combine our innate knowledge with learned skills make data easier to understand. In The Visual Display of Quantitative Information, Edward Tufte presents a data set and representation that supports this claim. Compare the tabular data to the scatterplot representation to see how the patterns become immediately clear when presented in the second format.

[1] Stuart K. Card, Jock Mackinlay, and Ben Shneiderman, Readings in Information Visualization: Using Vision to Think (San Francisco: Morgan Kaufmann, 1999), 1.

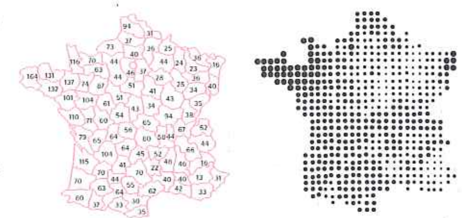
X	Y
10.0	8.04
8.0	6.95
13.0	7.58
9.0	8.81
11.0	8.33
14.0	9.96
6.0	7.24
4.0	4.26
12.0	10.84
7.0	4.82
5.0	5.68



X	Y
10.0	9.14
8.0	8.14
13.0	8.74
9.0	8.77
11.0	9.26
14.0	8.10
6.0	6.13
4.0	3.10
12.0	9.13
7.0	7.26
5.0	4.74



In his book Semiology of Graphics: Diagrams, Networks, Maps, Bertin presents another clear example of the communicative power of visual representation.



The maps of France on the left and right both present the same sociographic data, divided by canton (a French territorial subdivision). The representation on the right replaces each number with a circle sized to correspond to the numerical value. We can spend time analyzing the left map to see where there are concentrations of larger numbers, but on the right map we instantly comprehend the increased density in the upper left.

In the same book, Bertin introduces a series of variables that can be used to visually distinguish data elements: size, value, texture, color, orientation, and shape. For example, a bar chart distinguishes data through the height of each bar, and different train routes on a transit map are typically distinguished with color. For visualizations using only one variable, each element can be used in isolation. For multivariate visualizations (containing more than one variable) elements are combined.

When applying form to data, there are always questions about goodness of fit, meaning how well the representation fits the data. Visualizations can mislead as well as enlighten. As Tufte warns in Visual Explanations: Images and Quantities, Evidence and Narrative, "There are right ways and wrong ways to show data; there are displays that reveal the truth and displays that do not."[2] In Bertin's maps of France, the goodness of fit of the

[2] Edward R. Tufte, Visual Explanations: Images and Quantities, Evidence and Narrative (Cheshire, CT: Graphics Press, 1997), 45.

Vinec,
by Catalogtree, 2005
This graphic shows data for 10,000 cars crossing a bridge between the cities of Arnhem and Nijmegen in the Netherlands between 07:36 and 09:13 am. The horizontal axis for each unit shows the distance between the vehicles and the vertical axis shows the difference between the measured speed and the speed limit.

Flocking Diplomats,
by Catalogtree, 2008
Between 1998 and 2005, diplomats in New York City were responsible for 143,703 parking violations. This visualization shows each violation at its location within the city, with the United Nations at the epicenter.

Newsmap, by Marcos Weskamp, 2004–9
The Newsmap visualizes the changing landscape of the Google News aggregator. The interface allows users to view the news from one or more of eleven countries and to view articles from seven categories.

representations reveals information hidden within the data. Because each piece of data derives from a particular canton, associating that data with its location on a map allows us to see regional patterns. Presenting the data in a table that is organized alphabetically would not reveal this pattern. By applying the same visualization technique to a different map—a map of Europe, for example—would also not work as well. The Bertin map works because each canton is roughly the same size, but the size differences among European countries would dilute the visual patterns needed for interpretation. In this case, the data is tightly linked to a source (geography), but in other instances data can be more abstract, such as when revealing patterns in language. The visualization techniques that follow in the next section present other options for revealing patterns in data.

There are hundreds of distinct visualization techniques that can be organized into categories, including tables, charts, diagrams, graphs, and maps. When creating a new visualization, one technique is selected instead of another based on the organization of the data and what the visualization is meant to convey. Data representations that commonly appear in newspapers, such as bar charts, pie charts, and line graphs, were all developed before people relied on software; in fact, most commonly used data representation techniques are only useful for representing simple data (1- and 2-D data sets). These techniques are automated within frequently used software tools such as Microsoft Excel, Adobe Illustrator, and related programs. Visualizing information, once a specialized activity, is becoming a part of mass culture.

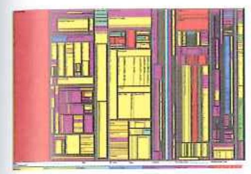
Writing new software is one approach to move beyond common data representations. New visualization techniques emerge as researchers and designers write software to fulfill their growing needs. The treemap technique is a good example to demonstrate the origins and evolution of a new visualization. It also shows how techniques often arise within a research group and are visually refined as designers use them in diverse contexts.

A treemap is a visualization that utilizes nested rectangles to show the relations between one or more data elements. They are effective because they allow for easy 2-D size comparisons. The story of the first treemaps, from their origin to the present, is documented by the technique's originator, Ben Shneiderman, a professor at the University of Maryland.[3] The first treemaps were developed in 1991 as a way to show the memory usage on a computer hard drive. After subsequent applications and further development, the public at-large was introduced to treemaps by Map of the Market, an Internet application created by smartmoney.com in 1998. This application introduced the innovation of making the tiles close to square, rather than using the thin tiles of previous treemaps, to increase legibility. The circular treemap technique explored by interface designer Kai Wetzel in 2003 pushed the form of treemaps even further. Wetzel worked on this representation as one of many ideas for a Linux operating system interface. He recognized that the approach wastes space and the algorithm is slower, but the aspect ratio of each node is the same. The 2004 Newsmap application by Marcos Weskamp applied treemaps to the headlines of news articles compiled from the Google News aggregator. The treemap representation makes is easy to see how many articles are published within each news category. For example, the visualization makes clear that, in England, the highest volume of published articles is world news rather than national stories, while in Italy, the reverse is true. By 2007, through the refinement of these and other initiatives, the treemap technique had become so ubiquitous that it was used in the New York Times with the expectation that a general audience can understand it.

The era of modern data analysis began with the 1890 U.S. Census. The Census Bureau

