Seventh Framework Programme
Research Infrastructures
e·**nventory** - European eInfrastructures Observatory
Grant Agreement no. RI-261554

Deliverable D4.1

Indicators visualisation and representation modelling and techniques

Authors:
Maria Roussou
Dimitris Nastos
Vivi Katifori
Yannis Ioannidis
Nektaria Berikou
Nikos Vogiatzis
Jorge-A. Sanchez-P.

Status –Version:
Final – 3.0

Date:
March 22, 2011

Distribution - Type:
Public

Abstract: Deliverable D4.1 "Indicators visualisation and representation modelling and techniques" compiles and describes several information visualisation and representation techniques for data indicators, in accordance with reference/good examples implemented internationally. Although primarily a "technical" deliverable, it is meant to assist the e·**nventory** consultation activities with a selection of appropriate visualisation engines and interface front-ends for the **European eInfrastructures Observatory**.

Document Revision History

Date	Issue	Author/Editor/Contributor	Summary of main changes / Status
October 4, 2010	1.0	M. Roussou	ToC for approval
October 11, 2010	1.1	M. Roussou	First working draft of the deliverable circulated and presented for comments at consortium meeting #7
December 20, 2010	1.2	M. Roussou	Working draft of the deliverable circulated and presented for comments at consortium meeting #10
January 30, 2011	1.3	M. Roussou	Working draft of the deliverable circulated and presented for comments at consortium meeting #11
February 21, 2011	1.4	M. Roussou, D. Nastos, Y. Ioannidis	Working draft of the deliverable circulated
February 25, 2011	1.5	V. Katifori, M. Roussou	Additional information on case studies
February 27, 2011	1.6	N. Vogiatzis	Cross-reading; editorial improvements; updated executive summary, intro, conclusions, Removed former chapter 3. Revised text in Chapter 2 and 3.
February 28, 2011	1.7	J.-A. Sanchez-P., N. Vogiatzis, N. Berikou, M. Roussou, D. Nastos	Approval of deliverable contents at consortium meeting #13
March 2, 2011	1.8	M. Roussou	Final additions and corrections
March 7, 2011	2.0	J.-A. Sanchez-P.	Final version submitted to the EC
March 20, 2011	2.1	QAT	Recommendations proposed by QAT
March 22, 2011	3.0	J.-A. Sanchez-P., M. Roussou	Incorporation of QAT recommendations

Table of contents

List of figures

List of tables

References

[1] *Annex I – Description of Work*, **e·nventory** project, grant agreement 261554, June 2010.

[2] *Quality assurance plan and risk assessment*, D1.1, **e·nventory** project, October 2010.

[3] *Dissemination and communication plan*, D5.1, **e·nventory** project, December 2010.

[4] *International observatories, repositories and indicators study*, D2.1, **e·nventory** project, January 2011.

[5] *European eInfrastructures Observatory Advisory Board charter*, D3.1, **e·nventory** project, February 2011.

[6] Ahlberg, C., Shneiderman, B., *Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays,* proceedings ACM CHI'94: Human Factors in Comp. Systems, pages 313-317, 1994.

[7] Bederson, B., Shneiderman, B., *The Craft of Information Visualisation: Readings and Reflections,* Morgan Kaufmann Publishers, 2003.

[8] Card, S., Mackinlay, J., Shneiderman, B., *Readings in Information Visualisation: Using Vision to Think*. Morgan Kaufmann Publishers, 1999.

[9] DeFanti, T., Brown, M., McCormick, B., *Visualisation: expanding scientific and engineering research opportunities,* IEEE Computer22, pages 12-25, 1989.Last retrieved at http://www.cs.panam.edu/~rfowler/vis-papers/DeFanti_1987_visualisation-expanding.pdf

[10] Dix, A., Ellis, G., *Starting Simple - adding value to static visualisation through simple interaction*, Proceedings of Advanced Visual Interfaces - AVI98, L'Aquila, Italy, ACM Press. pp. 124-134, 1998. Last retrieved at http://www.comp.lancs.ac.uk/~dixa/papers/simple98.

[11] Keim, D., *Information visualisation and visual data mining*, IEEE Transactions on Visualisation and Computer Graphics, 7(1), pages 100-107, 2002.

[12] Friendly, M., Denis, D., *Milestones in the history of thematic cartography, statistical graphics, and data visualisation*, 2001. Last retrieved athttp://www.datavis.ca/milestones.

[13] Geisler, G., *Making Information More Accessible: A Survey of Information Visualisation Applications and Techniques*, 1998. Last retrieved athttp://www.ischool.utexas.edu/~geisler/info/infovis/paper.html.

[14] Shneiderman, B., Plaisant, C., *Designing the User Interface: Strategies for Effective Human-Computer Interaction,* 5th ed., Boston, MA: Addison-Wesley, 2009.

[15] Owen, S., *History of Visualisation*, 1999. Last retrieved at http://www.siggraph.org/education/materials/HyperVis/visgoals/visgoal3.htm.

[16] Spence, R., *Information Visualisation: Design for Interaction,* 2nd ed., p. 304, Prentice Hall, 2007.

[17] Tufte, E. R., *The Visual Display of Quantitative Information,* Connecticut: Graphics Press, 1983.

[18] Tufte, E. R., *Envisioning Information,* 4th ed., p. 126, Graphics Press, 1990.

[19] Tufte, E. R., *Visual Explanations: Images and Quantities, Evidence and Narrative,* p. 156, Graphics Press, 1997.

[20] Tufte, E. R., *Visual & Statistical Thinking: Displays of Evidence for Decision Making,* p. 32, Graphics Press, 1997.

[21] Tufte, E. R., *The Visual Display of Quantitative Information,* 2nd ed., p. 197, Graphics Press, 2001.

[22] Ward, M. O., Grinstein, G., Keim, D., *Interactive Data Visualisation: Foundations, Techniques, and Applications,* p. 513, A K Peters Ltd, 2010.

[23] Ahlberg, C., Shneiderman, B., *Visual Information seeking using the FilmFinder*, ACM CHI 1994, pages 433-434, 1994.

[24] Tweedie, L., Spence, B., Williams, D., Bhogal, R., *The Attribute Explorer*, ACM CHI 1994, pages 435-436, 1994. Last retrieved at http://ieg.ifs.tuwien.ac.at/~aigner/teaching/ws04/infovis_ue/aufgabe3/references/p435-tweedie.pdf

[25] Bier, E., Stone, M., Pier, K., Buxton, W., DeRose, T., *Toolglass & Magic Lenses: The See-Through Interface*, ACM CHI 1994. Last retrieved at http://www2.parc.com/istl/projects/MagicLenses/93Siggraph.html.

[26] Yi, J., Melton, R., Stasko, J., Jacko, J., *Dust & Magnet: multivariate information visualisation using a magnet metaphor*, Information Visualisation, palgrave Macmillan Ltd., pages 239–256, 2005.Last retrieved at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.157.6563&rep=rep1&type=pdf

[27] Smith, G., Czerwinski, M., Meyers, B., Robbins, D., Robertson, G., Tan, D. S., *FacetMap: A scalable search and browse visualisation*, IEEE transactions on visualisation and computer graphics, *12*(5), 797-804, 2006.

[28] Ahlberg, C., *Spotfire: an information exploration environment*, *ACM SIGMOD Record*, *25*, pages 25-29, 1996.

[29] Williamson, C., Shneiderman, B., *The Dynamic HomeFinder: Evaluating dynamic queries in a real estate information exploration system*, Proceedings ACM SIGIR Conference, pages 339-346, 1992. Last retrieved at http://hcil.cs.umd.edu/trs/92-01/92-01.html.

[30] *Timelines and Visual Histories*. Last retrieved at http://www.math.yorku.ca/SCS/Gallery/index.html.

[31] *The Digital History web project*. Last retrieved at http://www.digitalhistory.uh.edu/timeline/timelineN.cfm

[32] *Human Development Report*, UNDP. Last retrieved at http://hdr.undp.org.

[33] *European Cluster Observatory*. Last retrieved at http://www.clusterobservatory.eu.

[34] *Data Visualizer*, World Bank. Last retrieved at http://devdata.worldbank.org/DataVisualizer.

[35] *The World at a Glance*, World Bank. Last retrieved at http://data.worldbank.org.

[36] *International Institute for Sustainable Development*. Last retrieved at http://iisd.ca.

[37] *The SIMILE project*. Last retrieved at http://www.simile-widgets.org/timeline.

[38] *Interactive stacked histograms*. Last retrieved at http://www.meandeviation.com/dancing-histograms/hist.html.

[39] *Excentric labelling from Maryland*. Last retrieved at http://www.cs.umd.edu/hcil/excentric.

[40] *Visual Thesaurus*. Last retrieved at http://www.visualthesaurus.com.

[41] *Interactive house finder*. Last retrieved at http://www.housingmaps.com.

[42] *Cellphone chooser with dynamic queries*. Last retrieved at http://www.myrateplan.com/cellphones.

[43] *Diamond, finder with dynamic queries*. Last retrieved at http://www.bluenile.com/diamond-search?track=dss.

[44] Gary Klass, *Just Plain Data Analysis*. Last retrieved at http://lilt.ilstu.edu/jpda.

Acronyms

3D	Three dimensional
ECO	European Cluster Observatory
GIS	Geographic Information Systems
HCI	Human Computer Interaction
HDI	Human Development Index
InfoVis	Information Visualisation
NYT	New York Talk Exchange
UCD	User-Centred Design
WEF	World Economic Forum
WWW	World Wide Web

Preface

The **e·nventory** project targets the formation of the **European eInfrastructures Observatory**, a single-entry-point and one-stop-shop data warehouse, capable of representing multiple primary and convoluted indicators and benchmarks, and a yardstick tool for progress monitoring, impact assessment, post-mortem analysis and ex-ante evaluation of eInfrastructures at regional and national levels across the European Union and beyond.

The aim of the **e·nventory** project is to carry out a design study that will set the grounds towards the **European eInfrastructures Observatory**; through the collection and utilisation of appropriate indicators, the project will be able to monitor the status quo and evolution over time of eInfrastructures development and communicate all findings to related stakeholders but also to the public-at-large, in a seamless and impartial way.

The **e·nventory** project will carry out extensive consultation with eInfrastructures stakeholders and research and innovation indicator experts and will extend prior benchmarking efforts (e.g. EARNEST) by including an extensive set of eInfrastructures components (i.e., computing, communication and services), eventually deploying a prototype web platform dealing with a critically-selected subset of indicators, through intuitive, interactive and user-friendly mappings, plots and graphics.

The project action plan is structured so that it can achieve:

- identification of a core set of **benchmarking indicators** for the **European eInfrastructures Observatory** that will be the baseline for **monitoring** eInfrastructures development progress,
- eInfrastructures **stakeholders' feedback and consensus** on the proposed structure and functionality of the **European eInfrastructures Observatory**, and
- **European eInfrastructures Observatory** functionality **demonstration through a prototype web platform** that will be available to all eInfrastructure communities and to the general public.

The **e·nventory** project responds to the need of eInfrastructure stakeholders by delivering a prototype tool to aid impact assessment of related eInfrastructure initiatives and programmes. In that respect, the **user communities** that will effectively use the **European eInfrastructures Observatory** as a yardstick tool and translate the project outputs into a real influence on eInfrastructure policy, are the **eInfrastructure stakeholders** themselves, including:

- the **European Parliament**, the **European Commission** and **National Governments** that are sponsoring eInfrastructure initiatives and place a high value in their policy agenda on the impact assessment, post-mortem analysis and ex-ante evaluation, to help design better and more successful future eInfrastructure programmes;
- **research funding bodies** that plan the development of strategies to address the digital divide and digital opportunity issues and need to utilise and assess the impact of eInfrastructures contributing to this goal;
- **eInfrastructure policy bodies** that have a clear mandate of supporting the development and the sustainability of eInfrastructures and need an impact assessment tool to monitor the progress of achieving their objectives;
- **eInfrastructure projects** that carry out specific eInfrastructure activities and need to monitor their impact during deployment as well as retrospectively;
- **scientific/research communities** that are empowered by eInfrastructures and are eager to comprehend their strengths and weaknesses that influence their everyday work.

