Information Visualisation

Overview | Information design | 49
Web design | Dynamic displays
Statistics | Browsers
Communication
Navigation

Modern-day map-making may be the best way our of Web design's stasis

By Nico Macdonald

pages 52-57

In 1569 the Flemish cartographer Gerardus Mercator published a map of the world in which parallels and meridians were rendered as straight lines at right angles, spaced so as to produce at any point on a map an accurate ratio of latitude to longitude. While he never travelled beyond northern Europe, Mercator's work in developing visual presentation helped to change the perception of the new world that was being explored. For the budding mercantile classes of the low countries, mapping the earth made it possible for new industries and modes of commerce to emerge.[1]

Over the past 50 years we have created another new world using micro-processors and networks. Digital technology – from cameras to keyboards, sensors to measurement devices – allows us to gather unprecedented amounts of data. Computers manipulate, store and present this data on a gigantic scale; the Internet gives its users access at high speed and low cost.

Society has become increasingly complex. Businesses co-ordinate more resources in more places 'just-in-time'. Financiers monitor companies, industries, markets and commodity prices in real time. Researchers and scientists analyse everything from chemical compounds to geological data and the human genome in increasingly pressured environments. Governments at the centre of more dynamic populations need to plan everything from housing to transportation, healthcare to employment.

Yet, most of us still try to comprehend this glut of data using representations from the era of print – many based on text rather than image. Not only that, we have been viewing them on monitors that display less information than a letter-size sheet of paper. There is a clear danger of 'information overload' or maybe 'understanding underload' – a failure to focus, to prioritise, and to apply wisdom to information. Our ability to present information in a useful and intelligent manner is falling behind our ability to create and distribute the raw data.

For the designer, this is the challenge of 'information visualisation', the process of revealing useful, easily grasped insights by transforming abstract data into visual and manipulable forms . In other words: making new maps for the new, information-rich universe.

Evolution has left humans with brains that are as much visual as they are analytical, able to distinguish and group objects by size, colour, shape and spatial location. Our brains are adept at identifying patterns. 'Millions of leaves on thousands of plants – and we can still pick out the ripe bananas', notes designer Brad Paley (see TextArc sidebar, p.57).

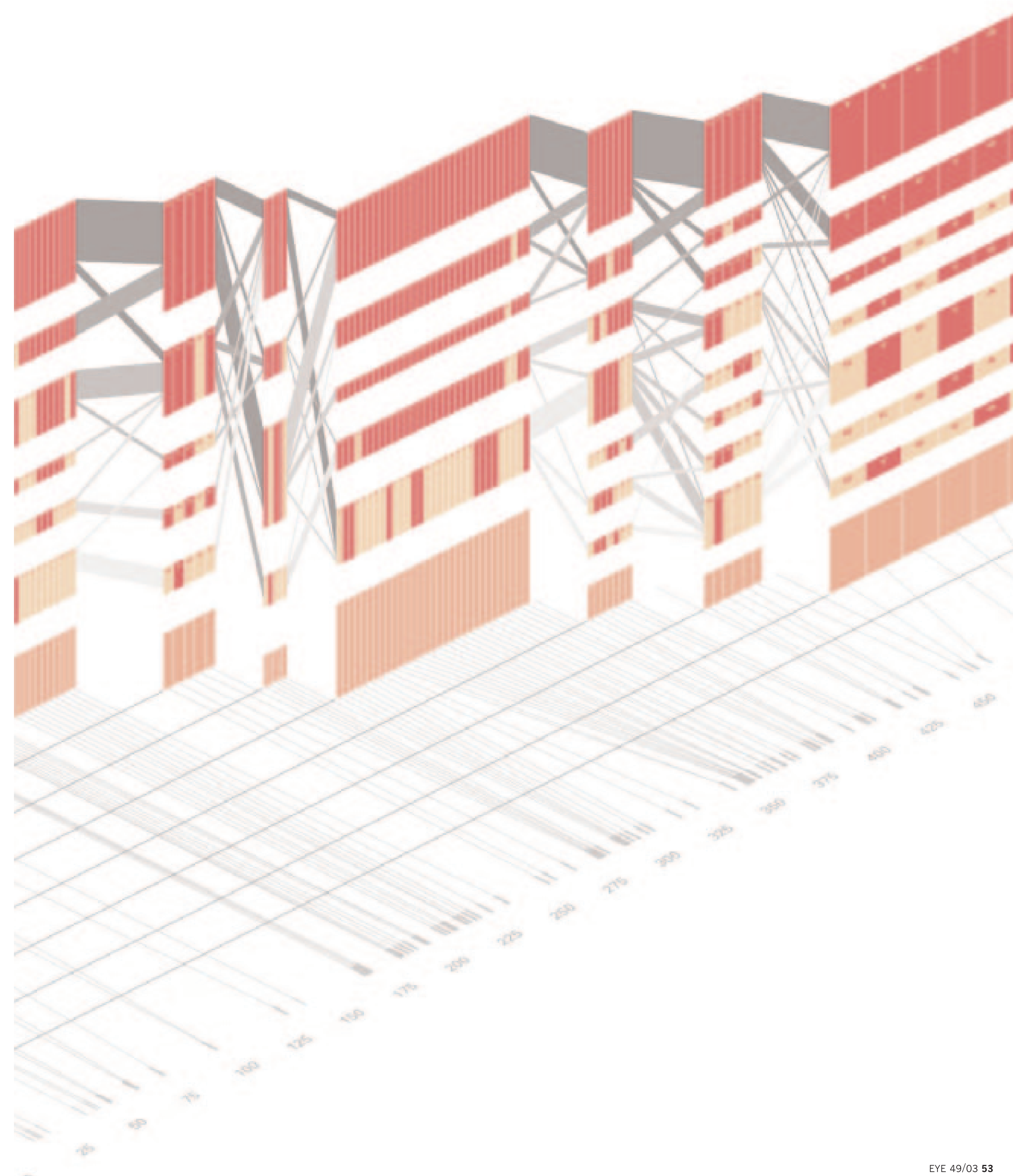
Information visualisation taps into these powerful cognitive skills in ways that textual and numerical information can't. A visualisation is more universally readable than any single language, and it can be 'read' more quickly. A good example might reveal multiple dimensions of information, present trends and patterns, and show information in context. Or show the qualitative state of complex systems 'at a glance', allowing for quick identification of status and changes in the system.