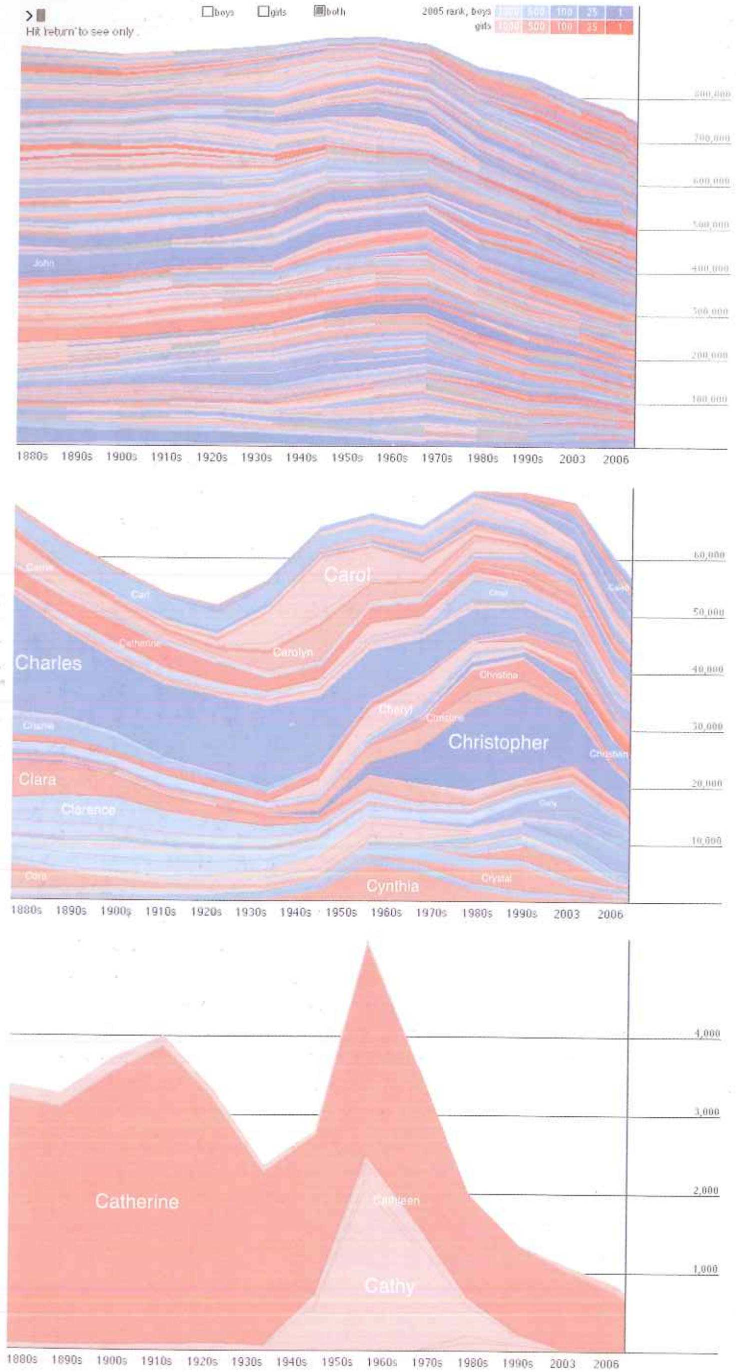
[3] Ben Shneiderman and Catherine Plaisant, "Treemaps for space-constrained visualization of hierarchies," http://www.cs.umd.edu/hcil/treemap-history/.

VISUALIZE



Treemaps, by Ben Shneiderman and Brian Johnson, 1991
This visualization depicts the file contents of a networked computer's hard drive that is shared by fourteen people. It shows who was using the most space and which large files (represented by large rectangles) could be deleted to make space.

Circular Treemaps, by Kai Wetzel, 2003
Like the original Treemaps, this visualization also depicts file space on a hard drive. In this variation, the age of the file is shown with color. Red is applied to new files and the oldest files are a soft yellow.



NameVoyager, by Laura and Martin Wattenberg, 2005
The NameVoyager shows the popularity of baby names in the United States from the 1880s to the present. The thickness of each name's color band shows how many babies were given that name each year. Names are selected by clicking on its band or by typing. Type a single letter to refine the search to only names beginning with that letter. Type an additional letter to further narrow the search.

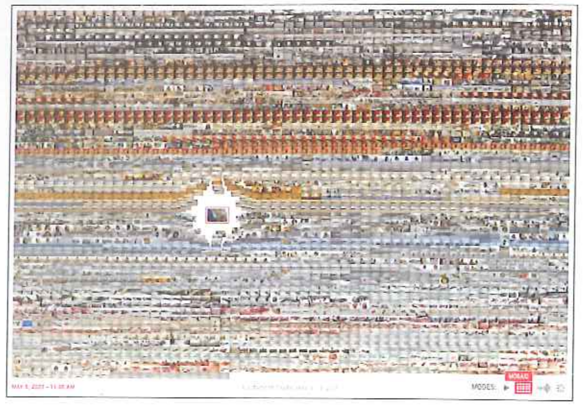
DYNAMIC FILTERS

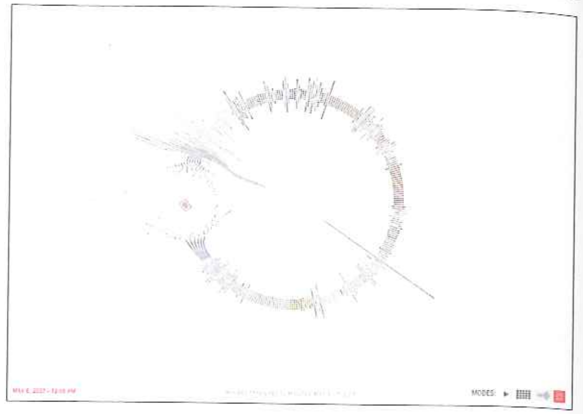
realized that the population was growing so quickly that, using their existing analysis method, the newly collected data would be out of date by the time it was evaluated. Herman Hollerith was commissioned to build a machine to automate the process. He succeeded by developing punch cards to store data and a machine to read them. Since that time, computing machines have enabled an unprecedented ability to acquire data. Today, it's no longer possible to rely solely on static, paper representations. The deluge of data created by the information age has made new forms of software analysis necessary.

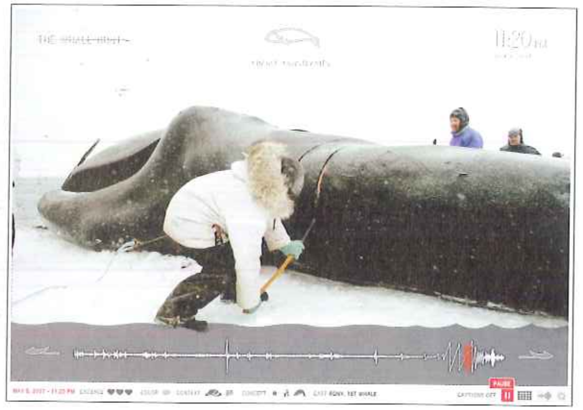
When there is more data than can be viewed at one time, it's necessary to limit the amount that is displayed. This process is called filtering or querying. An internet search is an act of filtering. The internet is such a large data source that to get value from it, search terms are used to limit what is seen at any given time. A tool for filtering can provide different levels of control, often corresponding to its use at a basic or expert level. For instance, when searching the web, it's possible to make a simple search where only keywords are input, but it's also possible to select a broad range of criteria such as file type and date to further refine the results. A real estate database is a good example of a large data set that becomes more useful with a filter. It doesn't make sense to look for all available listings for apartments, houses, and condominiums if you're only looking for one of the options. A tool like this becomes even more useful to search by a specific price, size, or neighborhood. In fact, a 1992 research project called the Dynamic Home Finder, developed by Christopher Williamson at the University of Maryland, was an early prototype for such a system.

Another example is the NameVoyager by Laura and Martin Wattenberg. It's a clear example of a continuous visualization that is accessed through a dynamic filter. The project presents a simple way to discover the popularity of nearly 5,000 baby names used in the United States from the 1880s to the present. For example, by typing the name Deanna, we see this name originated in the 1920s, reached its peak in the late 1960s, and has since become less common. The interface begins with a stacked-graph representation of all names in the data set. When a letter is typed, the search narrows to reveal only those names that begin with the specified letter. For example, typing the letter C reveals the historic popularity of Christopher and Charles. Further inputting an A and T reveals that Catherine is the most popular name that begins with Cat and also shows the relative obscurity of Catalina and Catina, among others.

