

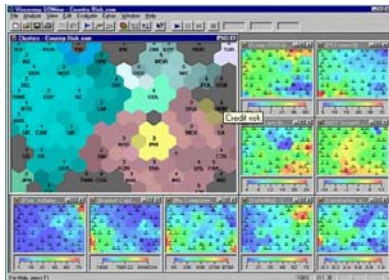
INFORMATION VISUALISATION THROUGH THE MARKET

With the development of Self Organizing Maps by Teuvo Kohonen from 1982 onwards, the method has been used in many data mining applications. Because of its proven efficiency and success, it has been preferred as a popular solution to a lot of problems in the data analysis field. In this review, commercial tools using SOM and then other tools related to “document clustering and visualization” using other methods will be explored.

1. Commercial Tools using SOM in visualization and data mining

1.1 *Viscovery@SOMine*

The developer:	Eudaptics Software Company
Homepage:	http://www.eudaptics.com
Required platform:	Windows NT (4.0 or higher) and for Windows 95/98
Memory required:	Not valid
Space required:	Not valid
Pricing:	Standard Edition: US\$ 289, Enterprise: US\$ 1,995
Technology Paradigm:	Kohonen SOM, Ward, SOM-Ward

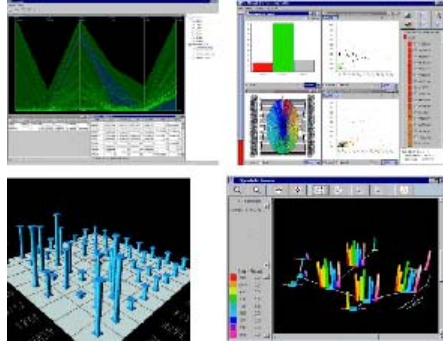


The tool provides:

1. Interactive display
2. File Import/Export
3. Process Administration
4. User Support

1.2 *Kensington Discovery Edition*

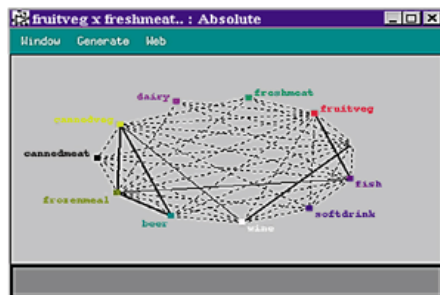
The developer:	Inforsense Software Company (spin-off from the long-established Data Mining Research Group, affiliated to the Parallel Computing Centre at Imperial College of Science, Technology and Medicine)
Homepage:	http://www.inforsense.com
Required platform:	Platform independent (Java client)
Memory required:	-
Space required:	-
Pricing:	Not valid
Technology Paradigm:	Decision Tree-Based Classification, Bayesian Classification with Parallel AutoClass, Parallel SOM for Clustering, Back-Propagation Neural Networks Association Rule Discovery



The tool provides 2D and 3D interactive visualization running on different databases, spreadsheets, free text, Web documents, XML documents and various document management systems. Also it supports information retrieval, context knowledge management and dynamic data warehousing. It has the capability of reporting.

1.3 Clementine

The developer:	SPSS Company
Homepage:	http://www.spss.com/clementine/
Required platform:	NT, Unix
Memory required:	40 MB
Space required:	65 MB
Pricing:	\$ 43,000
Technology Paradigm:	Kohonen Neural Networks, Regression, back-propagation, ID3/C5.0 Machine Learning Algorithm



It supports various data analysis methods. Also it provides import/export data capability, reporting, data pre-processing, modelling, visual programming.

1.4 IBM Intelligent Miner

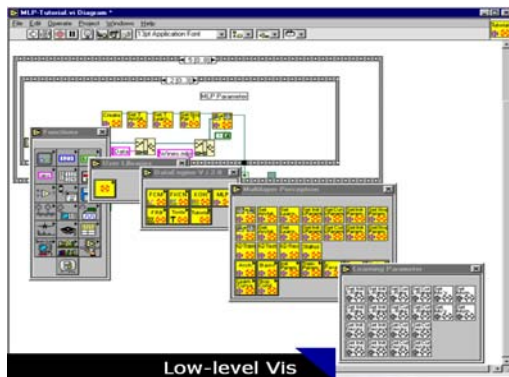
The developer:	IBM Company
Homepage:	http://www-4.ibm.com/software/data/iminer/
Required platform:	AIX, OS/390, OS/400, Solaris, Windows 2000, Windows NT and z/OS Apart from it, DB2 must be installed
Memory required:	128 MB (for WinNT)
Space required:	100 MB (for WinNT)
Pricing:	\$ 60,000
Technology Paradigm:	Kohonen Neural Network, decision trees, various classification- association- sequence discovery- regression methods