The concept is not new. From Mercator through Minard (and the other lesser heroes celebrated by Edward Tufte) via Colin Ware[2] to contemporary protagonists such as Richard Saul Wurman[3] the concept of visualisation has been fervently evangelised – but only in print. In the digital world its development has been slow, though the idea of graphical displays, epitomised by Ivan Sutherland's work at MIT and the University of Utah, dates back to the early 1960s. Sutherland's successors in the field include Xerox PARC veterans Stu Card and Jock Mackinlay and University of Maryland HCI Lab founding director Ben Shneiderman.[4]

The graphical user interface has been with us for some time, joined ten years ago by the graphical Web browser, but in these cases 'graphical' refers to the use of visual techniques for task execution and information navigation. Actual information is usually presented as text or a static graphic. Though there is a need to develop information visualisation techniques and

1. Genetic data, DESIGNERS?, YEAR?
The Human Genome Project has been addressed as a subject for information visualisation in a project entitled Genomic Cartography. This visualisation compares a section of 'haplotype' genetic data for 500 people. Vertical size is proportional to the frequency with which a sequence is found in the group, while the horizontal blocks represent a grouping of related sequences of letters. MORE DETAIL?

1. The Mercator map, centred around the Greenwich Meridian, shaped many people's view of the world for centuries. In the 1960s, geographer Arno Peters created a map, the Peters Projection, in which each country's surface area was represented more accurately.
2. US-based UK émigré and author of *Information Visualization: Perception for Design.*
3. Richard Saul Wurman (see *Eye* no. 28 vol. 7) created the pioneering Access map series, edited *Information Architects* and initiated Understanding USA.
4. Stu Card, Jock Mackinlay and Ben Shneiderman co-edited *Readings in Information Visualization: Using Vision to Think* (Morgan Kaufmann, 1999). More recently, Shneiderman co-authored *The Craft of Information Visualization: Readings and Reflections* (Morgan Kaufmann, 2003).



2. AT&T Labs SWIFT-3D visualisation, DESIGNERS, YEAR. A ‘landscape’ visualisation of phone network activity in the US. MORE DETAIL, PLEASE.