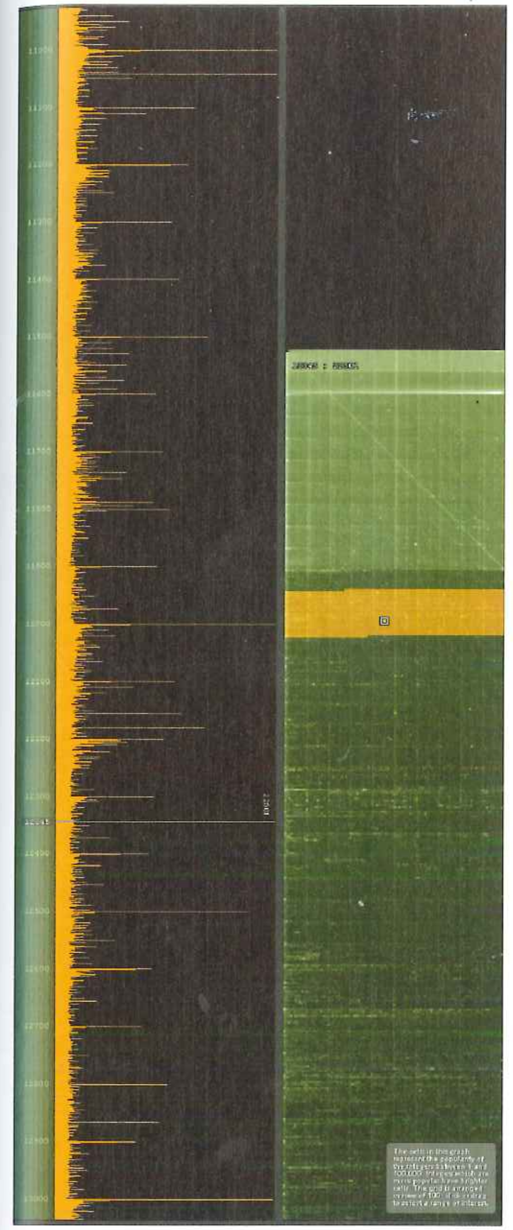
A group of photographs is also a database that may be filtered and navigated. The proliferation of online digital photographs has created a fascinating data set to be explored through image searching. Jonathan Harris's The Whale Hunt is an elegant example of navigating a large series of images. He travelled to Alaska to participate in a whaling expedition with a family of Inupiat Eskimos. He took 3,214 photos from the time he left New York City to returning home nine days later. By default, the images are presented sequentially, but he also allows viewers to navigate in other ways. It's possible to jump to any image in the timeline, pause, and change the pace. But more interestingly, it's possible to select a subset of images based on categories. By using a fluid and clear interface, images can be filtered according to a concept (e.g. blood, boats, buildings), context (e.g. New York City, Alaska, the Patkotak family house), or a member of the cast (e.g. Abe, Ahmakak, Andrew). After one or more selections are made, the image sequence is automatically edited and the story told through the images is changed to reflect the filters.

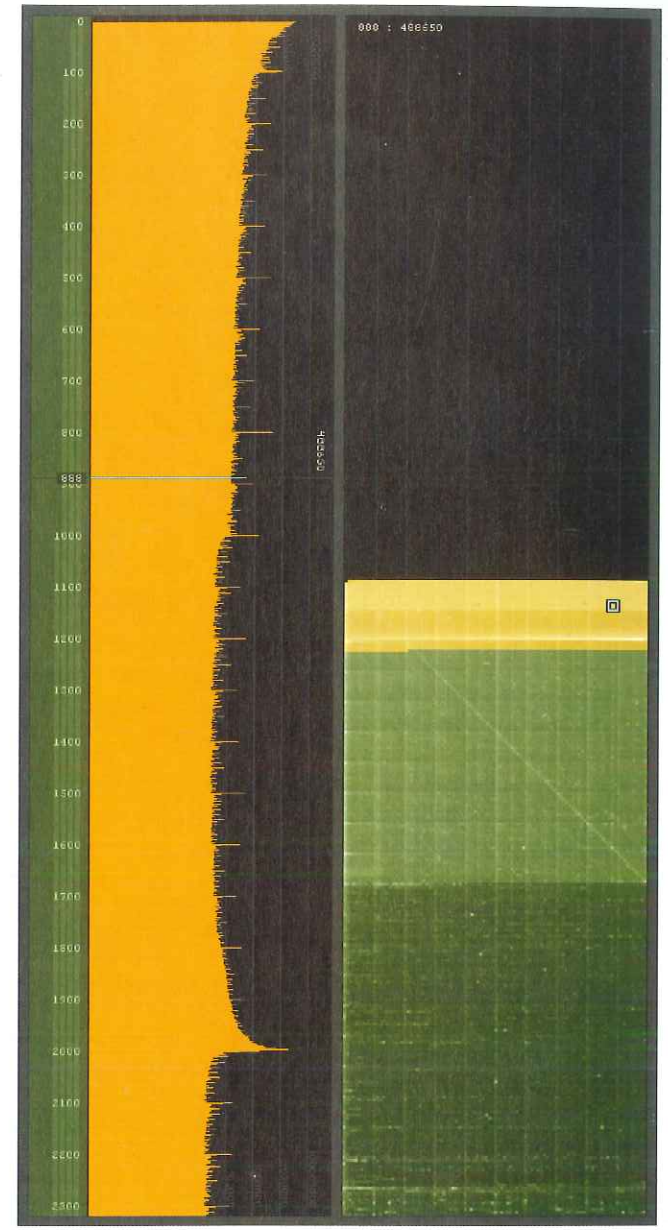




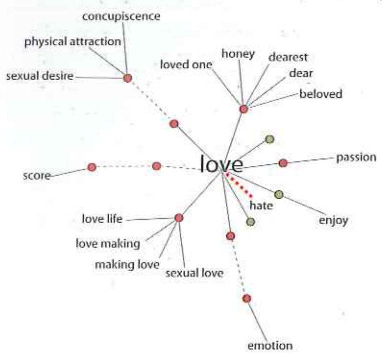


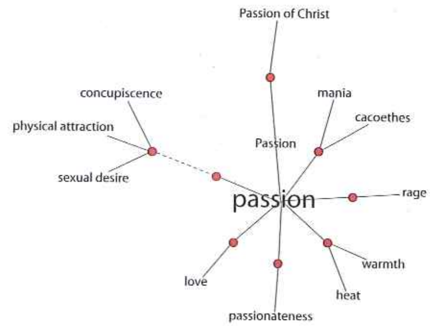
The Whale Hunt, by Jonathan Harris, 2007
The Whale Hunt is an experimental interface for storytelling. Viewers may rearrange the photographic elements of a story to extract multiple sub-stories focused around different people, places, topics, and other variables.