High utilisation of the **European eInfrastructures Observatory** from the broad eInfrastructure community is a key measure of the project's impact; therefore, the entire project workplan has been designed to maximize **stakeholders' engagement** and awareness.

The **e·nventory** project kicked-off in September 2010 and is planned to be completed by August 2012. The project is co-financed by the European Commission's Seventh Framework Programme for Research Infrastructures.

Executive summary

What is the focus of this Deliverable?

The objective of WP4 is to create the necessary tool/web platform, albeit at a pilot/prototype level, through which access will be given to the **European eInfrastructures Observatory**. The portal will allow for various representation and statistical data aggregation methods, including graphs, distributions, and other visualisation means that will enable accurate and user-friendly utilisation of the **e·nventory** data repository.

Deliverable D4.1 compiles and describes several information visualisation and representation techniques for data indicators, in accordance with reference/good examples implemented internationally. Although primarily a "technical" deliverable, it is meant to assist the **e·nventory** consultation activities with a selection of appropriate visualisation engines and interface front-ends for the **European eInfrastructures Observatory**.

What are the deliverable contents?

Deliverable 4.1 provides the foundation on information visualisation, including its definition but more importantly its principle features: representation and interaction. Chapter 2 presents selected representation techniques related to the **e·nventory** work. Chapter 3 presents the most notable interaction techniques. Then, Chapter 4 reviews reference case-studies of information visualisation available on the web. The choice of the reference case-studies has been made to reveal interactive visualisation features adopted by data observatories that exhibit certain similarities to the **e·nventory** approach.

What are the deliverable main conclusions?

Several international data observatories have adopted information visualisation techniques that provide their global users with multiple interactive options. In some cases, the visualisation tools available on-line are highly interactive and visually appealing. However, many of these tools are also too complex for a broader public and, consequently, suffer from usability problems. Furthermore, the interactive visualisation of eInfrastructures-related data, i.e. of data concerning grids and supercomputers, is sparse.

This document is the first pillar of **e·nventory** towards creating the necessary tool/web platform through which access will be given to the **European eInfrastructures Observatory**. The review of international data observatories representation and interaction techniques and usability will help shape the design and development (i.e. the functional specifications) of the **European eInfrastructures Observatory**'s pilot application in visualising eInfrastructure indicators online. This document provides a smooth transition to the follow-up **e·nventory** web platform specifications.

How does this deliverable contribute to the project quality metrics?

This deliverable sets the ground for achieving the following project quality metrics:

- *At least 2 use cases supported by the prototype web platform, i.e. at least 2 different user-group needs addressed through the interactive visualisation of indicators developed.*

How does this deliverable contribute to the avoidance/mitigation/exploitation of the project risks?

This deliverable contributes to the avoidance, mitigation or exploitation of the following project risks:

- *Project high-level objectives not clear and/or agreed within the consortium*
- *Not accomplishing effective interactive web visualisation of the information space*
- *The on-line interactive visualisation tool not being user-friendly enough*

What is next in the process to deliver the e·nventory results?

- Deliverable D2.2 (M06) documents the first set of the **European eInfrastructures Observatory** benchmarking indicators that will be used as input for the next phases of the project, especially as the baseline for discussion for the consultation activities in WP3.

- Deliverable D3.2 (M09) will collect the **European eInfrastructures Observatory** **Advisory Board's** initial assessment and directions upon concluding a first round of consultation process with the Advisory Board. This will serve as an intermediate input for the **e·nventory** workplan.
- Deliverable D4.2 (M12) will document the first set of draft specifications for the prototype implementation of the web platform, namely its architecture, web engine, data collection, storing, and updating mechanisms, as well as its visualisation and representation methodology.
- Deliverable D5.2 (M12) will report (mid-project) on the implementation of dissemination and outreach activities, including assessment of their impact on the targeted eInfrastructure groups
- Deliverable D5.3 (M14) will present an update of the dissemination plans for the project as well as a potential revision of the project communication package.
- Deliverable D3.3 (M15) will document, after concluding the entire consultation process, with the **European eInfrastructures Observatory** **Advisory Board**, the final reflections that will serve as a key milestone and final input for the **e·nventory** workplan.
- Deliverable D2.3 (M16) will provide the final set of the **European eInfrastructures Observatory** benchmarking indicators after the finalization of consultation activities in WP3.
- Deliverable D4.3 (M20) will document the final set of specifications for the prototype web platform dealing with a critically-selected subset of indicators and relevant repositories, its functionality and the user's guide. It will further include the results of the usability and accessibility tests that will be conducted with the platform. This deliverable will also serve as a technical reference for any follow-up (production-level) version of the portal after the completion of the project.
- Deliverable D5.4 (M23) will take the form of an inauguration workshop that will demonstrate the **European eInfrastructures Observatory** to representatives of the eInfrastructures community at large.
- Deliverable D5.5 (M24) will provide a final report on the implementation of dissemination and outreach activities, including assessment of their impact on the targeted eInfrastructure groups.

1. Information visualisation: definition and purpose

1.1. Visualisation

Visualisation is any technique for creating images, diagrams, or animations to communicate a message, reinforce cognition, hypothesis building and reasoning. Visualisation through visual imagery has been an effective way to communicate both abstract and concrete ideas since the dawn of man. Numerous definitions of the term visualisation have been proposed in the past decades [8], [14].

The act of transforming data – in whatever form - into pictures is not new, as exemplified by the age-old art of geographic mapmaking. A few thousand years after the first maps on clay, mapmaking evolved into the construction of "thematic maps", i.e. cartographic and statistical data maps [17], [18].

Computers, however, and the processing power and potential for real-time interactivity that they bring, have expanded the techniques for the visual depiction of data. The recent emphasis on visualisation started in 1987 with the special issue of Computer Graphics on Visualisation in Scientific Computing. Since then there have been several conferences and workshops, devoted to the general topic, and special areas in the field, for example volume visualisation.

Visualisation today has ever-expanding applications in science, education, engineering (e.g. product visualisation), interactive multimedia, medicine, etc.

Scientific visualisation is a well-established field with a specific focus on data that describe physical objects and scientific simulations or experiments with an implicit or explicit geometric structure, to allow the exploration, analysis and understanding of the data[9], [13]. It is a way to visualise large sets of scientific data that are otherwise difficult to see and manipulate, such as molecular structures, and to enhance scientists' ability to see phenomena in the data, such as the flow of air over wings. It's a very important part of visualisation and maybe the first one, as the visualisation of experiments and phenomena is as old as science itself. Traditional areas of scientific visualisation are flow visualisation, medical visualisation, astrophysical visualisation and chemical visualisation.

Educational visualisation is using a simulation normally created on a computer to create an image of something so it can be taught about. This is very useful when teaching about a topic which is difficult to otherwise see, for example, atomic structure, because atoms are far too small to be studied easily without expensive and difficult to use scientific equipment.

Information visualisation, sometimes called InfoVis, refers to the use of interactive visual representations of abstract non-physically based data to support the discovery and interpretation of information and, consequently, the human's ability to gain insight and make decisions. The term "information visualisation" was originally coined by the User Interface Research Group at Xerox PARC.

Information visualisation, is a rather recent advancement (which has seen a meteoric rise in the last 15 years) with broader applications in areas such as digital libraries (collections of documents), data mining, financial data analysis, market studies, manufacturing production control, drug discovery [7], educational data, even crime mapping[1]. In these cases, the data represents non-physical, abstract information without any obvious spatial mapping.

The principal task of information visualisation is to allow *information to be derived from data*. Spence in [16] emphasizes this distinction between data and information. Data can be of many different types: numerical or quantitative data (e.g. temperatures, indicators), ordinal data, or categorical data. The modern problem is that of data overload not information overload. Spence considers information visualisation to be essentially the formation of a mental model of data, thus emphasizing the role of human performance and perception.

Hence, the purpose of information visualisation techniques and applications is to enable a wide variety of computer users to more easily navigate *information spaces*, to better display retrieved information, and to improve their understanding of information.

In other words, information visualisation is useful to the extent that it increases our ability to discover, make decisions, and explain or, to put it more succinctly, "*the purpose of information visualisation is insight, not pictures*" [8].

---

[1] http://en.wikipedia.org/wiki/Crime_mapping.

Information visualisation applications rely on basic features that the human perceptual system inherently assimilates very quickly: colour, size, shape, proximity, and motion. These features can be used by the designers of information systems to increase the data density of the information displayed [18].

Because we perceive such features so readily, and because each feature can be used to represent different attributes of data, good visualisation enables us to not only perceive information more easily but also to *perceive more information at one time*. We can immediately see patterns in data that indicate trends, recognize gaps in the data, discover outliers or errors in the data, pinpoint minimum and maximum values, and identify clusters. As a result, information visualisation applications enable us to better understand complex systems, make better decisions, and discover information that might otherwise remain unknown.

Practical application of information visualisation involves selecting, transforming and representing abstract data in a form that facilitates human interaction for exploration and understanding. Important aspects of information visualisation are dynamics of visual representation and the interactivity. Strong techniques enable the user to modify the visualisation in real-time, thus affording unparalleled perception of patterns and structural relations in the abstract data in question.

Many of these techniques have recently been used in interactive information visualisation applications that can be found on-line and which are presented in this document. The following chapters will present the methods and techniques commonly used to represent and interact with information visualisations.

1.2. Representation and interaction

Numerous techniques have been used to map data to visual form, to lay it out in available and suitable display areas and time (**representation**) and to then allow the user to control the visual representation, to move from one view of data to another (**interaction**). These techniques should present the user with information in a way in which the global, high-level structure is apparent and yet the low-level detail is still accessible.

Representation is what attracts all of the attention, yet interaction is where the action is. It is very challenging to come up with innovative, new visual representations but a lot of interesting work can be done with how the user interacts with the view or views. This is what distinguishes information visualisation from static visual representations on paper. In other words, the principal feature that separates modern information visualisation tools than, say, static graphs or maps is interaction, i.e. the facility to change our view of the available data.

Card et al. in [8] have proposed a simple reference model to show the mappings between data, their visual form and the human perceiver (Figure 1).

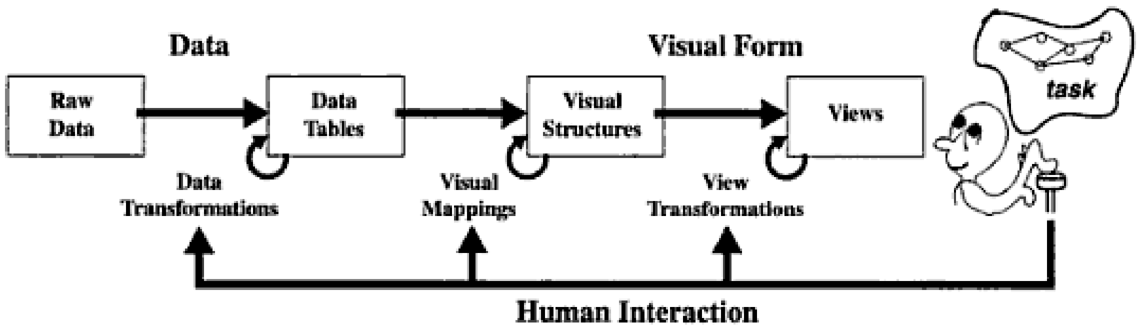

Figure 1: Reference model of the mappings of data, their visual form and the human perceiver

The reference model approximates the basic steps for visualising information. Data (on the left) is eventually transformed into Visual Form for the human (on the right) through a series of data transformations:

- Firstly, the **Raw Data** (i.e. data in some idiosyncratic format) is mapped to **Data Tables**, which are relational descriptions of data and metadata.
- **Visual Mappings** then transform the tabular data into **Visual Structures** (i.e. structures that combine spatial substrates, marks, and graphical properties). This is the actual visualisation process, or the **Representation stage**.
- Finally, **View Transformations** are used to create **Views** of the **Visual Structures** by specifying parameters such as position, scaling, and clipping.