3-8 (opposite). Kryptästhesie in Vernetzen Systemen [semantic search tool] by Martin Grothmaak, Jürgen Späth, 1999 (www.projekttriangle.com). The dynamic research tool ‘Cryptaesthesia in networked systems’ looks for information in big databases or the Internet, evaluates it and displays it geometrically. The dynamic model illustrates both the content-based relationships between search criteria and the generation of search results. The results appear not as a list but as data clouds in the form of points in a circle around the central search word. BRIEF EXPLANATION OF ‘DIGGING DOWN’ PROCESS?

models, several factors have limited development and there are many misconceptions – often based on the print heritage.

Complexity demands that we devise new ways of seeing data. ‘There are a large number of conventions about visualising information, especially in business’, notes Paul Kahn, giving financial data as an example: ‘The typical data visualisations – bar charts, pie charts – become so conventional that they more or less look the same and hence are less useful’. He believes that the ‘Map of the Markets’ tool created by *SmartMoney* magazine (see sidebar, p56), represents the kind of progress that is needed. This was a landmark, combining a cognitively powerful visualisation with elegance and aesthetics to create a high quality experience.

‘People think of data visualisation as an insert into an application’, observes Mark Schindler of Visual I/O. ‘This is where Tufte stops – he focuses on static ways of representing data that is often dynamic’. Schindler, who describes traditional data visualisation as the translation of one set of facts to another, argues for an approach in which data can be seen in context from a number of different vantage points. User interaction with data presenting results in real time – ‘direct manipulation’ according to Shneiderman’s terminology – is a powerful element of information visualisation.

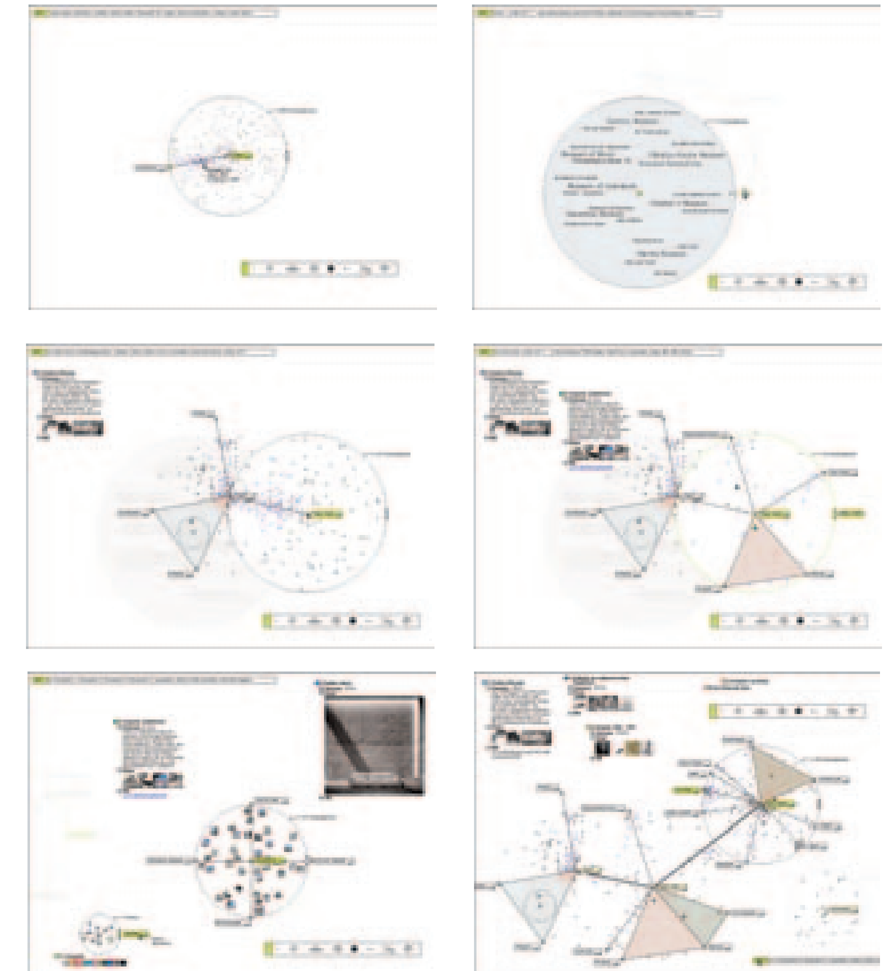
Three-dimensional visualisation is another area that escapes the limitations of print. One recent initiative is an interface for the knowledge mapping tool Antartica, which presents a three-dimensional interface to its geographical metaphor for information organisation. Kahn sees value in the use of an extra dimension for highlighting information, but regrets the way that 3D navigation ‘undermines your understanding of what you are looking at and where it is’.