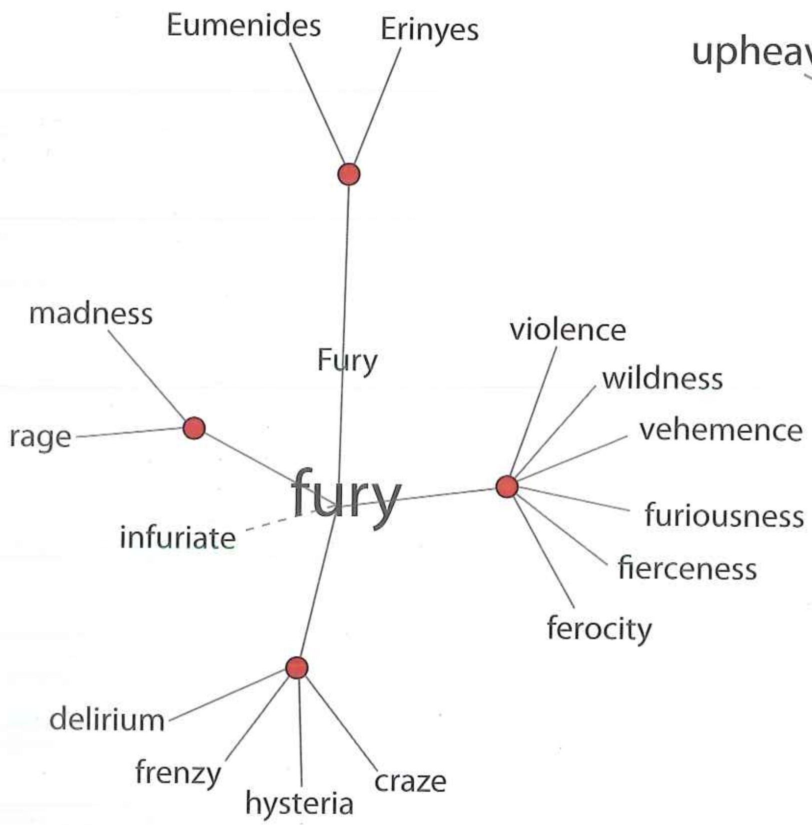


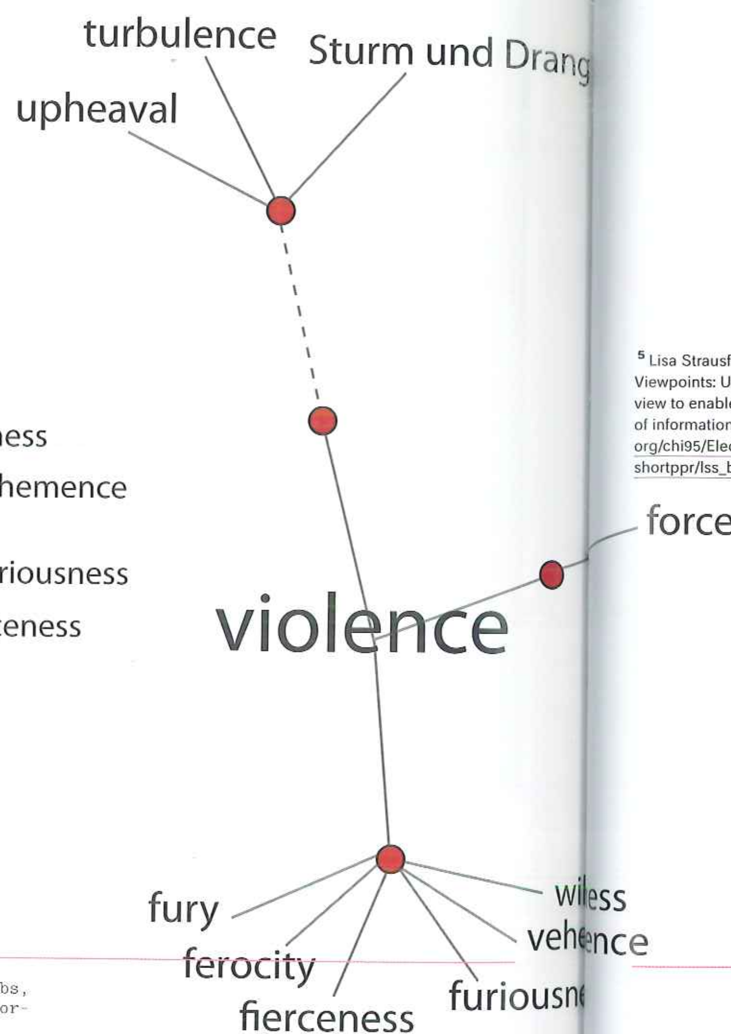


The Secret Lives of Numbers, by Golan Levin with Martin Wattenberg, Jonathan Feinberg, Shelly Wynecoop, David Elashoff, and David Becker, 2002
This web-based project reveals the relative popularity of every integer between zero and one million according to their frequency on web pages. The interface allows for exploration of this massive data set and provides a novel lens for viewing social patterns. Levin states, “Certain numbers, such as 212, 486, 911, 1040, 1492, 1776, 68040, or 90210, occur more frequently than their neighbors because they are used to denominate the phone numbers, tax forms, computer chips, famous dates, or television programs that figure prominently in our culture.”









Visual Thesaurus, by Thinkmap, 1998-present
When a word is selected, a new set of related words emerges. In these images, follow the path from love to passion to rage to fury to violence. Related words are connected with gray lines and antonyms are connected with red-dashed lines. Different parts of speech (verbs, nouns, etc.) are color-coded. Try it at www.visualthesaurus.com.

NAVIGATION

Some early forms of data navigation might have included flipping through clay tablets, moving through a room painted with hieroglyphs, and rolling and unrolling a scroll. Early books improved upon scrolls because they allowed the reader to move quickly between sections in the text and could be smaller and therefore easier to carry. Book conventions such as the index, page numbers, and table of contents developed slowly. Despite thousands of years of refinement and the widespread proliferation of the Internet, we're still scrolling and viewing data on pages. The unique tool for looking at and navigating pages on the web is the hyperlink, a link from one page to another. Ted Nelson coined the phrase hypertext in the 1960s to describe this concept. Since that time, designers and researchers have pushed forward this and other innovative navigation concepts by writing software.

As an example, imagine the data inside a thesaurus. There's a list of thousands of words in addition to all of the relations from each word to others. To explore this data, you look up one word, which you may then follow to another, and so on. Even the small amount of time needed to hunt for the next word can break the flow. The Visual Thesaurus software, written by Thinkmap, makes navigating language relations a more fluid experience. The software shows a network of words related to the currently selected word. Clicking on one of the outlying words makes it the center, and new words appear that relate to it, while the former relations disappear. The interface allows the user to see the context around the current selection, but avoids overwhelming the senses with additional layers of nonrelevant information.

Spatial navigation is an emerging technique for exploring data, but it has roots that are thousands of years old. The ancient memorization technique Method of Loci, sometimes called a Memory Palace, places information inside imagined rooms within the mind to enhance recall by associating data with a mentally navigable space.

The sci-fi novels of William Gibson introduced intriguing concepts for spatial data navigation. In Neuromancer, published in 1984, he wrote about "rich fields of data" and described a vision of cyberspace: "A graphic representation of data abstracted from the banks of every computer in the human system."[4] Although Gibson's world was fictional, a related real-world concept was developed by designer Lisa Strausfeld in 1995. She describes her software, Financial Viewpoints, as follows:

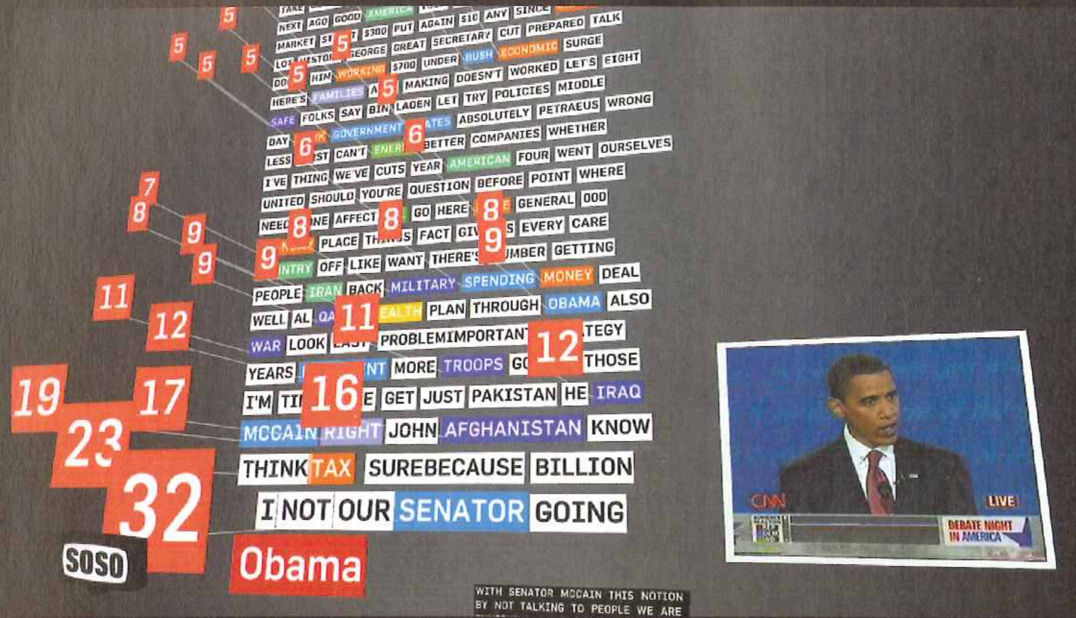
> Imagine yourself without size or weight. You are in a zero-gravity space and you see an object in the distance. As you fly towards it, you are able to recognize the object as a financial portfolio. From this distance the form of the object conveys that the portfolio is doing well. You move closer. As you near the object, you pass through an atmosphere of information about net assets and overall return statistics. You continue moving closer. Suddenly you stop and look around. The financial portfolio is no longer an object, but a space that you now inhabit. Information surrounds you.[5]

At that time, Strausfeld was a research assistant in the Visible Language Workshop (VLW) at the MIT Media Lab. The research group was directed by Muriel Cooper, who set out to discover what graphic design could mean in the new era of communications, through the use of software applications. David Small, another researcher in the group, used software to present large bodies of text within a single navigable environment. His Virtual Shakespeare presents the entire works of William Shakespeare within one continuously navigable space. From the long view, only the names of individual plays, such as Hamlet and Henry V, are visible, but as you zoom closer, the acts come into view as rectangular textures, and finally it's possible to read the dialog and stage directions.