These transformations are controlled by the user through **Interaction** with the visualisation (e.g. pressing a button, moving a slider, selecting a geographical region, etc). Interaction to create Views essentially allows to increase the amount of information that can be visualised or to restrict the Views to certain data ranges.

1.2.1. Raw data and data types

Raw data comes in many forms, from numbers to the text of novels. For this project, a distinction is made between the usual numerical data and indicators, which represent the kinds of data **e·nventory** will be handling.

A definition for **indicators** is given by the International Institute for Sustainable Development: "An indicator quantifies and simplifies phenomena and helps us understand complex realities. Indicators are aggregates of raw and processed data but they can be further aggregated to form complex indices"[36].

Examples of traditional indicators or indices used in measuring the social, economic and environmental welfare are: the gross national income (Figure 2), unemployment rates, price indices and life expectancy.

Research in information visualisation has produced a wide variety of applications [13]. The look and function of an application is determined by the type of user tasks the visualisation is designed to support, the kind of data being visualised, and the way the application uses basic features such as colour, size, shape, proximity, and motion to represent data.

While there are several ways to survey this range of information visualisation applications, perhaps the most fundamental method is to categorize them by the underlying data types that the applications attempt to visualise.

The relation between two or more items of data is referred to as *structure* (e.g. the use of coloured lines to represent the connection between stations on a London Underground map). The dimensions or attributes of data that must be represented is what defines the data's complexity. Shneiderman in [14] identifies:

- univariate data or one-dimensional (where only one attribute is involved)
- bivariate data or two-dimensional
- trivariate data or three-dimensional
- hypervariate data or multi-dimensional

1.2.2. Data tables

The power of information visualisation stems to a considerable degree from our ability to take raw data and represent it (present it again) in a different way with the aim of informing a user.

The first step in the visualisation process is to transform this data into a relation or set of relations that are more structured and thus easier to map to visual forms.

This is commonly done by organising the data in tables, or data tables [8]. Data tables may include, in addition to the raw data, descriptive information about the data, otherwise known as *metadata*.

Although raw data can be visualised directly, data tables are an important intermediate step when the data are abstract, without a direct spatial component. The use of tables is pervasive throughout numerous disciplines as they provide a familiar way to convey information that might otherwise not be obvious or readily understood. One of the advantages of data tables is that they clearly depict the number of variables associated with a collection (Figure 2).

1.3. Presentation -usability

Apart from how data is represented, a decision has to be made about how that representation is to be displayed to the user, i.e. how it will be offered for view by a user, how it will be *presented*. In the case of interactive systems and applications, presentation pertains to the interface between the human and the digital world; it is this interface through which the user of an interactive information visualisation tool will experience the representations (visual structures) and interaction capabilities of the tool.

Presentation may be influenced by the task that a visualisation tool is designed to support or by the context (in Figure 3, fuel tanks for presenting the price of gasoline in the past decade, an office for Facebook, etc). However, contemporary Human-Computer Interaction (HCI) practice dictates that the way an interface is designed should directly take into account the needs and goals of the users that will be using it.

Gross national income per capita

US dollars, current prices and PPPs

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Australia	20 729	21 569	22 743	23 909	25 389	26 525	27 521	28 770	30 255	31 212	32 640	34 113	35 938	36 897
Austria	23 135	24 161	24 623	25 723	26 529	28 285	28 187	29 854	30 775	32 347	33 083	34 920	36 100	37 256
Belgium	22 808	23 220	24 295	24 777	25 826	28 260	28 951	30 417	30 704	31 484	32 350	33 656	34 997	35 523
Canada	21 936	22 512	23 704	24 700	26 217	27 740	28 500	29 145	30 497	32 121	34 377	36 451	37 963	38 593
Czech Republic	12 787	13 455	13 613	13 702	13 975	14 655	15 633	16 103	17 215	18 240	19 452	20 743	22 316	22 875
Denmark	22 747	23 744	24 902	25 844	26 699	28 216	29 027	30 393	30 238	32 444	33 659	35 839	36 699	37 323
Finland	18 173	18 775	20 607	22 127	23 373	25 493	26 584	27 643	27 478	30 063	30 813	32 907	35 329	35 837
France	20 260	20 915	21 890	22 966	23 990	25 623	26 965	27 853	27 568	28 493	29 884	31 120	32 840	33 309
Germany	22 326	22 935	23 408	23 999	24 873	25 709	26 592	27 246	28 364	30 182	31 738	33 602	35 390	36 017
Greece	15 044	15 504	16 393	16 859	17 160	18 462	20 058	21 655	22 570	23 977	24 224	25 787	26 981	27 947
Hungary	8 738	9 073	9 551	10 241	10 651	11 560	12 917	14 009	14 705	15 462	16 027	16 882	17 429	18 407
Iceland	22 570	23 609	25 501	27 225	28 071	28 046	29 492	31 033	30 282	32 323	33 674	32 309	33 035	22 515
Ireland	16 161	17 749	19 423	21 279	22 304	24 717	25 795	27 422	29 501	31 273	33 164	35 873	37 997	35 581
Italy	20 787	21 526	22 437	23 563	24 091	25 406	26 953	26 594	26 912	27 253	28 056	29 467	30 795	30 774
Japan	22 586	23 774	24 572	24 296	24 557	25 935	26 593	27 252	27 965	29 581	31 027	32 843	34 759	35 258
Korea	13 286	14 344	15 111	13 978	15 491	17 131	18 132	19 668	20 198	21 694	22 762	24 699	26 623	27 839
Luxembourg	35 969	37 325	39 182	39 633	43 897	46 516	47 893	47 726	47 060	56 760	58 668	58 806	65 342	63 978
Mexico	7 196	7 628	8 256	8 645	9 028	9 811	9 926	10 216	10 696	11 376	12 260	13 193	13 936	14 305
Netherlands	21 872	22 844	24 417	25 215	27 226	30 044	31 026	32 236	32 059	34 092	35 280	38 173	40 165	39 983
New Zealand	16 042	16 318	17 168	17 705	18 625	19 355	20 291	21 028	21 668	22 494	22 897	23 968	24 982	24 997
Norway	23 310	25 746	27 669	27 105	29 550	35 643	37 131	37 166	38 532	42 331	47 646	52 079	54 189	59 253
Poland	7 375	8 065	8 805	9 403	9 940	10 530	10 925	11 524	11 867	12 653	13 523	14 342	15 727	16 900
Portugal	13 045	13 545	14 264	14 954	15 843	16 668	17 278	18 065	18 537	18 865	20 255	20 886	21 801	22 346
Slovak Republic	8 345	9 091	9 753	10 324	10 340	10 912	12 061	12 909	12 911	14 056	15 706	17 402	19 734	21 545
Spain	15 895	16 546	17 535	18 705	19 638	21 143	22 230	23 703	24 458	25 608	26 991	29 145	30 743	30 648
Sweden	21 355	22 127	22 907	23 919	25 546	27 523	27 696	28 905	30 330	31 990	32 249	34 903	37 674	37 780
Switzerland	27 376	28 190	29 916	31 110	31 961	33 946	33 588	34 469	35 774	36 994	38 822	41 107	42 338	39 735
United Kingdom	19 561	20 799	22 358	23 527	24 141	26 026	27 747	29 315	30 251	32 240	33 272	34 298	35 432	36 259
United States	27 520	28 881	30 467	32 024	33 652	35 659	36 415	37 012	38 322	40 605	43 091	45 610	46 867	47 320
EU27 total	17 329	18 044	18 895	19 737	20 501	21 810	22 923	23 825	24 424	25 717	26 845	28 375	29 920	30 511
OECD total	19 554	20 479	21 522	22 290	23 197	24 683	25 440	26 144	26 924	28 469	30 024	31 815	33 216	33 748
Chile	..	7 822	8 370	8 686	8 566	8 950	9 382	9 583	9 847	10 381	11 115	11 396	12 311	13 299
Estonia	6 283	6 923	7 753	8 319	8 632	9 541	10 259	11 475	12 676	14 046	15 902	17 487	19 174	19 402
Israel	16 301	17 035	17 290	17 787	20 461	21 935	22 433	22 579	21 360	22 850	23 149	24 974	26 726	27 448
Russian Federation	5 550	5 443	5 577	5 225	5 680	6 634	7 259	7 884	8 984	10 043	11 569	12 811	..	..
Slovenia	13 081	13 834	14 792	15 621	16 647	17 482	18 385	19 593	20 288	21 940	23 295	24 543	26 007	27 222

Figure 2: Raw data (in this case indicators) presented in a data table form[2]

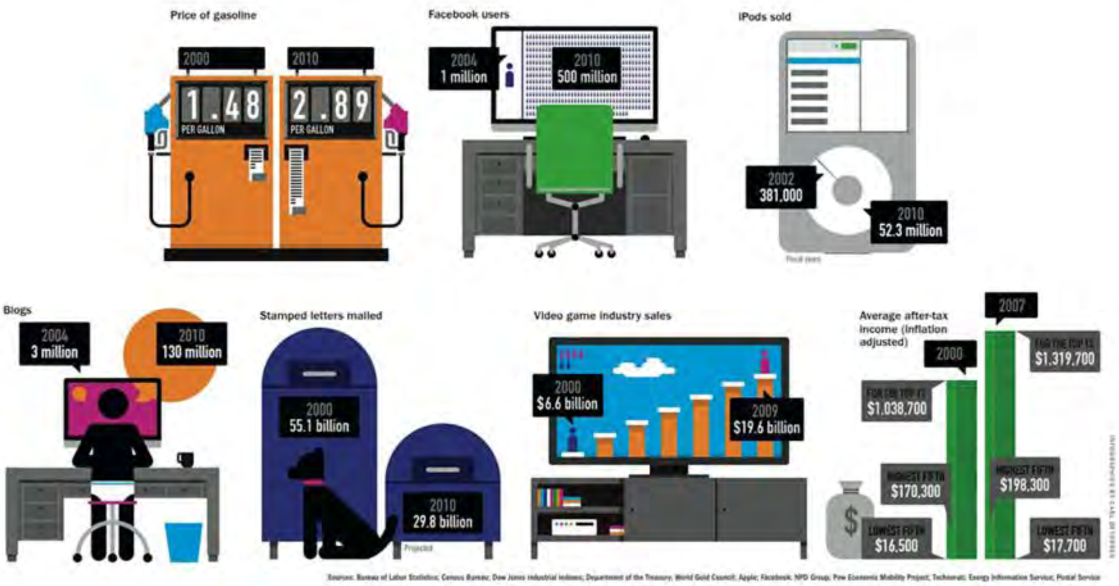

Figure 3: A creative way of displaying simple data by showcasing the context[3]

[2]http://www.oecd-ilibrary.org/economics/oecd-factbook_18147364

[3]In Time Magazine, 24 November 2010.

This approach to designing interfaces is named **user-centred design (UCD)** and relates to a process in which the needs, wants, and limitations of end users of a product are given extensive attention at each stage of the design process.

The first step to achieving a UCD approach is by performing a user requirements analysis, i.e. an analysis of the potential users of the system either through discussion with the potential users themselves and/or with others (who work with them, who are affiliated to them, etc.). Collecting user needs then leads to the formation of functional specifications, i.e. a list of the functionality required by the system to accomplish the goals of the project and the potential needs of the users.