So why should clients bother with information visualisation? Szeto sees the value for organisations, noting that analysis – of marketplace dynamics, consumption patterns, and the like – is a core business activity and clients need to ‘figure out if the business is performing on plan or off plan and why, and to validate past decisions and support future ones’. An article in the *Economist* observes that ‘after rounds of lay-offs, companies have fewer people to take complex decisions – a shortage that better software tools can help to alleviate.

In the area of research and development, information visualisation can be the determining factor in success. Yet data analysis is not keeping pace with technological advances in other areas. The Human Genome Project is an example of the challenge of understanding large amounts of data.

Bob Quinn of Battelle, owner of visualisation tool developer OmniViz, believes that data visualisation is a way for business to gain competitive advantage: since many companies receive exactly the same data at the same time, ‘the winners will be those who can quickly transform raw data into sources of high value information’. Another area of business that can benefit is communication. Kahn sees a role for information visualisation in sharing knowledge by ‘helping to create a common mental model among groups of people who need to understand it’.

So if information visualisation is an appropriate solution to a design problem, the data is there, and the client and stakeholders are convinced of its value, how are designers to proceed? They need to begin by understanding the nature of the information, where it comes from, how and where it is used, what insights users will be looking for and their relative importance. Financial traders will be more interested in seeing changes in existing data relationships (see Thomson Financial example), while researchers may wish to find new data relationships. The needs of beginner and expert users may also



SPOTFIRE

Many areas of government, business and research create enormous volumes of data: from censuses to through financial information to genomic research. Computers and networks allow us to create and exchange this data, while our greater understanding of the world produces more phenomena to explore, and collect data about. How can we make use of unimaginably large datasets?

At the University of Maryland Human-Computer Interaction Lab, Dr Ben Shneiderman set his students the task of describing a possible interface for finding a film from a library of 10,000 videotapes.

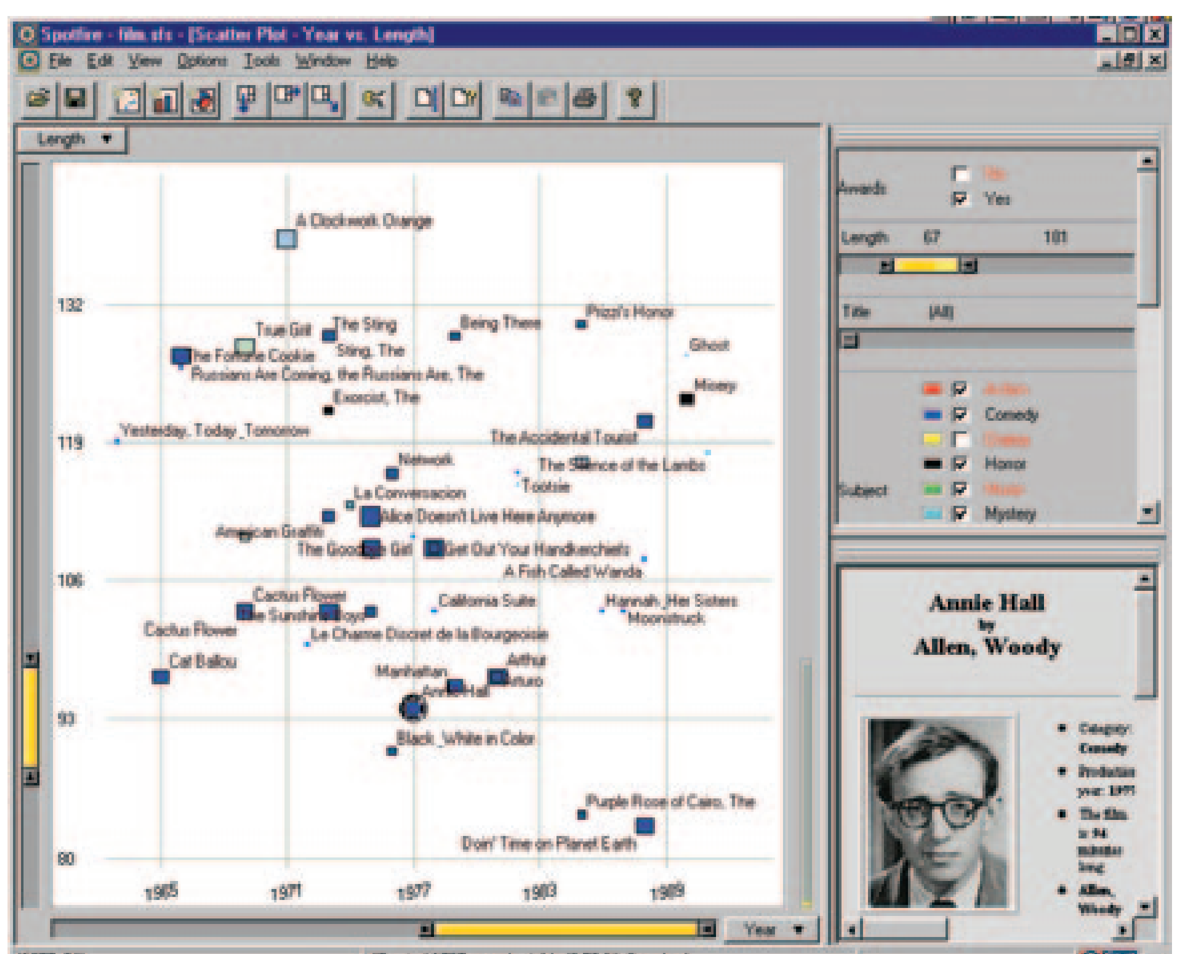
One of the solutions developed from this is Chistopher Ahlberg’s Film Finder (right), which has a range slider that allows filtering by each film’s length and year of release. Shneiderman calls this ‘direct manipulation’. Buttons allow for selection by ratings and awards.