[4] William Gibson, Neuromancer (New York: Ace, 1984), 51.

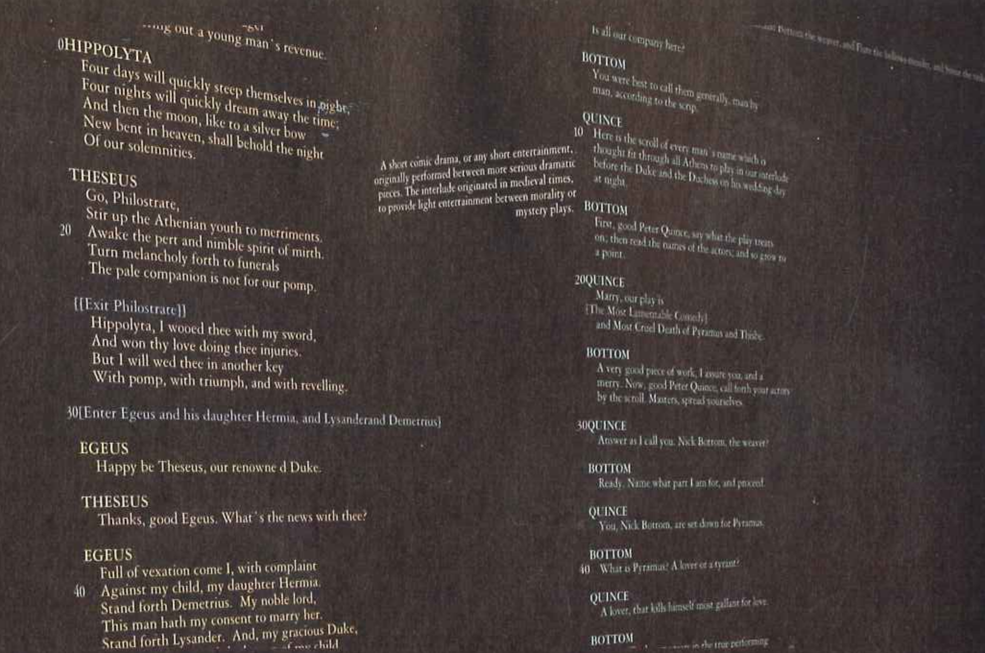
[5] Lisa Strausfeld, "Financial Viewpoints: Using point-of-view to enable understanding of information," http://sigchi.org/chi95/Electronic/documnts/shortppr/lss_bdy.htm.

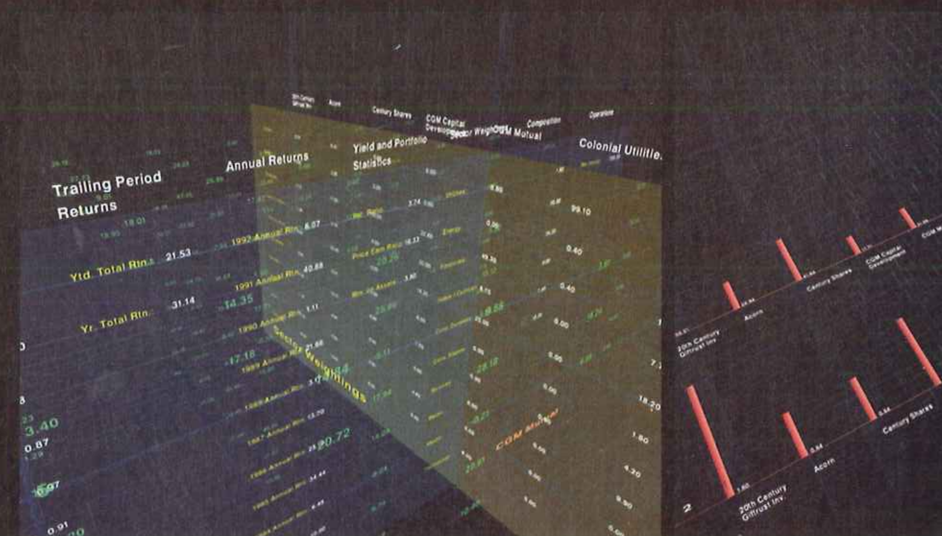
VISUALIZE



ReConstitution, by Sosolimited (Eric Gunther, Justin Manor, and John Rothenberg), 2008

This software was written for use with live performances during three presidential debates between Barack Obama and John McCain. Algorithms were applied to the live images and closed captions to dynamically visualize the way language was used during the debates.