For **e·nventory**, a number of target user groups have already been identified in [3]. These groups (e.g. policy makers, student researchers, the management of eInfrastructure policy bodies, etc.) may have different needs and goals, which could possibly require “adjustments” or even different renditions of the interface in order to accommodate them.

2. Representation techniques

Representation concerns the graphical presentation or portrayal of data, in other words, the mapping of data to visual form. The purpose of the data graphical form is to present the data, thus it is an instrument to help people reason about information and to facilitate inferencing from data.

The core of the reference model proposed in Figure 1 is the mapping of a data table to a visual structure. There are a number of different ways in which values in a data table can be encoded visually, e.g. by using colours, lines, dots, crosses, etc.

According to Tufte [17], to be effective, the graphical display of data should:

- Show the data;
- Induce the viewer to think about the substance rather than about methodology, graphic design, the technology of graphic production, or something else;
- Avoid distorting what the data have to say;
- Make large data sets coherent;
- Encourage the eye to compare different pieces of data;
- Reveal the data at several levels of detail, from a broad overview to the fine structure;
- Serve a reasonably clear purpose: description, exploration, tabulation, or decoration;
- Be closely integrated with the statistical and verbal descriptions of a data set.

Visual structures are made from some kind of spatial positioning, marks (the visual points, lines, areas, volumes, etc. that occur in space), and the marks' graphical properties. The most fundamental aspect of a visual structure is its use of space, which is described in terms of axes and their properties. For example, the following types of axes are elementary: nominal axis for names, such as, for example thematic categories; ordinal (categorical) axis for data with ordering, such as the days of the week; quantitative axis (Q) for numerical data, such as a percentage.

In a nutshell, effective representation of data is the well-designed presentation of interesting data – a matter of substance, of statistics, and of design. Effective representation consists of complex ideas communicated with clarity, precision, and efficiency. At their best, visual structures are instruments for reasoning about quantitative information. Often the most effective way to describe, explore, and summarize a set of numbers – even a very large set - is to look at pictures of those numbers [17].

A selection of such representation techniques or visual structures, related to the work of **e·nventory,** follows.

2.1. Maps

2.1.1. Geographic maps

Historically, over 5,000 years ago, the first visual structures were maps drawn on clay tablets [13]. However, it wasn't until the 17thcentury that cartographic and statistical skills came together to construct data maps [17]. The traditional geographic map is an excellent example of how large amounts of data can be efficiently displayed visually.

Nowadays, the most common types of two-dimensional data visualisations are geographic information systems (GIS) and web mapping. Large commercial GIS systems are being used for regional planning, transportation planning and management, weather forecasting, and mapping. The use of the web as a dissemination medium for maps can be regarded as a major advancement in cartography: simple GIS applications are becoming fairly common on the web in the form of customized maps that are displayed to show the location of an address in response to a query to a search engine.

Mapquest[4] was the first such popular online address matching and routing service with mapping output. Figure 4 shows a street map returned in response to a query that supplied a zip code and other relevant information. Since then, numerous online map services have been deployed: OpenStreetMap, Google Maps,

[4]http://www.mapquest.com

Map24, Bing Maps, ViaMichelin, Multimap.com, Ask.com Maps, MapQuest, WikiMapia, Ovi Maps, NearMap, Mappy, and many others with local coverage.

Figure 4: A snapshot from the original Mapquest

Among the online map services, Google Maps[5] launched in 2005 and its API allow developers to integrate Google Maps into external websites and overlay specific data on them. Although initially only a JavaScript API, the API has expanded to include an API for Adobe Flash applications, a service for retrieving static map images, and web services for performing geocoding, generating driving directions, and obtaining elevation profiles. Over 350,000 web sites use the Google Maps API, making it the most heavily used web application development API. As Google Maps is coded almost entirely in JavaScript and XML, some end users have reverse-engineered the tool and produced client-side scripts and server-side hooks which allowed a user or website to introduce expanded or customized features into the Google Maps interface. Using the core engine and the map/satellite images hosted by Google, such tools can introduce custom location icons, location coordinates and metadata, and even custom map image sources into the Google Maps interface.

Among all the visualisation techniques of geographic/cartographic maps the ones that are most commonly used in data observatories are geographic maps with colour coding/highlighting according to the value of an index (Figure 5) and simple geographical maps with an overlay of custom location icons placed on the map (Figure 6) or other information (e.g. ranking in Figure 7).

Figure 5: Geographical maps with a colour coding according to the value of an index[6]

[5] http://maps.google.com

[6] http://www.imd.org/research/publications/wcy/wcy_online.cfm

e·nventory has used YOOmaps[7] (a lightweight module for Joomla) to integrate Google Maps into the **e·nventory web platform** to offer the **European eInfrastructures Observatory Organisation Mapping Service**. The user can choose to display NRENs, NGIs, Supercomputer Centres and European eInfrastructures coordinating entities on a geographic map (Figure 6).

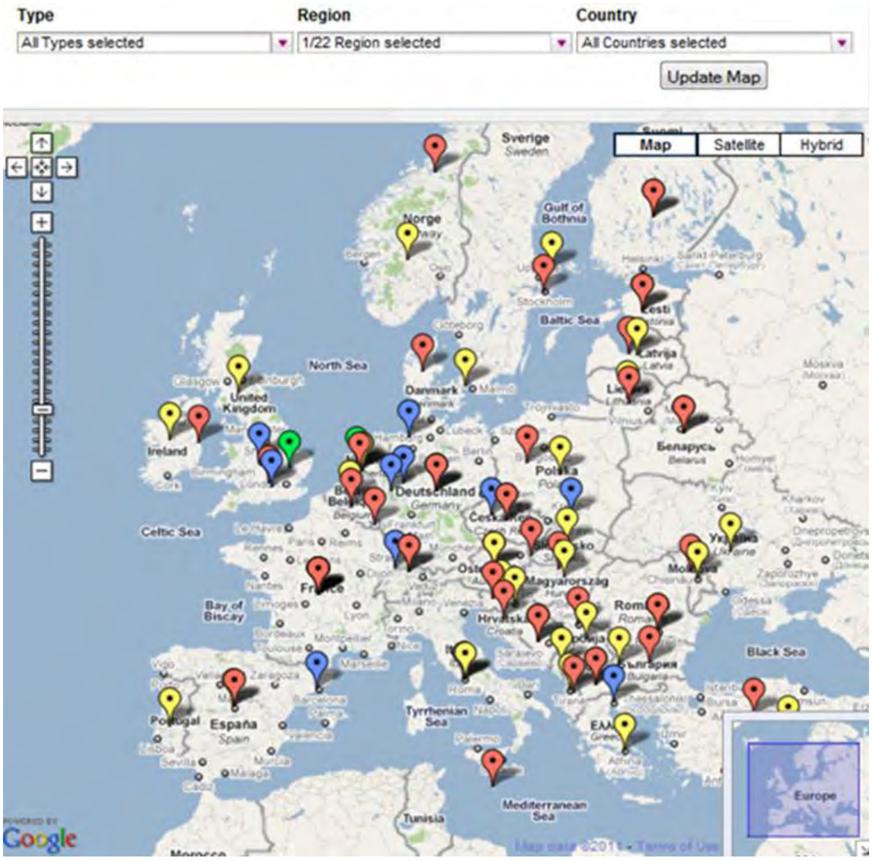

Figure 6: Geographical map with custom location icons - e·nventory's organisation mapping[8]

Figure 7: Geographical map with an overlay of location-related additional information – WEF[9]

[7]http://www.yootheme.com/tools/extensions/yoomaps

[8]http://www.enventory.eu

[9]http://www.networkedreadiness.com/gitr/main/analysis/mapview.cfm

2.1.2. *Cartograms*

Cartograms are a rather interesting visualisation technique that morphs geographical maps according to the value of an index. The geometry or space of the map is distorted in order to convey the information of this alternate variable. There are two main types of cartograms: area (value-by-area or isodemographic) and distance (central-point) cartograms.

This technique has been used in the FedEx Experience[10] project. On-line visitors explore information on a map of the world (Figure 8), on which the different countries change in size according to the value of the index that is represented at any given time. The indices that are visualised belong to numerous different categories (24 categories each comprised of 3 different indices, therefore 72 different maps). The category topics range from air travel (number of world airports, number of departures, etc.) and social media penetration (Facebook users, etc.) to the "coffee effect" (i.e. coffee imports, labour productivity, even insomnia cases). Another interesting feature is the ability to compare any one map to another.

This technique has been also utilized in the Online Atlas of the Millennium Development Goals (Figure 9). The user is able to explore maps of key indicators for each of the eight MDG goals, where countries resize to reflect comparative values.

Figure 8: Cartogram - FedEx Experience project

Figure 9: Map morphing according to the value of an index- World Bank online atlas

[10] http://www.experience.fedex.com

2.1.3. Treemaps

Treemapping is an effective non-geographical spatial map technique for displaying hierarchical (tree-structured) data as a set of nested rectangles. Each branch of the tree is given a rectangle, which is then tiled with smaller rectangles representing sub-branches. A leaf node's rectangle has an area proportional to a specified dimension on the data. Often the leaf nodes are coloured to show a separate dimension of the data. When the colour and size dimensions are correlated in some way with the tree structure, one can often easily see patterns that would be difficult to spot in other ways. A second advantage of treemaps is that, by construction, they make efficient use of space. As a result, they can legibly display thousands of items on the screen simultaneously.

The Map of the Market[11] (Figure 12) visualises more than 500 stocks at once, grouped by sector and industry, with data updated every 15 minutes. Each coloured rectangle in the map represents an individual company. The rectangle's size reflects the company's market cap and the colour shows price performance (green means the stock price is up; red means it's down, dark colours are neutral). By moving the mouse over a company rectangle, more information is displayed in a pop up "tooltip". The controls on the right side provide a number of choices: although the map starts off displaying price changes since the previous market close, one can change to view over three other time periods; to pick out the market movers, to highlight the top five gainers or losers; to find a particular stock; to spotlight a company, type its name or symbol; to change the colours to a blue/yellow colour scheme for a clearer view.

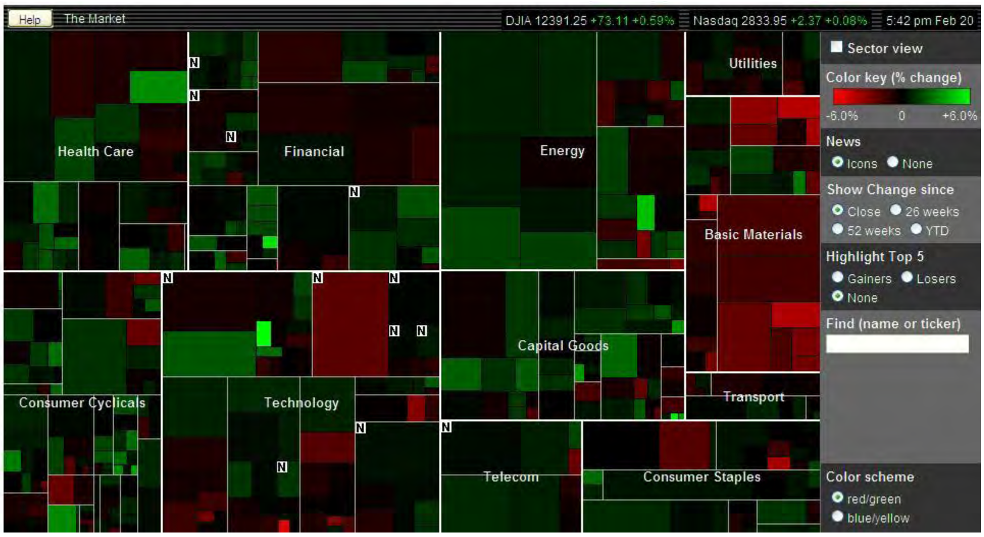

Figure 10: Treemap - the Map of the Market

2.1.4. Self-organizing feature maps

A self-organizing map or self-organizing feature map is a type of artificial neural network that is trained using unsupervised learning to produce a low-dimensional (typically two-dimensional), discretised representation of the input space of the training samples, called a map.