The results are displayed as a ‘starfield’ that allows the user to narrow down the selection and see emerging patterns. A click on one of the spots displays a box with details of each film and a picture of one of the actors.

There are many areas of activity where the winnowing of similarly large datasets is critical. One area is the pharmaceutical industry, where researchers need to narrow the scope of a large number of compounds and represent the results clearly. A dataset of compounds may be refined by molecular weight, stability, charge, particular properties and by many other criteria to leave a found set that would be realistic to test further, or move on to the next stage of drug development.

Spotfire, the company founded by Ahlberg, now counts sixteen of the twenty major pharmaceutical companies among its clients.

www.spotfire.com



VISUAL I/O

As part of their development of an information visualisation system for a pharmaceutical research and development organisation Visual IIO developed an example of their approach that addressed a stop / go decision with data to support a historical analysis. They chose the task of deciding whether to pull a pitcher from a baseball game (right).

The data for the example is drawn from game statistics. A dynamic ‘opinion gathering’ interface allows users to adjust weightings using direct manipulation.

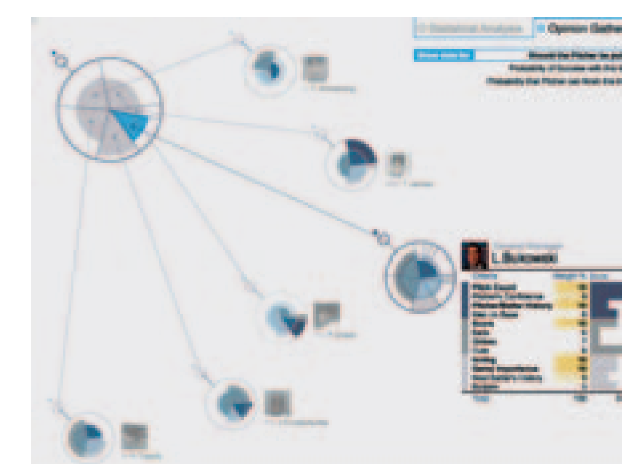
The central element of the interface is a circular diagram that shows an aggregate ‘finding’ of whether the pitcher should be removed from the game in this circumstance. Each ‘slice’ of the circle represents the data from one of the criteria. The size of the slice represents its weighting, and the radius shows the value of that factor.