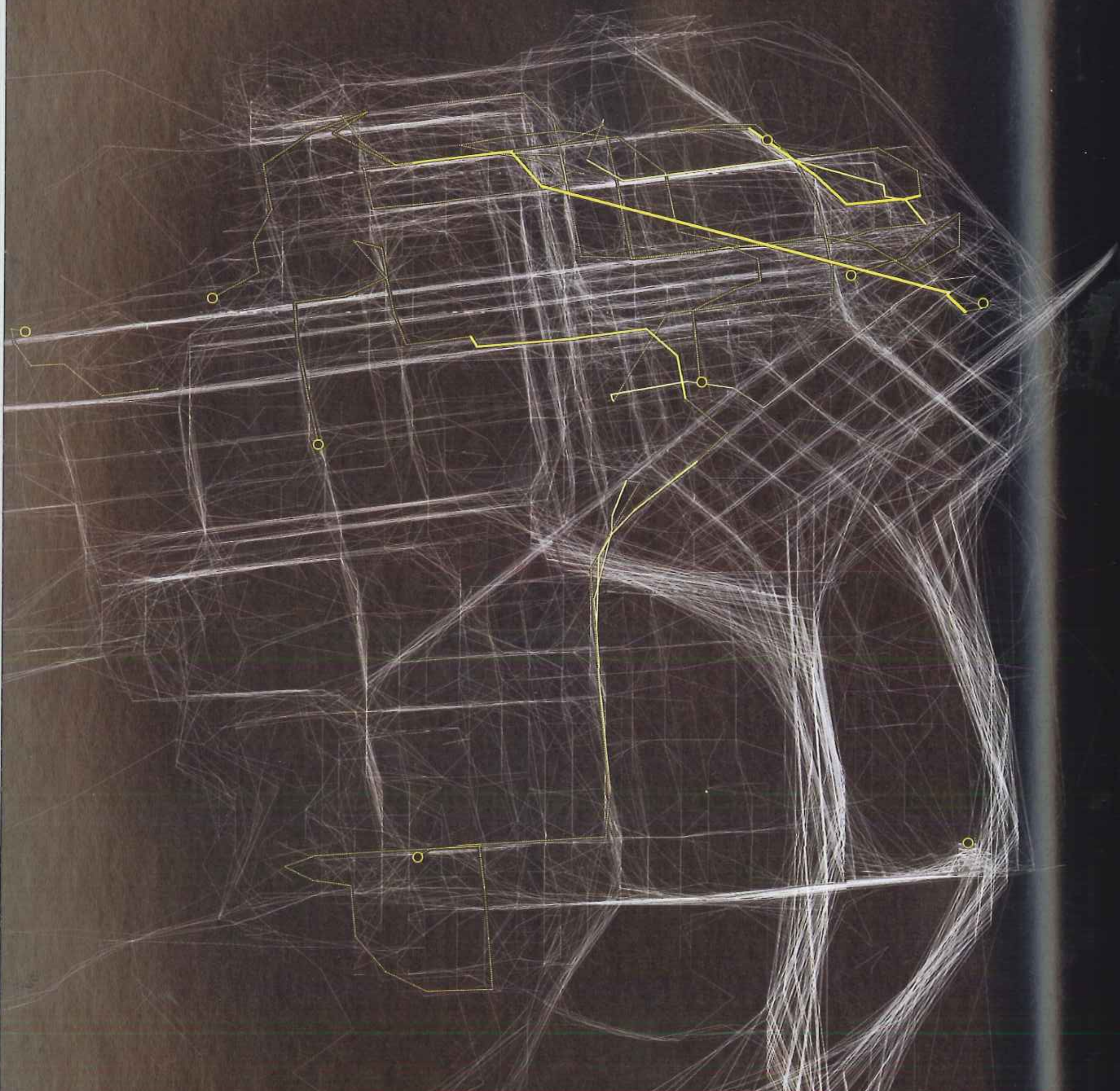




VISUALIZE

Virtual Shakespeare, by David Small, 1994
Utilizing software's ability to dynamically scale typography, Small presents all of Shakespeare's plays within a navigable environment. From far away, each play looks like a column of lines. As you navigate closer, each character's dialog is revealed.

Financial Viewpoints, by Lisa Strausfeld, 1995
Strausfeld defines her project as "an experimental interactive 3-D information space that spatially and volumetrically represents a portfolio of seven mutual funds. A 3-D point of view is used to represent context, and context shifts in the information to allow users to view multiple representations of the information in a continuous environment."

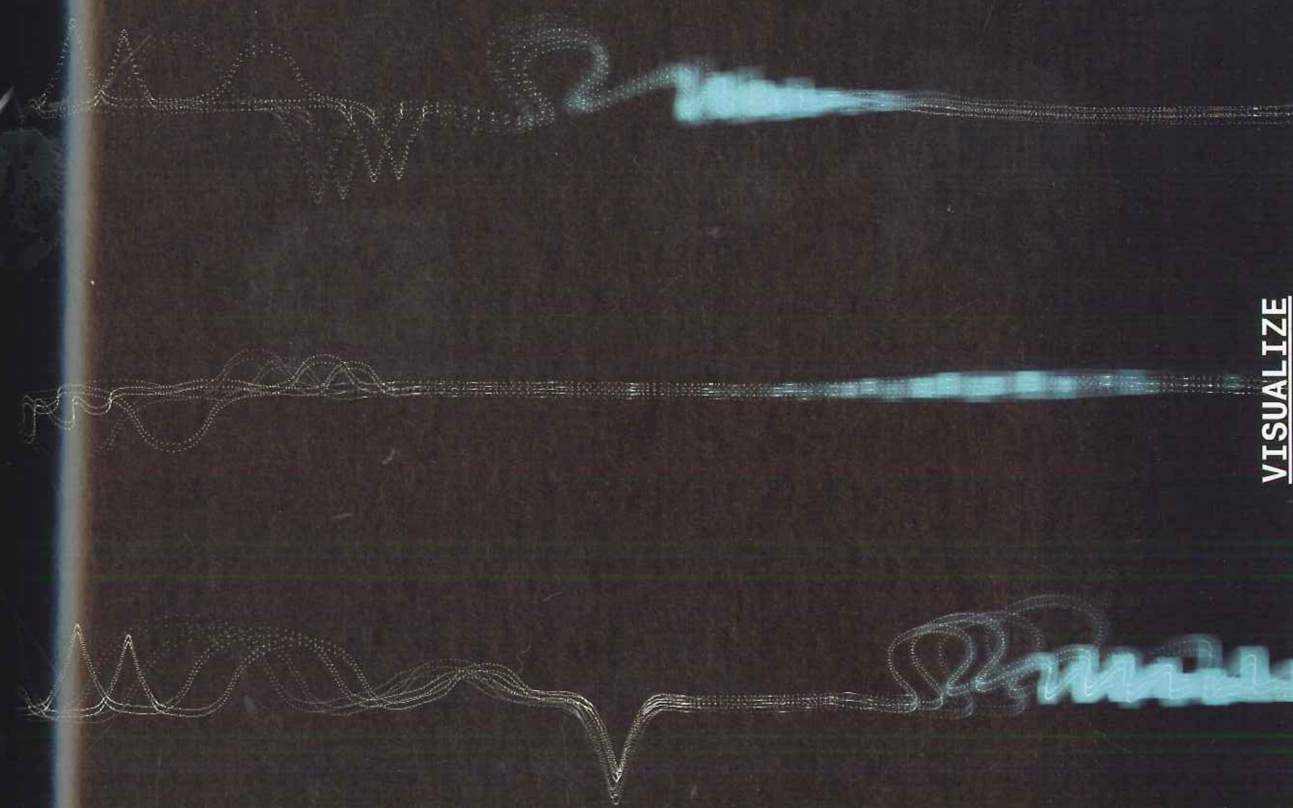


Cabspotting,
by Stamen design with Scott Snibbe, Amy Balkin, Gabriel Dunne and Ryan Alexander, 2006–8

To track the movement of taxis around the San Francisco Bay Area, Global Positioning System (GPS) devices were attached to the vehicles. Lines are drawn to connect the GPS data points, showing the taxis' paths as they navigate the city. The city grid and circulation patterns are revealed through line density.

VISUALIZATION TECHNIQUE
TIME SERIES

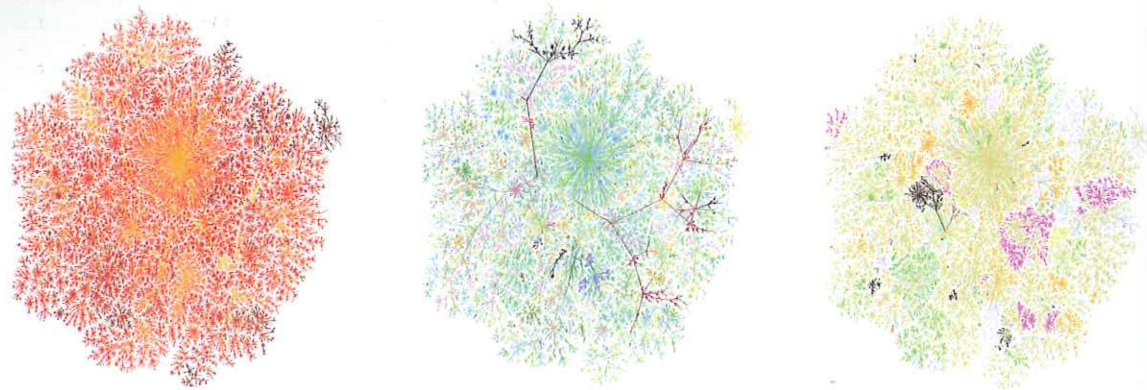
A time-series visualization shows data collected over a long period within a single image. It compresses many moments into a single frame. A time-series image can be a single, static image or it can be an animated image that combines data through motion. By using time as the ordering principle, changes become clearer.