In Figure 11, the geographic map metaphor has been used in a more abstract way, such as in the example of the SiteMap, a method of visualising groups of documents in a collection, using size and location as the primary attributes. Documents in a collection are indexed and assigned to groups corresponding to the frequency of terms used within the documents and the collection. The relative importance of each group within the collection is represented in an abstract space (a large rectangle that serves as the map's outer boundary) by the size of the group's area within the bounding rectangle. In addition, the location of an area within the map indicates its relationship to other groups; neighbouring areas are more closely related than groups farther apart from each other. Although intuitively appealing, those graphical overviews of large document spaces have yet to be shown to be useful and understandable for users.

[11] http://www.smartmoney.com/map-of-the-market

Figure 11: Self-organizing feature map

2.1.5. Other mixed techniques

An example is the New York Talk Exchange (NYTE)[12], a visualisation of volumes of long distance telephone and IP (Internet Protocol) data flowing between New York and cities around the world. Specifically, the "World Within New York" visualisation (Figure 12) shows how different neighbourhoods reach out to the rest of the world via the AT&T telephone network. The city is divided into a grid of 2-kilometer square pixels where each pixel is coloured according to the regions of the world wherein the top connecting cities are located. The widths of the colour bars represent the proportion of world regions in contact with each neighbourhood. Encoded within each pixel is also a list of the world cities that account for 70% of the communications with that particular area of New York.

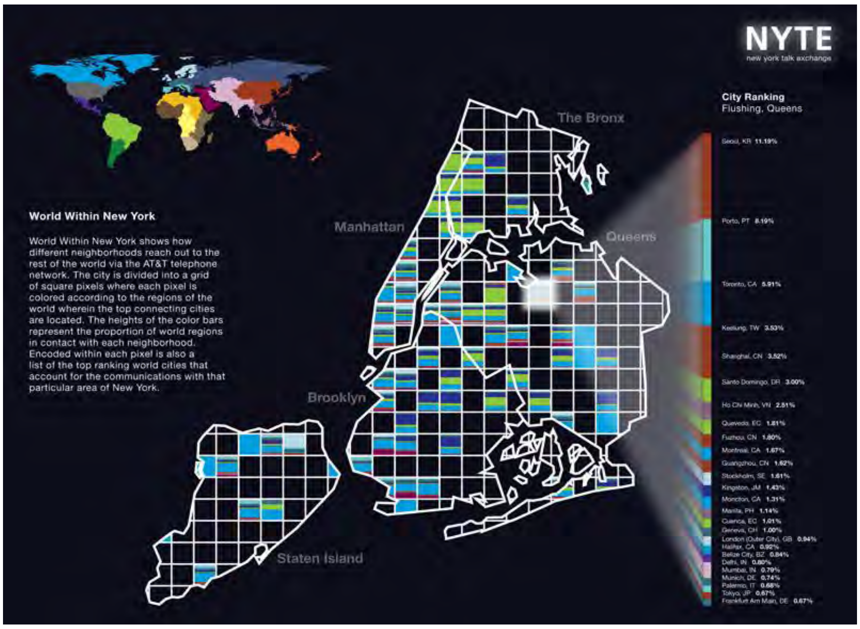

Figure 12: Mixed map techniques - New York Talk Exchange visualisation of volumes of traffic

[12] http://senseable.mit.edu/nyte/visuals.html

2.2. Charts

2.2.1. Line charts

A line chart (Figure 13) is a type of graph, which displays information as a series of data points connected by straight line segments. It is an extension of a scatter graph, and is created by connecting a series of points that represent individual measurements with line segments. A line chart is often used to visualise a trend in data over intervals of time – a time series – thus the line is often drawn chronologically.

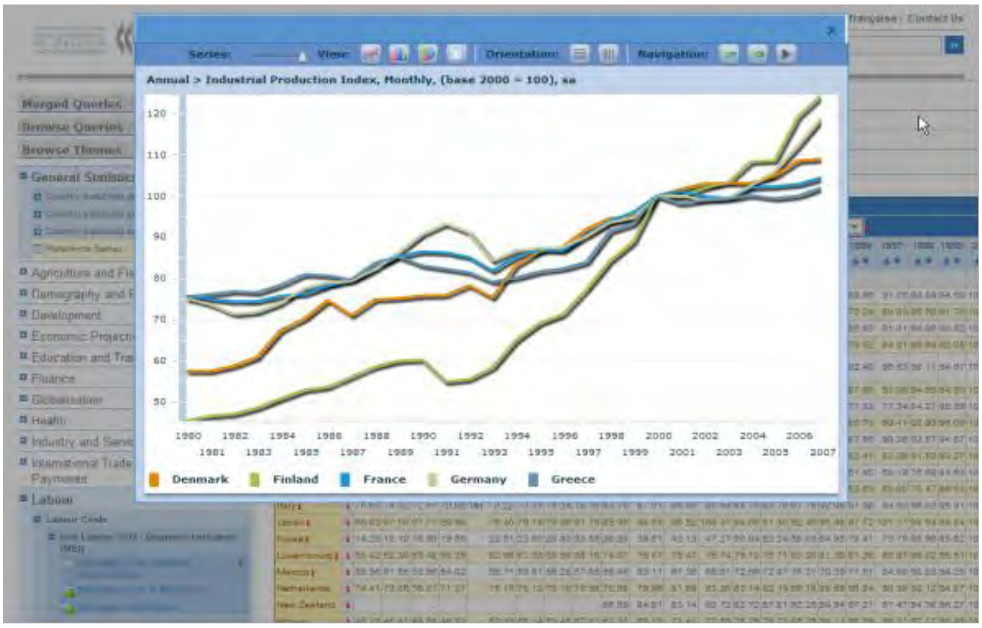

Figure 13: Line chart - OECD statistical information[13]

2.2.2. Bar charts

Bar charts (Figure 14) are charts with rectangular bars with lengths proportional to the values that they represent. They are useful for visualising summary data across various categories. Bar charts are used for plotting discrete (or 'discontinuous') data i.e. data which has discrete values and is not continuous. The bars can be plotted vertically or horizontally.

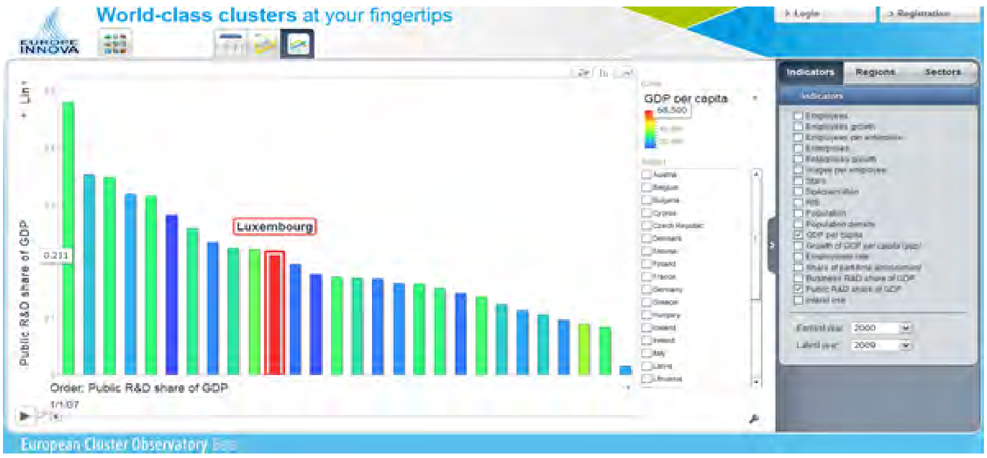

Figure 14: Bar chart – European Clusters Observatory[14]

[13] http://stats.oecd.org

[14] http://www.clusterrobservatory.eu

2.2.3. Pie charts

A pie chart is a circular chart divided into sectors, illustrating proportion. In a pie chart, the arc length of each sector (and consequently its central angle and area), is proportional to the quantity it represents. The pie chart is an age-old visualisation technique (Figure 15) which continues to be very popular. Nowadays, there are quite a few variants of pie charts, from exploded pie charts to 3D pie charts. A polar area diagram is an enhanced form of a pie chart.

The pie chart is perhaps the most ubiquitous statistical chart in the business world and the mass media. However, it has been criticized, and some recommend avoiding it, pointing out in particular that it is difficult to compare different sections of a given pie chart, or to compare data across different pie charts. This solution is also impractical for the colour blind [10]. Pie charts should be avoided and three-dimensional pie charts should never be used [44].

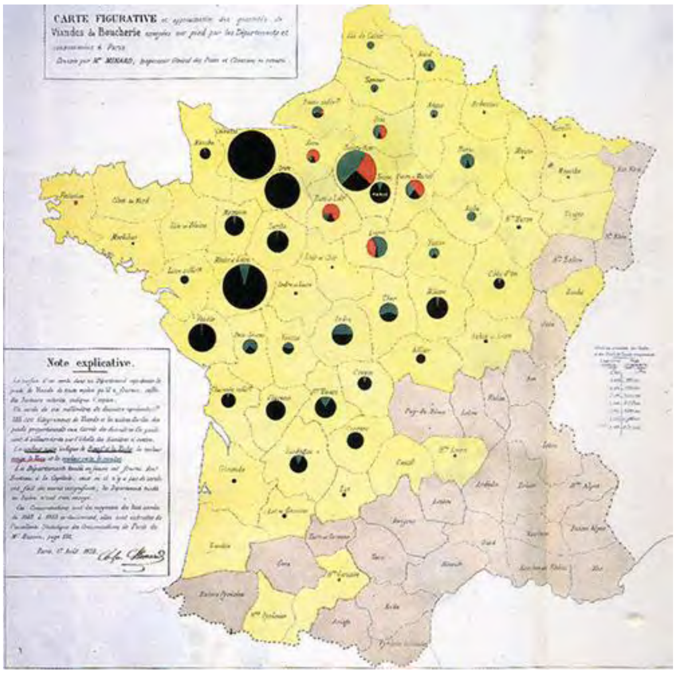

Figure 15: Minard's map using pie charts to represent the cattle sent from France's regions to Paris (1858)[15]

2.2.4. Bubble charts

A bubble chart is a type of chart where each plotted entity is defined in terms of three distinct numeric parameters. The entities displayed on a bubble chart can be compared in terms of their size as well as their relative positions with respect to each numeric axis. Since both X and Y axes of the bubble chart are numeric scales, the position of plot is an indicator of two distinct numeric values. The area of the plot depends on the magnitude of a third numeric characteristic. One concern when rendering data with a bubble chart is that the area of a circle is proportional to the square of the radius. So if one scale the radius with the third data point, will disproportionally emphasize the third factor. To get a properly weighted scale, one should take the square root of the magnitude of this third metric. Many bubble charts are rendered without this correction.

A bubble chart can be considered a variation of a scatter plot, in which the data points are replaced with bubbles. This type of chart can be used instead of a scatter chart if the data has three data series, each of which contains a set of values.

The Data Visualizer tool of the World Bank employs bubble charts to display data in four dimensions (Figure 16). In each chart, the size of the country circle represents a volume measure, such as population or GDP. The position of the bubbles is determined by the indicators selected for the horizontal and vertical axes. The time series used is a subset of 2010 World Development Indicators database. In some cases, bubble charts are used in combination with geographical structures (Figure 17).