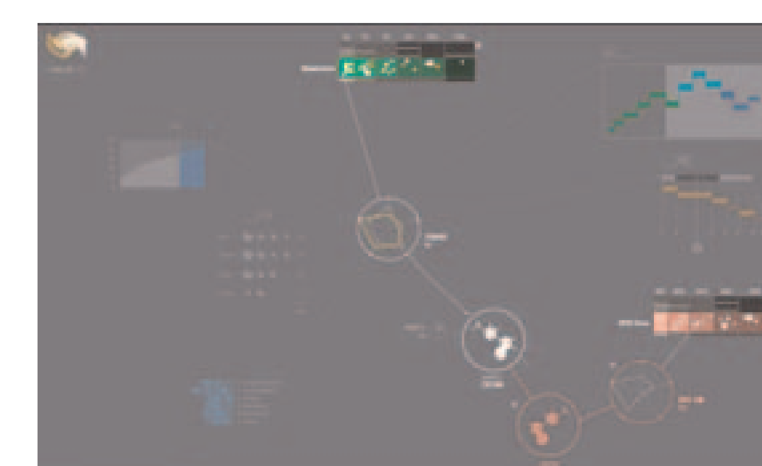
The criteria are grouped into shades of grey and there is a median represented by the light blue transparency. Yellow is used to indicate that a tranche is being edited. The more filled in the circle, the stronger the suggestion that the pitcher should be pulled.

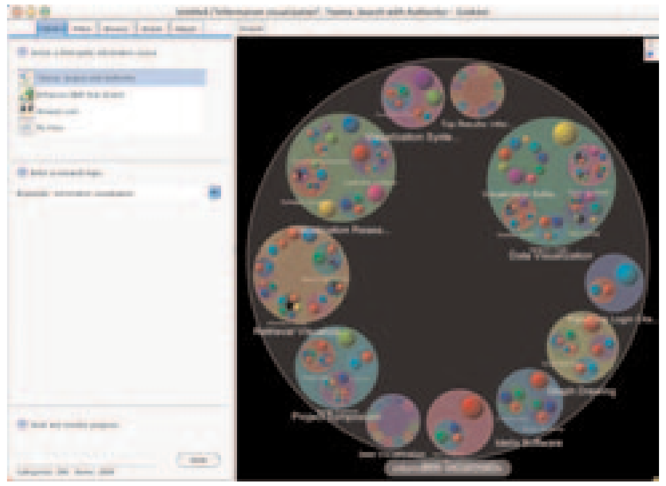
This demonstration tool, which was created in Flash, can be found at:

www.manifesto9.com/baseball

NEED COPY TO EXPLAIN
VisualIO_PharmaInterface.jpg







9. Grokker, DESIGN CREDIT, YEAR, presents users with richer ways of browsing large volumes of data. MORE DETAIL?? www.groxis.com

10-12. Thomson Workspace delivers financial products via the Web. Creative director: Rodney Edwards. Lead User Interface Design: Franco Riccitelli. Technical architect & software engineer: Javier Garcia Flynn. Interaction Design: Keith Marsh. This re-creates an application environment in a standard Internet browser by transcending typical browser capabilities and building on users' familiarity with operating system tools such as Windows, Frames, Dialogs, Menus, Folders, and interaction techniques such as drag and drop and direct manipulation.
10. Typical Workspace layout display showing financial information. Here, the Chart window has been selected.
11. User has selected the 'Insert' menu in order to select new content types.
12. The user can drag and drop a financial instrument from one window to another. In this instance the user is dropping the HSBC instrument into the Chart window to compare performance against another stock in the same sector – Lloyds TSB.

need to be considered, particularly if training will be limited or some people will only use the tool infrequently .

Issues of perception, cognition and visual psychology need to be taken into account when considering how to present, contextualise and prioritise information, and support its manipulation. These relate to the number of dimensions of data presented, spatial positioning and relationships, colour and tone, and dynamic changes in these elements. At the level of interface and human-computer interaction, the designer needs to analyse the tasks to be supported, and address comprehensibility, navigability and the effectiveness of any solution, including its usability. There may be a visual metaphor around which an interface can be structured. Visual I0's Angel Shen-Hsieh develops metaphors by 'discerning the mental picture that expert decision-makers have.'