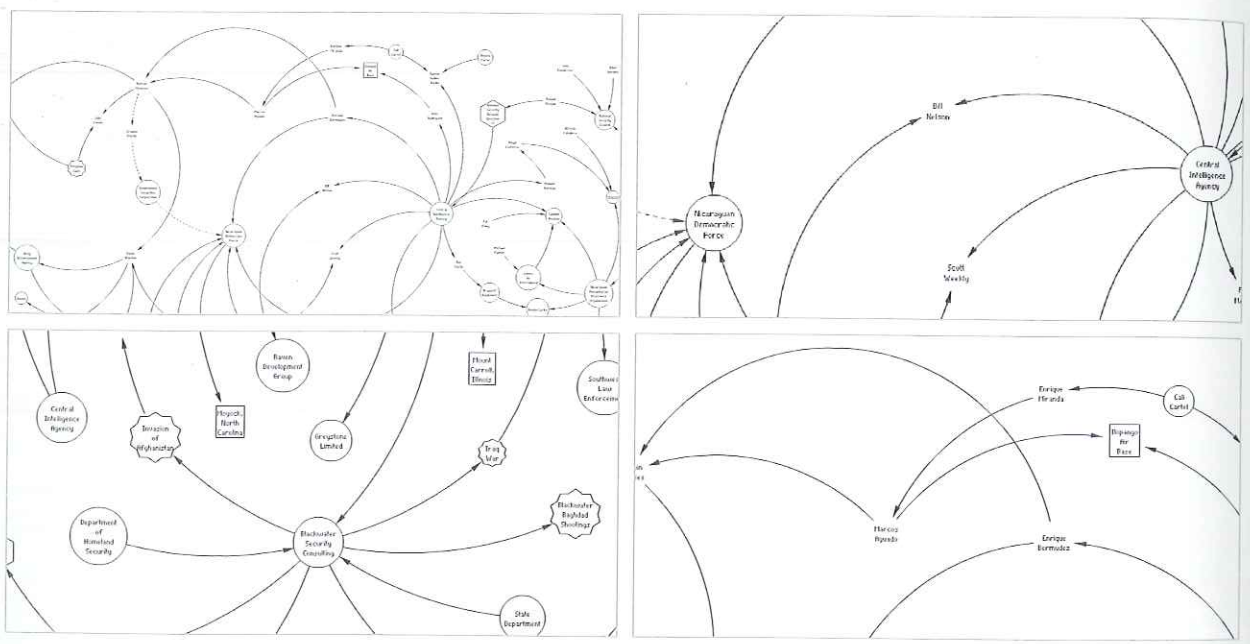


Takeluma,
by Peter Cho, 2005
Takeluma is an innovative alphabet that displays the rhythms and emphases of human speech. Each letter relates to the visual quality of a consonant or vowel sound. The characters are different from any known alphabet, but each glyph has an analog within the English alphabet. Language flows into a single expressive line.

This image represents part of the Neil Armstrong quote, "That's one small step for man, one giant leap for mankind."

Flight Patterns, by Aaron Koblin, 2005–9
The animated path of each plane is shown as a line delineating where the plane is and where it was, therefore implying its destination. This gives insight into the nature of the invisible high-ways far above the ground. Population clusters are visible through the densities of patterns.

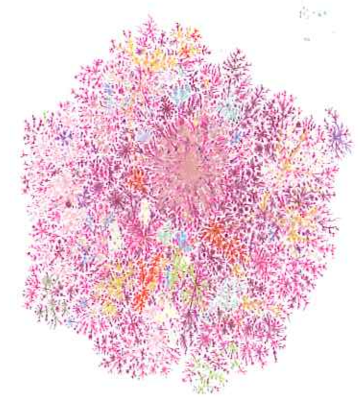




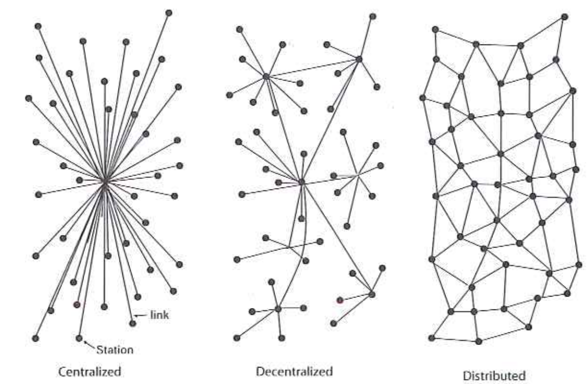
The Internet Mapping Project,
by Bill Cheswick and Hal Burch, 1998
These images were the first attempts to assign an appropriate visual form to the Internet. They captured the public's imagination; from left: color is applied in relation to the distance from the test host, as a function of the network address, colored by the top-level domain (the black areas are .mil sites) and by the Internet service provider.

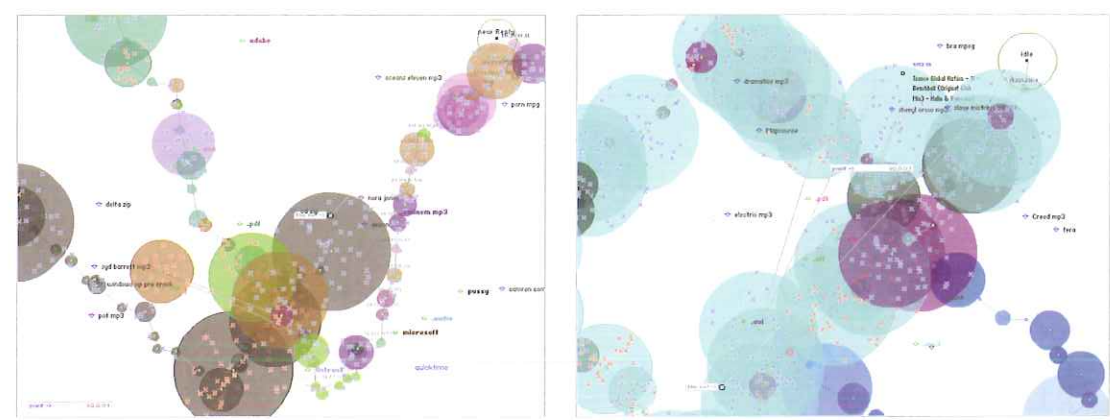
Power Structures,
by Aaron Siegel, 2008
These drawings build on the work of Mark Lombardi, an artist who depicted crime and conspiracy information as visual networks. It is a relational database and mapping tool that allows users to contribute data and draw relationships.

VISUALIZATION TECHNIQUE
NETWORKS



As social, political, and technical networks become denser and more complex, widespread interest in visualization is growing. Provocative visualizations help us to better understand the sometimes invisible relationships that affect our world. Network diagrams frequently include two types of elements: nodes and connections. A node is an individual element (a person, country, or computer) and connections show relationships between the nodes. Visualizations help us to see different types of networks: centralized (star), decentralized (hybrid), and distributed (grid or mesh). These different organizations were elegantly diagrammed in 1962 by Paul Baran, one of the conceptual architects of the Internet.



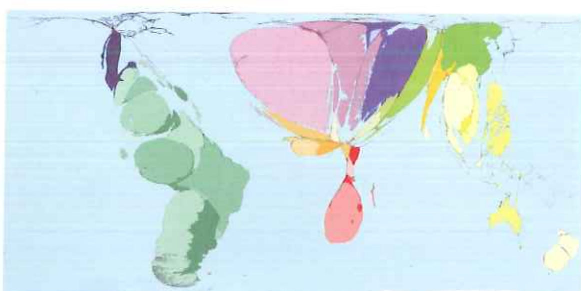


Minitasking,
by Schoenerwissen/OfCD, 2002
With the rise of Internet file sharing, the duo of Anne Pascual and Marcus Hauer built Minitasking to reveal the ad hoc networks created through the Gnutella network, a peer-to-peer protocol for sharing files. Data fed into the visualization as the transactions occurred, revealing the structure, filenames, and rhythm.