ova.controlled	Similar	Docc
controlled.numeric	100.0%	0.1%
control.frame	Similar	Docc
database.main	Similar	Docc
database.sub	100.0%	0.1%
database.update	Similar	Docc
memory.unit	Similar	Docc
function.peripheral	72.3%	0.2%
sequence.unit	Similar	Docc
function.membership	Similar	Docc
fuzzy.logic	11.3%	0.7%
fuzzy.inferred	Similar	Docc
rule.element	Similar	Docc
compute.fault	11.8%	0.4%
compute.kolmogor	Similar	Docc
color.data	Similar	Docc
color.color	10.7%	0.4%
look.table	Similar	Docc
output.signal	Similar	Docc
signal.unit	0.9%	0.7%
input.signal	Similar	Docc

It's tightly coupled with DB2 database but it supports input resources from ASCII. Besides this, pre-processing and data warehouse tools must be obtained individually.

1.5 DataEngine

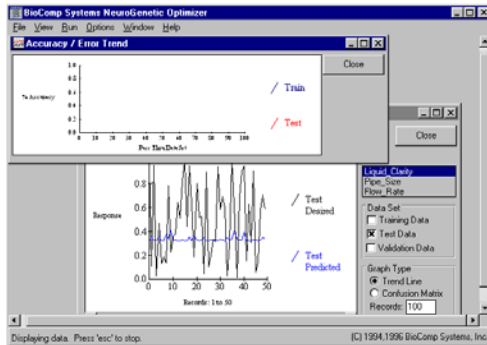
The developer:	M.I.T(Management Intelligenter Technologien GmbH)
Homepage:	http://www.dataengine.de
Required platform:	Server(Windows2000), Client (Win95/98- WinNT)
Memory required:	32 MB
Space required:	47 MB
Pricing:	Server (30,000€), Client(5,000€)
Technology Paradigm:	Neural networks (Kohonen), statistical methods, fuzzy technologies



The tool provides pre-processing of the data, data acquisition and visualization. It's especially designed for clustering. It contains an extensive list of various mathematical and statistical functions.

1.6 *NGO NeuroGenetic Optimizer*

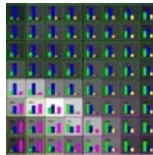
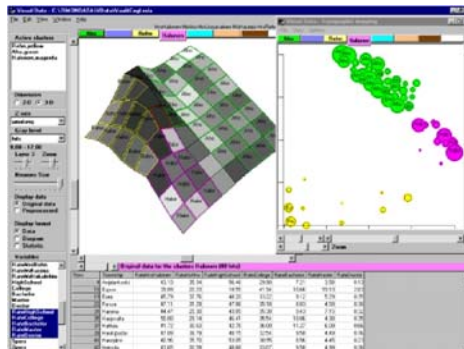
The developer:	BioComp Systems, Inc.
Homepage:	http://www.bio-comp.com/
Required platform:	IBM Compatible Personal Computer (Windows '95, '98 or NT 4.0)
Memory required:	32 MB
Space required:	10MB
Pricing:	\$295 and up
Technology Paradigm:	Neural Networks(SOM and Temporal SOM, fast Back Propagation), Probabilistic, Generalized Regression, Genetic Algorithms



The major attribute of the tool is the ability of multiple modelling simultaneously. They optimise the neural networks by using genetic algorithms.

1.7 *Visipoint*

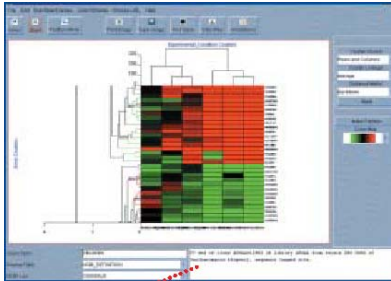
The developer:	Visipoint Company
Homepage:	http://www.visipoint.fi
Required platform:	Win95/98 – NT
Memory required:	Not valid
Space required:	Not valid
Pricing:	Not valid
Technology Paradigm:	Kohonen SOM



The tool supports the input from any text or spreadsheet format. It provides pre-processing and different visualization representations except SOM such as bar, histograms.

1.8 *GeneSight*

The developer:	BioDiscovery, Inc.
Homepage:	http://www.biodiscovery.com/
Required platform:	Windows 95/98/NT4/2000
Memory required:	256MB
Space required:	5GB
Pricing:	Not valid
Technology Paradigm:	SOM, hierarchical- K means clustering, time series analysis, Principal Component Analysis



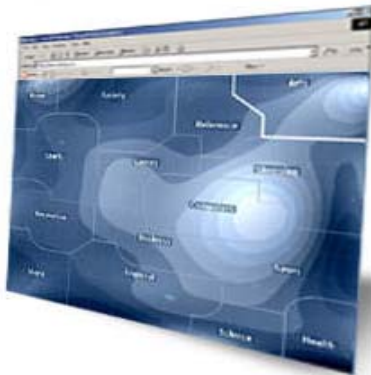
It provides various visualization methods like bar, histogram, scatter points, GenePie plots. It provides import/export files and pre-processing.