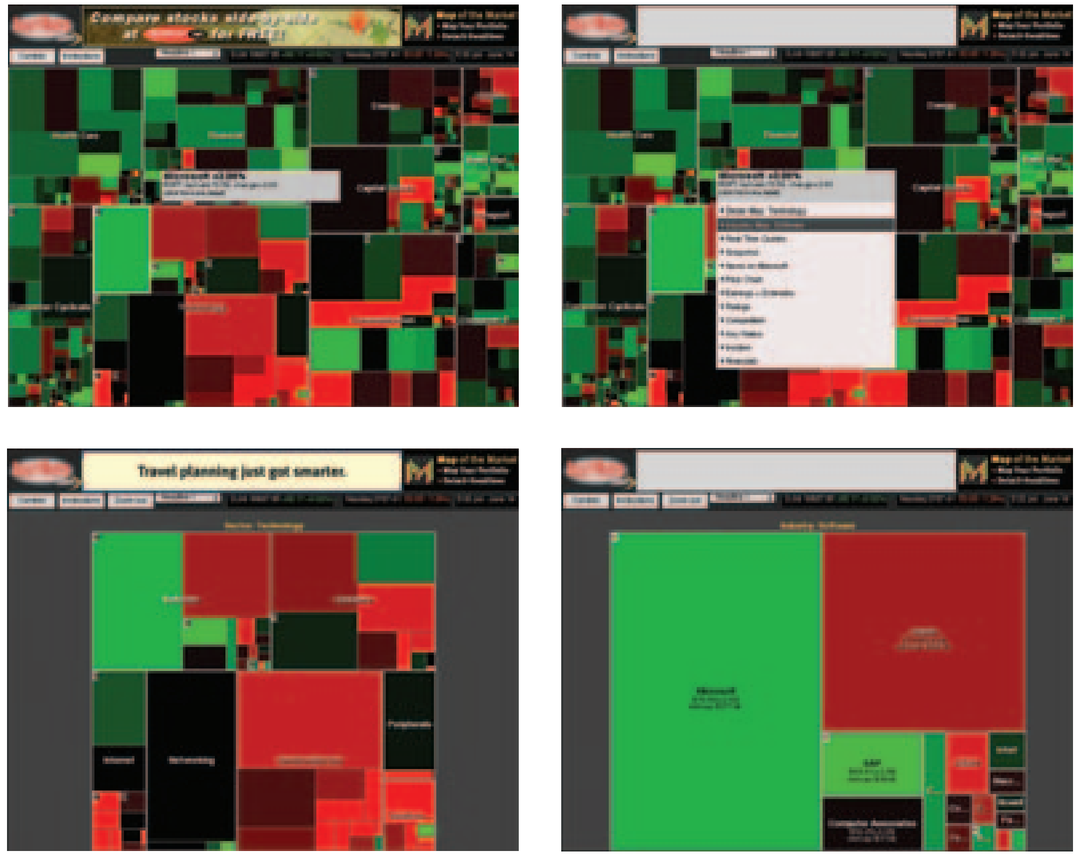
Developing interface concepts may begin with sketching but will likely require more sophisticated prototyping as, unlike many areas of design, sketches don't provide a good basis for user (or even heuristic) testing. It is also particularly important to work with the people who manage the data the tool will draw from, and those responsible for its implementation, to ensure the data can be delivered in the form it is needed and that the tool will perform quickly and reliably. The designer will also need to address the overall quality of experience of the product, considering visual appeal, information design and typography and interaction style.

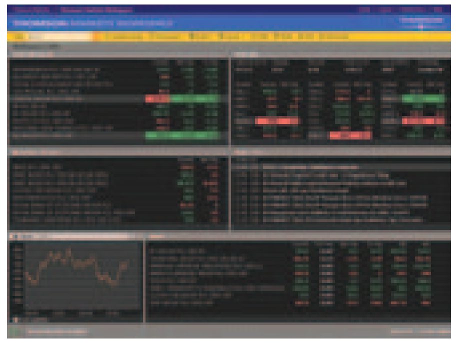
Yet information visualisation is not well served by tools that address these challenges: designers find that they have to custom-create or re-combine tools for anything other than simple interfaces. Clients also need tools to input, structure and annotate information at the back end – a design deliverable that is often forgotten on even the most mundane projects.