[15]"Des chiffres et des cartes: la cartographie quantitative au XIXè siècle". Paris: Comité des travaux historiques et scientifiques, 1858

Figure 16: Bubble chart – World Bank's Data Visualizer[16]

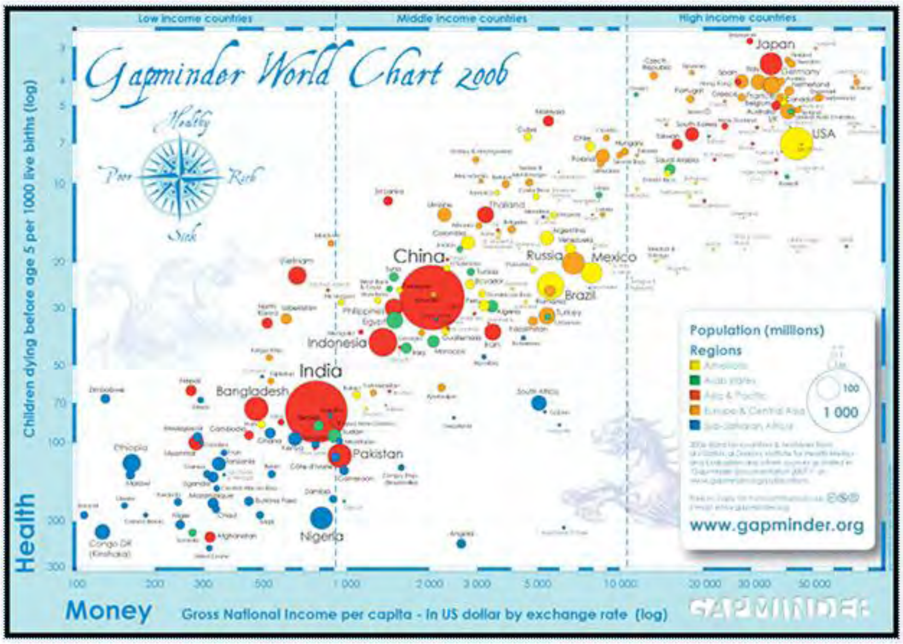

Figure 17: A combination of a bubble chart and a geographical visual structure[17]

2.2.5. Scatterplots

A conventional approach to the representation of values for two variables for a set of data (i.e. bivariate data) is a *scatterplot* or *scattergraph*. The data is displayed as a collection of points, each having the value of one variable determining the position on the horizontal axis and the value of the other variable determining the position on the vertical axis (Figure 18). A scatter plot can suggest various kinds of correlations between variables with a certain confidence interval. Correlations may be positive (rising), negative (falling), or null (uncorrelated). A line of best fit (alternatively called 'trend line') can be drawn in order to study the correlation between the variables.

An example of an enhanced scatterplot is the FilmFinder (Figure 19). The FilmFinder is an application created by the Human-Computer Interaction Laboratory of the University of Maryland for exploring a film database. It combines different kinds of encoding and interaction techniques to represent hypervariate (multi-dimensional) data in a single interface. The FilmFinder is designed to allow a user to find information about a

[16] http://devdata.worldbank.org/DataVisualizer

[17] http://www.gapminder.org/world

film. Films are represented on the main (scatterplot) display as coloured squares (each film's "FilmID" variable is mapped to a coloured square-point). Each colour encodes a film genre (drama, comedy, etc), shown also as buttons on the lower part of the screen. The scatterplot includes two orthogonal quantitative axes: the X-axis represents time (year of production) and the Y-axis a measure of popularity. The FilmFinder augments the scatterplot with a collection of user interface sliders and buttons, which can be used to specify other attributes such as Actor, Actress, and Director (using "alphasliders"), ratings (with checkboxes), film length (with a double box range selector), film genre (with large colour-coded buttons at the bottom), etc. These allow rapid query specification through direct manipulation, which is coupled with instantaneous feedback.

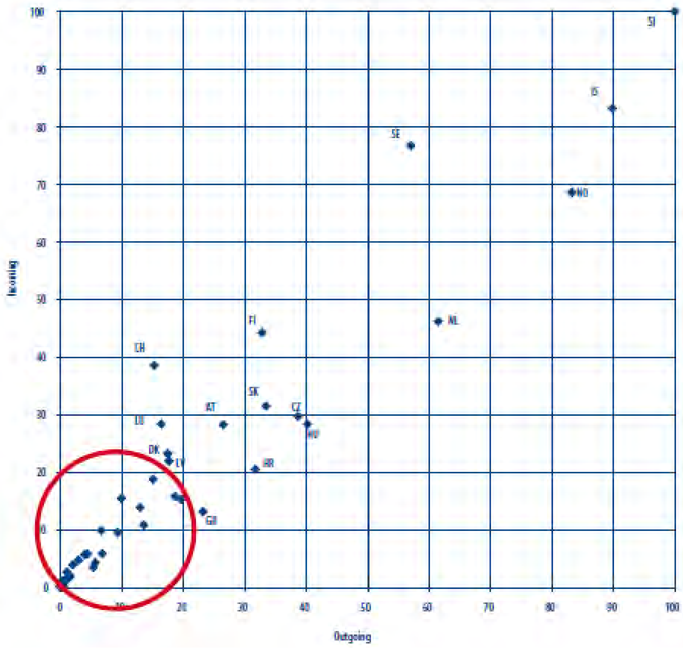

Figure 18: A scatterplot – EARNEST geographic study – outgoing vs. incoming NREN traffic per user[18]

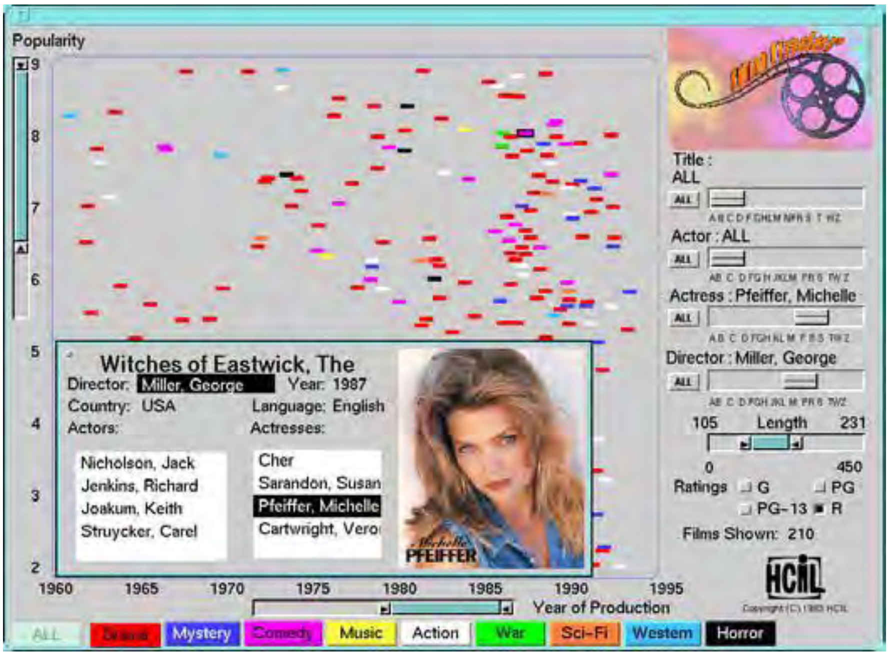

Figure 19: A scatterplot enhanced by additional and selective encoding - the FilmFinder

[18]http://www.terena.org/activities/earnest/geog.html

2.2.6. *Time-series plots*

A special case of the line chart, the bar chart or the scatterplot is a time series plot, in which one axis represents time (expressed as time of day, date, month, year, and so on) and the other some function of time (e.g. the daily closing value of the Dow Jones index). The time-series plot is the most frequently used form of graphic design. Most timeline charts use a 2D representation, time x {place or theme}. Some are more successful in integrating additional dimensions.

2.2.7. *Spider/Radar charts*

There are several other kinds of visual structures, which in some cases, can be considered as relatively recent visualisation constructs. An impressive such visual structure is the Spider or Radar chart. Spider charts are used to display multivariate data in the form of a 2D chart of three or more quantitative variables. These are represented on axes starting from the same point. Because shape can often effectively represent the combined attribute values of a single object, the points on each axis can be joined (Figure 20).

Spider charts are primarily suited for comparing specific objects, showing extreme observations and commonality, or when one chart is greater in every variable than another. They are mostly used for ordinal measurements, i.e. where each variable corresponds to "better" in some respect, and all variables on the same scale. E.g. the World Economic Forum (WEF) uses the spider chart to compare the performance of indices for a selection of countries[19].

Radar charts are visually striking, and can add interest to what would otherwise be a dry data presentation. Radar charts are primarily suited for strikingly showing outliers and commonality, or when one chart is greater in every variable than another, and primarily used for ordinal measurements – where each variable corresponds to "better" in some respect, and all variables on the same scale. Conversely, radar charts have been criticized as poorly suited for making trade-off decisions – when one chart is greater than another on some variables, but less on others. Further, it is hard to visually compare lengths of different spokes, because radial distances are hard to judge, though concentric circles help as grid lines. Instead, one may use a simple line graph, particularly for time series.

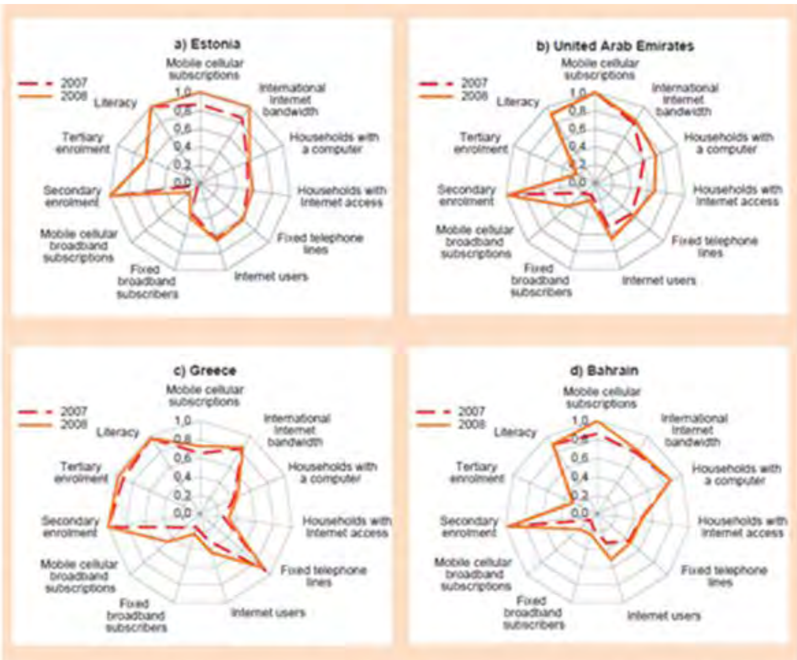

Figure 20: A spider or radar chart - ICT Development Index- ITU[20]

[19].World Economic Forum, last retrieved at http://gcr.weforum.org/gcr2010/

[20]http://www.itu.int/ITU-D/ict/publications/idi/2009

2.2.8. Box and whisker diagrams/plots

A box and whisker diagram (Figure 21) is another convenient way of graphically depicting groups of numerical data through their five-number summaries: the smallest observation (sample minimum), lower quartile (Q1), median (Q2), upper quartile (Q3), and largest observation (sample maximum). It is used to display a set of data so that one can easily see where most of the numbers are.

Box and whisker plots are uniform in their use of the box: the bottom and top of the box are always the 25th and 75th percentile (the lower and upper quartiles, respectively), and the band near the middle of the box is always the 50th percentile (the median). The ends of the whiskers can represent several possible alternative values.

The boxplot is a quick way of examining one or more sets of data graphically. Boxplots may seem more primitive than a histogram but they do have some advantages. They take up less space and are therefore particularly useful for comparing distributions between several groups or sets of data. Choice of number and width of bins techniques can heavily influence the appearance of a histogram.