2. Commercial Tools using other methods in information visualisation

Because there is a huge amount of information visualization methods, we'll consider only web based document visualization tools here.

2.1 *WebMap*

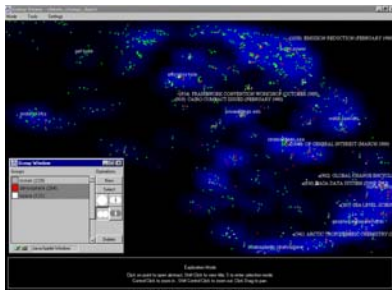
The developer:	WebMap Technologies
Homepage:	http://www.webmap.com
Required platform:	Platform independent
Memory required:	-
Space required:	-
Pricing:	Free
Technology Paradigm:	Information mapping technology



It's a built-in component in the browsers. It displays fixed size categories and with an interactive map you can browse these categories. It provides a search mechanism with the application.

2.2 *WebTheme*TM

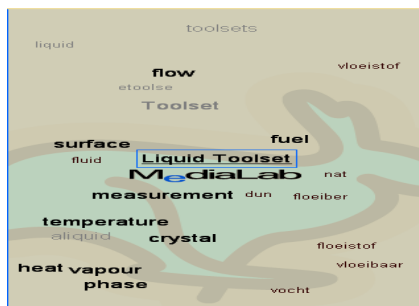
The developer:	SPIRE-Spatial Paradigm for Information Retrieval and Exploration
Homepage:	http://www.pnl.gov/infoviz/spire/spire.html
Required platform:	Plattform independent
Memory required:	-
Space required:	-
Pricing:	Not valid (still under development)
Technology Paradigm:	Built in SPIRE general technology, information retrieval and clustering methods are used.



The tool is being developed by the team of “Pacific Northwest National Laboratory”. It’s designed as an interactive map and has the capability of document browsing and search.

2.3 *Zoef*

The developer:	MediaLab and Headline Communication
Homepage:	http://www.zoef.nl/
Required platform:	Platform independent
Memory required:	-
Space required:	-
Pricing:	Not valid (under development)
Technology Paradigm:	Fuzzy systems

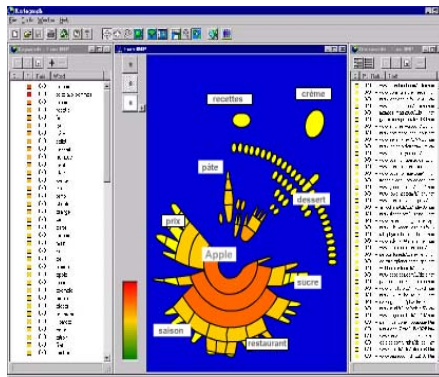


It’s based on MediaLab’s previous technology called Liquid Technologies. According to it, the application reads concept space from a file and displays the information with focus on the screen. The map is interactive so that the user can give focus to a different concept.

2.4 Umap Web

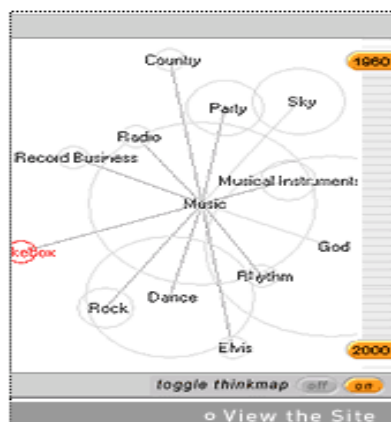
The developer:	Trivium Company
Homepage:	http://www.umap.com/highuk/index.htm
Required platform:	Windows 95, 98 or NT 4.0
Memory required:	32 MB
Space required:	25 MB
Pricing:	Free
Technology Paradigm:	Information retrieval and extraction methods

After the user submits her query, the program extracts all the keywords from generated web sites and creates an interactive map from these results.



2.5 Thinkmap

The developer:	Thinkmap Inc.
Homepage:	http://www.thinkmap.com/
Required platform:	Windows and Linux
Memory required:	128 MB
Space required:	40 MB
Pricing:	ThinkMap Studio 1,185\$, Application Server 4,995\$, Application Server 5,095\$
Technology Paradigm:	Information retrieval and extraction methods



The tool provides presentation capabilities, searching and browsing of the documents. The circles represent documents and the size of the circles shows the number of the documents belong to that keyword. Also it has a directory structure displaying the documents and selected document content.