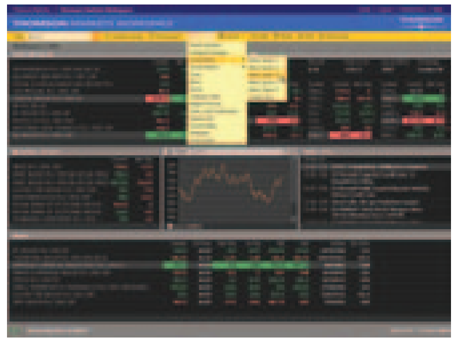
TREEMAP

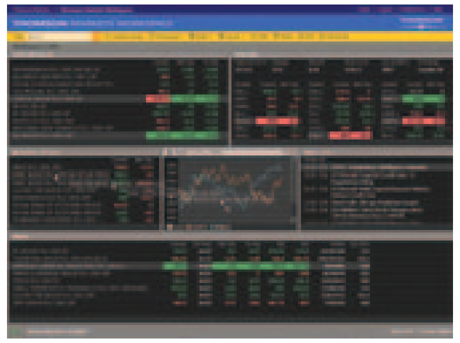
Map of the Market is based on the self-organising map model. In this model, also known as a TreeMap, a fixed area is sub-divided by rectangles whose areas are in proportion to the data they represent, allowing for an efficient use of space, with colour tones differentiating time-based change. This implementation shows 535 popularly held stocks, organised by industry groups, size-coded by market capitalisation and colour-coded to show rise or fall of market capitalisation.
The TreeMap concept was developed by Brian Johnson and Ben Shneiderman. A former researcher at the University of Maryland HCI Lab, Martin Wattenberg, now at IBM Research, brought the concept to *SmartMoney* and worked with them to implement it.
Paul Kahn notes that Map of the Market 'is a rare marriage of complex computer science [the self-organising map algorithm] with a well defined domain (top stocks in specific industries) and a well defined change over time (rise and fall of value).'

www.smartmoney.com/marketmap









Many of the developments that will enable information visualisation to become as commonplace in the future as the graphical user interface is today are either technological or infrastructural. Higher quality, cheaper and more portable displays will allow the combination of the qualities of paper and digital media. Pervasive networks will allow people to get to information where and when they need it. And structured languages will allow data from disparate sources to be combined for new and practical uses.* Others factors involve a greater commitment to the idea: on the part of clients to invest in the right solutions; and on the part of designers in developing skills in information visualisation, and in communicating its value to clients. Focusing on one industry sector's needs, and doing a lot of testing and evaluation, presents one way forward, as Spotfire creator Chris Ahlberg found with pharmaceutical research (see sidebar, p.54).

The Web design industry appears to have arrived at a fixed set of solutions that preclude real innovation. Meanwhile, businesses and other organisations have moved from using computers as document processing and basic communication devices to deploying them for enterprise management, logistics and manufacturing control, business and financial analysis, realtime decision-making, and strategic planning. If designers and engineers could use information visualisation to do for these organisations what the spreadsheet did for them over twenty years ago, design would be taken more seriously as a business resource and the spin-off benefits to society would be tremendous. They may do for the twenty-first century what Mercator did for the sixteenth. But if we get to this point, don't expect everyone to be talking about 'information visualisation': it will become the norm, barely recognised. We will merely marvel that people ever had to work in any other way.

TEXTARC

TextArc is described by Brad Paley, its creator, as a 'visual index / concordance', originally developed to help people get their bearings in text documents. The entire text under analysis (see example *The Communist Manifesto*, right) is arranged in a clockwise arc, starting at twelve o'clock. Words appear brighter according to their frequency, and their location is determined by averaging their position between their appearances in the text. A word that appears more often at the end of a text – 'social' in this case – is drawn towards the region of nine o'clock.
Colour is used to communicate information rather than for simple aesthetics. When a line of text from the outside is selected it appears in an attention-getting orange, and its constituent words turn a quieter green. When an individual word is selected, other forms of the word, for instance those with the same root, are picked out less-prominently.
The principle behind TextArc is that the reader can understand the structure of a text, get a sense of its character and pace, and begin to identify its key ideas and the style of writing. It provides a sideways look at familiar work, and has attracted some attention from the world of literature studies.
'I wanted it to show *all* of the words in a document, letting the reader decide which words and concepts were important', writes Paley. 'And I wanted to show the distribution of words in a document, so people would know where to go to read up on a certain idea, or where it was first introduced, or how relevant it was to the entire document.'

www.TextArc.org