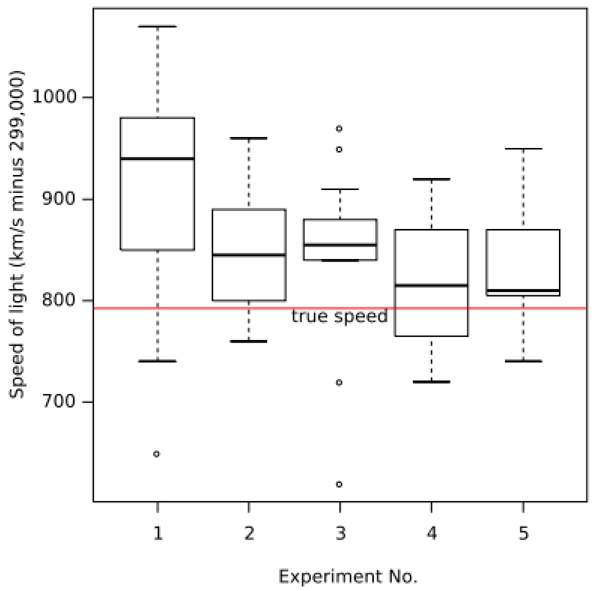

Figure 21: A box and whisker plot representing Michelson and Morley's data on the speed of light

Other charts, such as 3D plots (Figure 22), Polar area diagrams, Waterfall charts, Biplots are less relevant to the types of data dealt with in the **e·nventory** project.

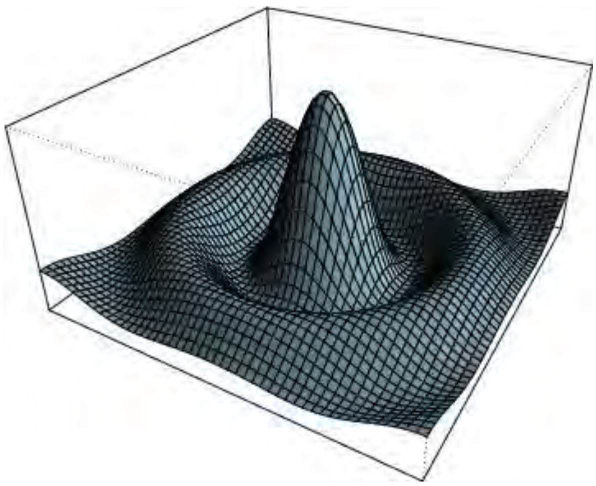

Figure 22: 3D plot of a surface[21]

[21] http://www.r-project.org

2.3. Timelines

A timeline is a graphic design showing a long bar labelled with dates alongside itself and (usually) events labelled on points where they would have happened. It is used to show events along a period of time.

Timelines are particularly useful for studying science, as they convey a sense of change over time. Wars and social movements are often shown as timelines. Timelines are also useful for biographies and project management. Timelines can make use of any time scale, depending on the subject and data. Most timelines use a linear scale, where a unit of distance is equal to a set amount of time.

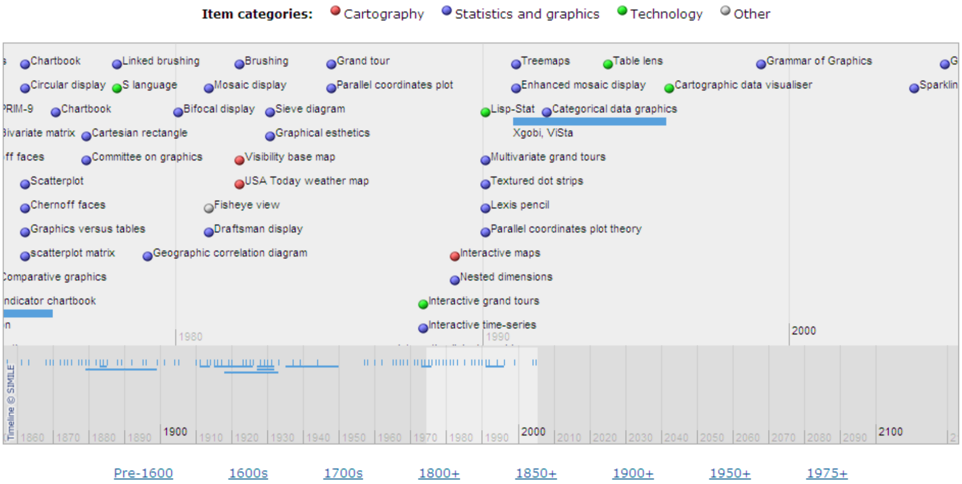

Figure 23: A set of observations taken at different points in time and charted in a time series[37]

An older example, the Digital History web project[31], designed and developed to support the teaching of American History in K-12 schools and colleges, deploys an interactive timeline which allows users to display political/diplomatic, social/economic, and cultural events on a map of the United States from 1590 to the present year (Figure 24). There is a gold-coloured "Life Span Bar" at the bottom of the screen to navigate through the years and events appear or disappear, represented by a blue square for political/diplomatic events, a green triangle for social/economic and an orange circle for cultural events. With mouse rollover, the circles, triangles or squares pop up a tooltip with event details. By clicking on each shape a new window opens with more information.

Figure 24:Digital History Timeline[31]

3. Interaction techniques

The principle feature which separates modern information visualisation from the graphical applications created on paper is interaction: the facility to change our view of the available data. According to the information visualisation reference model presented in Figure 1, interaction involves the transformations that provide the user with the ability to view different aspects of the visually encoded data; the user manipulates controls to change parameters in the chain of these transformations. These controls can be separate from the visualisation or integrated into it.

The ability to change one's view of a corpus of data is of enormous benefit since information spaces are so large that no single all-inclusive view is likely to lead to insight. The effect of a good visual representation is amplified by making it interactive.

Various applications exemplify the addition of interaction in information visualisations, including the FilmFinder [23], the Attribute Explorer, a graphical interactive tool for visualising the relationships within multi-attribute data sets [24], the Magic Lens and See Through Tools [25], Dust & Magnet [26], FacetMap [27], Spotfire [28], DataMaps[22], and others. Additionally, there are plenty of web resources where interactive visualisation is used [38], [39], [40], [41], [42], [43].

Interaction techniques for information visualisation are essentially a form of selection, selecting a subset of objects in a data table [8]. This allows them to be used to locate data, to reveal patterns in data, or to select the arguments of other transformations.

According to Card et al. in [8] there are three types of interaction (Table 1):

- Interaction with data transformations
  - Querying the database. Depending on the subset we want to show, different queries can extract different subsets of data.
  - Statistical transformations of and among the stored data. We could wish to visualize derived magnitudes, like differences, averages or other statistical magnitudes.
  - Structural changes in the database. In the same way we derive new magnitudes from existing data we can change or combine the structure of the database in order to find new patterns.
- Interaction with the type of graphic representation, the correspondence between data and graphic structures, i.e. the visual mapping and/or the type of correspondence between what is shown and what is meant.
- Interaction with the geometric transformations that affect the rendering of the representation, like rotations, translations and projections. One can change the perspective, the angle of rotation, the pan or zoom.

Table 1: Interaction techniques (Card et al., 1999).

Modifies Data Transformation	Modifies Statistical Transformations	Modifies View Transformation
Dynamic queries	Dataflow	Direct selection
Direct walk	Pivot tables	Camera movement
Details-on-demand		Magic lens
Attribute walk		Overview + detail
Brushing		Zoom & pan
Direct manipulation		

Dix and Ellis in [10] listed also: highlighting and focus, accessing extra info – drill down and hyperlinks, overview and context – zooming and fisheyes, same representation - changing parameters, linking representations – temporal fusion. Keim's taxonomy in [11] includes: projection, filtering, zooming, distortion and linking and brushing.

[22] http://www.datamaps.com

3.1. Interaction with data transformations

3.1.1. Dynamic queries

In this technique sliders and buttons in their diverse forms are used to select and set boundaries to the data to be visualised. The rest of the data remains hidden or is partially visualised (for example using transparency). Dynamic querying is probably one of the best-known and most useful information visualisation techniques. Dynamic queries are essentially database queries and, at the same time, visual means of specifying conjunctions.

Queries are made using widgets, such as buttons, range sliders (for continuous data attributes), alphanumeric sliders (for textual attributes), toggles (for binary attributes), and check boxes (for discrete multi-valued attributes), to specify each attribute (dimension) of the data. The choice of query controls depends on the variable type. In general, buttons/radio buttons are usually used for binary (nominal) variables and sliders for ordinal or quantitative variables, but also for nominal variables.

In the example of applications such as the FilmFinder or the HomeFinder, sliders (for nominal to ordinal values) or radio buttons are used to select value ranges for variables in the data table. The cases for which all the variables fall into the specified ranges are displayed, while the other cases are hidden.

Some of the positive aspects of dynamic querying include that it is a visual representation of the world of action including both the objects and actions (very natural interaction); allows for rapid, incremental and reversible actions; allows selection by pointing (not typing); and promotes reversing, undo, exploration providing immediate and continuous display of results.

Some of the problems with dynamic querying include that controls are global in scope; controls take up space and they affect everything; controls must be fixed in advance; as the data set gets larger, real-time interaction becomes increasingly difficult that one must learn the language; only exact matches are shown; the magnitude of the results is not visible in advance, if at all; no helpful context is shown; and reformulating to a new query can be slow.

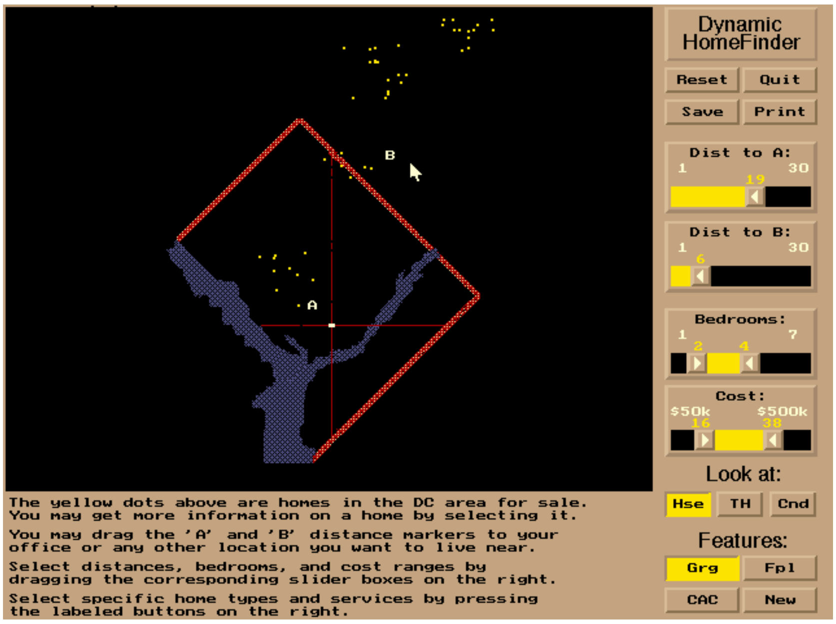

Figure 25: Dynamic queries - The Dynamic HomeFinder

3.1.2. Direct walk

Direct walk concerns a set of linkages between cases, where exploring one may lead to another. An example is following hyperlinks on web pages. Through a series of clicks on visualisations, the user can search for information or modify it.

3.1.3. Details-on-demand

This technique allows the user to expand the details of a data case or cases, i.e. expanding a small set of objects to reveal more of their variables. It allows more of the variables of the case to be mapped to the visualisation because there is more space to show them.

3.1.4. Attribute walk

In an attribute walk, the user selects some case and then searches for other cases with similar attributes. In the FilmFinder example, the details-on-demand window can be used to select cases with similar attributes.