Tourism expenditure

Child labor

Coal power



Fruit export

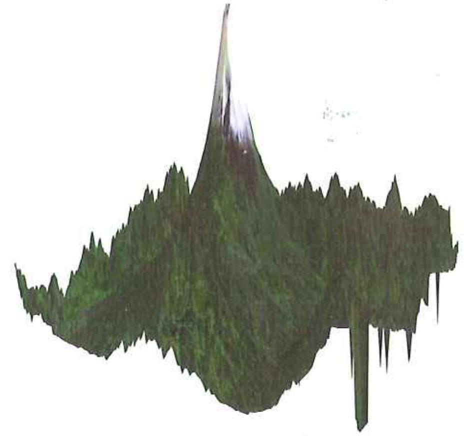
Population in year 1 CE

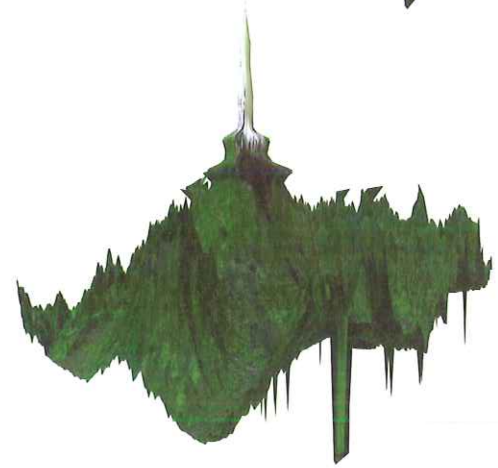
Worldmapper.org, 2006
In these cartograms, information is communicated through deformations in the size of each territory represented in the data being mapped. For example, in the map representing Coal Power, the territory size shows the proportion of worldwide electricity generated from the coal produced there. Because the United States produces a high amount of energy from coal, its size is proportionally larger than countries that produce a smaller amount. This type of map helps the viewer see differences immediately; clockwise from top, the maps show: tourist expenditure, coal power, population in the year 1 CE, fruit exports, and child labor.

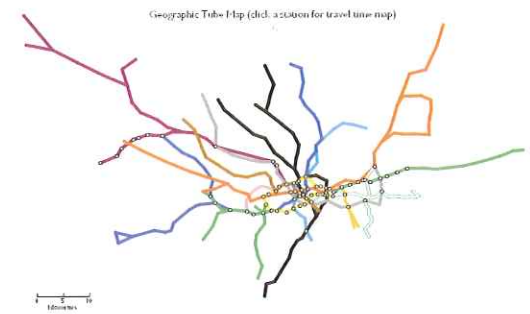
VISUALIZATION TECHNIQUE

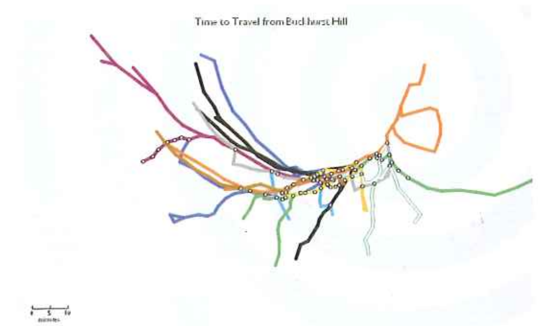
DYNAMIC MAPS

Most maps show many layers of information within a single surface. For example, a single map might show the locations of roads, landmarks, topography, and political borders. Because the sophisticated language and representation of maps are so familiar, they provide a good foundation for additional layers of information. Adding changes in time and geometric distortion are effective ways to push the conventions further.







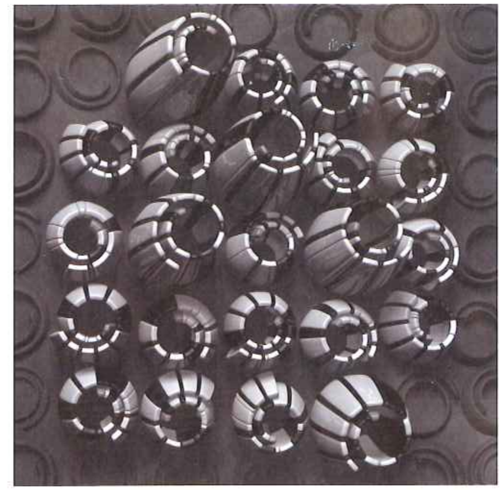


Impressing Velocity, by Masaki Fujihata, 1994
A 3-D map of Mount Fuji was distorted by the GPS data gathered from Fujihata's ascent (middle) and descent (bottom) from the mountain. The decreased climbing speed near the summit is reflected in extreme distortion at the peak.

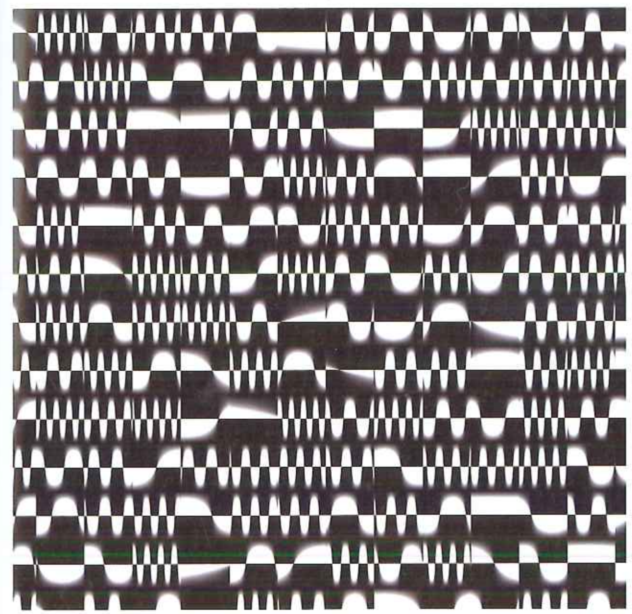
Travel Time Tube Map, by Tom Carden, 2005
This map reimagines the draftsman Henry Beck's classic Underground Map as a malleable space. It warps to show the time between stations, rather than simply their order. When a station is selected, it becomes the center of the diagram and all of the other stations are organized into concentric rings showing travel times from that place.

Superformula,
by David Dessens, 2008
Dessens has been using the superformula since 2006 as a basis for his dynamic visual compositions and animations. The graphic forms depicted in this image were generated using that equation.

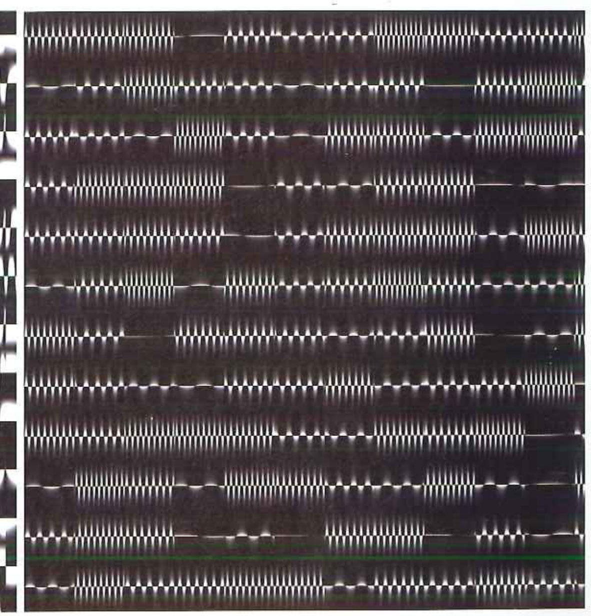
VISUALIZATION TECHNIQUE
MATHEMATICS VISUALIZATION



Images were created to think about mathematics long before computers were invented to calculate and visualize. For instance, Euclid (circa 300 BCE) constructed diagrams to show relationships between geometric elements and physical models. A paper model of a Möbius strip and a glass model of a Klein bottle can make these intriguing surfaces approachable to a wider audience, beyond those who understand the equations behind them. Before the era of the personal computer, the Mathematica exhibition, created by the Eames Office in 1961, presented diagrams, objects, and machines to demystify basic mathematical principles. Mathematica is also the name of a powerful program used within the sciences for calculations and visualizations. This program, along with other software development projects, cleared the path to new categories of mathematics visualization, including the popular fractal images of the Mandelbrot set. Mathematicians, artists, and architects are actively mining the structures of numbers and equations to produce visual images for pleasure and insight.



EPF:2003:V:A:997141,
by Kenneth A. Huff,
2003
Using the properties of prime numbers to determine the base structure, Huff adds the dimensions of depth, texture, and lighting to create fantastic forms. A different prime number determines the length of each unique segment.



Algorithmic Visualizations,
by George Legrady,
2002–5
These images are created from mathematical equations that have their origins in image-processing algorithms. Legrady "shapes" and "massages" the equations to affect their visual expression.