3.1.5. Brushing

Brushing is a very common technique in Information Visualisation that applies when there are multiple views of the same data; hence, selecting or highlighting a case in one view generates highlighting the case in the other views (Figure 26). For example, highlighting a case from a data table in one view selects the case in the other views.

Range sliders can be seen as an elementary form of brushing. In the FilmFinder example, for instance, the slider represents all the objects ordered by some variable (e.g. movie length); selecting some range of these movies on the slider selects those same movies on the scattergram.

3.1.6. Direct manipulation

In this technique dialogue elements embedded in visualisations allow the user to directly manipulate and modify transformations.

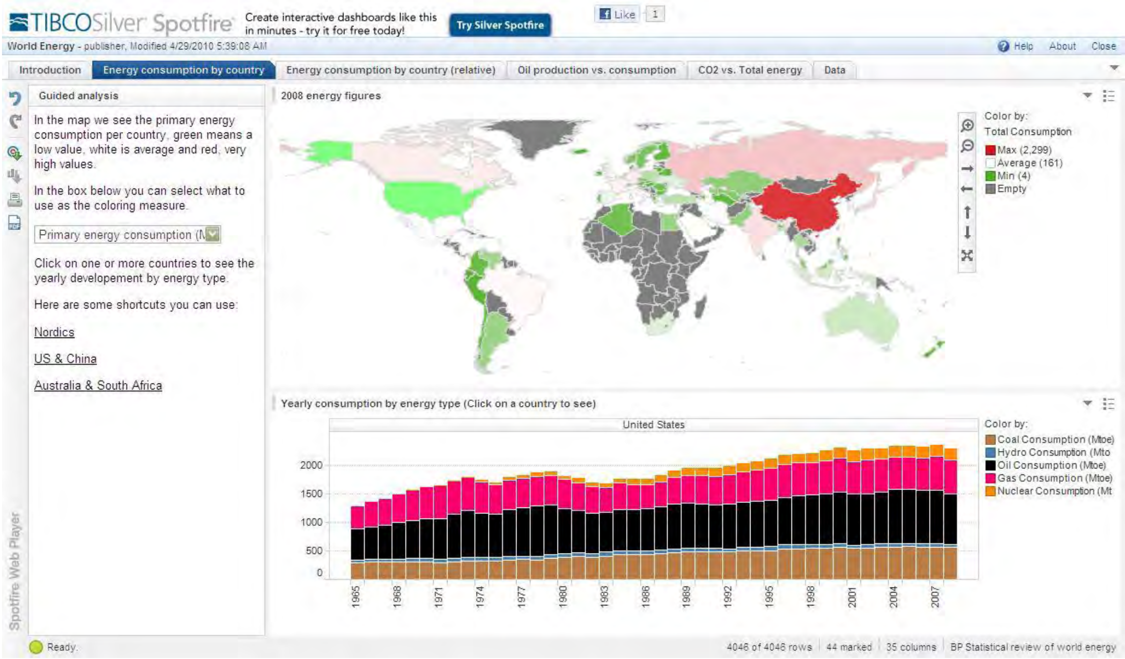

Figure 26: An interactive visualisation application where two different visual structures are linked together

3.2. Interaction with the correspondence between data and visual form

3.2.1. Dataflow

This type of interaction has been used in visual programming. It uses an explicit representation such as node-link diagrams to show the flow of data instead of showing the flow of logic (the sequence of programming sentences) so that the user can change it by manipulating the visual mapping.

3.2.2. Pivot Tables

This is a technique found in spreadsheets that lets the user easily manipulate the mapping of data to rows and columns.

3.3. Interaction with geometric transformations

3.3.1. Direct selection

Direct selection refers to a set of techniques for selecting and highlighting objects or groups of objects in a visualisation. Typically a pointer, a cursor, or some other visual element is used to select or to identify an item. Clicking on or moving the pointer over an item selects it and attributes of the data point are shown (Figure 27). Often a selection leads to a form of detailed information on that selection, which could be displayed as a "tooltip" or in a pop-up window. In the Table Lens example (Figure 28), the textual information contained in database tables are combined with graphical representations of the same data. This allows the user to focus on specific information in a specific part of a table while still getting a sense of how that information relates to the larger context of the entire table. This scheme gives the user a way of looking at and navigating through the information that makes understanding the information easier [13], allowing the selection of a film on the basis of type, duration, year of production and other attributes. The user's selected item is highlighted with a black circle – outline.

3.3.2. Camera movement

Accounts for the change in position of the centre of projection. This changes the perspective, the point of view of the observer avoiding in certain applications the occlusion that occurs between objects, especially in 3D images. This technique also allows seeing parts of the visualisation that can't be seen from another given perspective.

3.3.3. Magic lens

This technique selects objects according to the X, Y of their marks and then combines with further selection techniques like dynamic queries. They can combine also with other techniques transforming the view or the data.

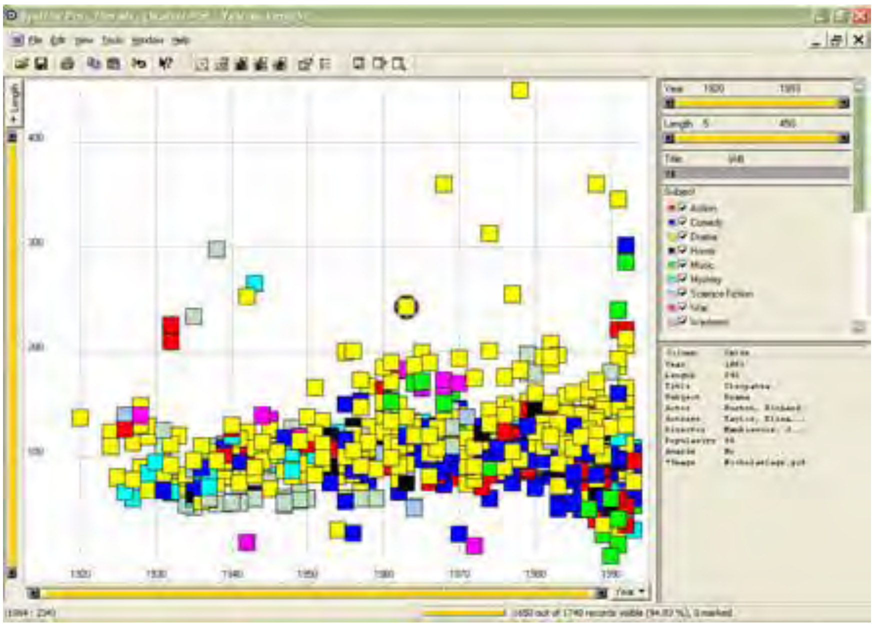

Figure 27: A scatterplot enhanced by additional selective encoding

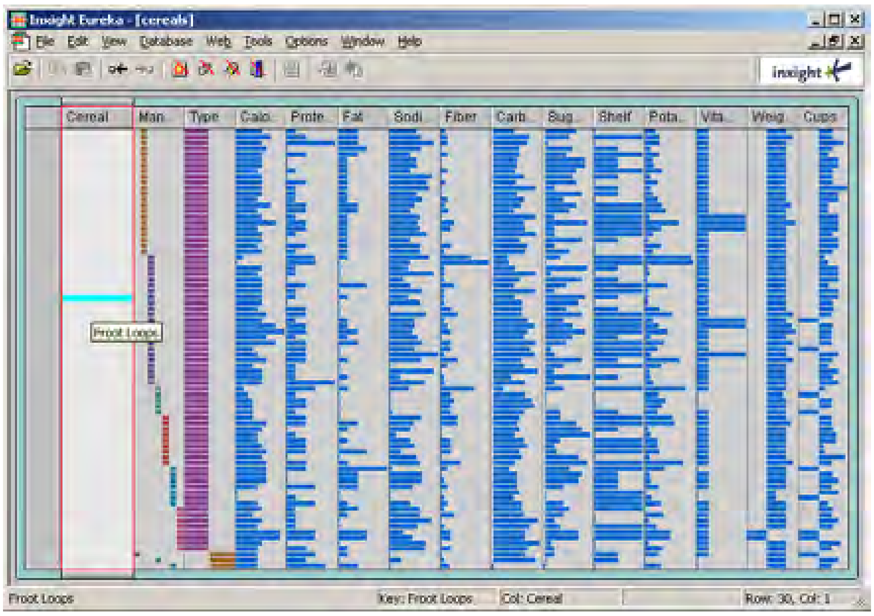

Figure 28: In the TableLens columns can be moved while a hovering mouse cursor brings up details of an item

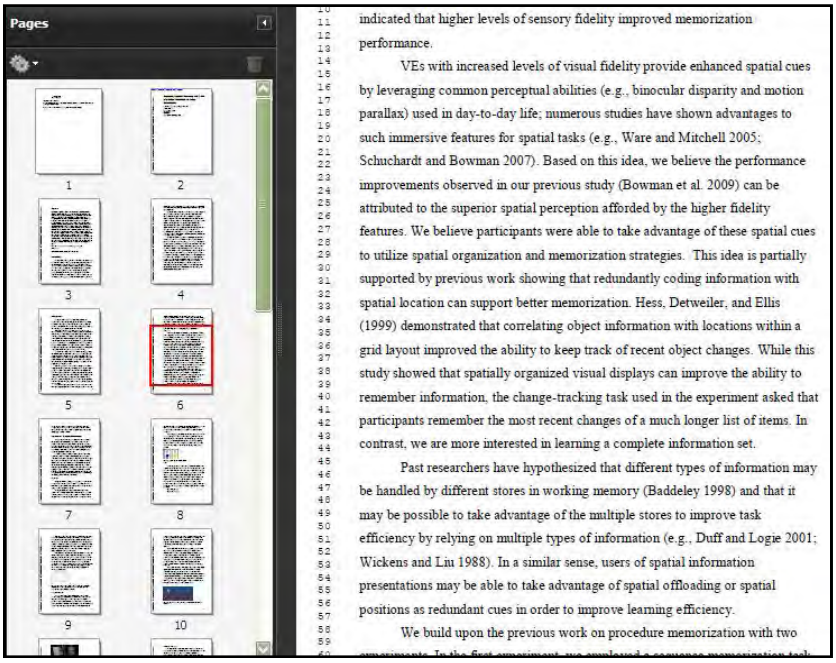

Figure 29: Overview + detail when viewing a pdf document: two views, one of context and the detail

3.3.4. Overview+detail

Overview+detail is a viewpoint control technique where two separate views are used side-by-side: an overview of the Visual Structure and a detail window that provides a magnified focus for one area. The overview provides the context for the detail view and acts as a control widget to change the detail view. A familiar example is the facility to see and overview comprising miniature pages of a document ranged vertically alongside one particular page from which detail is discernable (Figure 29). The user is afforded the opportunity to go into detail but to also retain an awareness of context, iterating their attention between the two views. The highlighted – red - frame around the miniature page provides a 'you are here' sign.

3.3.5. Zooming and panning

Many information visualisation systems provide zooming and panning capabilities on display. Other techniques, such as the overview+detail technique described above, are concerned with the most effective use of the available display area; zooming and panning are valuable processes to add to these. Zooming involves reducing the number of objects that are visible by possibly increasing the number of variables that are shown. In other words, it is the increasing magnification of a fraction of the viewing area, without changing the size of the viewing frame. Panning is the smooth, continuous movement of a viewing frame over a two-dimensional image of greater size [16]. The Google Maps application provides a familiar example of the application of both zooming and panning through embedded tools.

3.4. Other interaction techniques

3.4.1. Excentric technique

The widespread use of information visualisation is hampered by the lack of effective labelling techniques. The excentric technique is a dynamic technique to label a neighbourhood of objects located around the cursor without intruding into the existing interaction. It is appropriate for dealing with information that has too many labels and/or long labels, therefore facilitating the selection of objects.

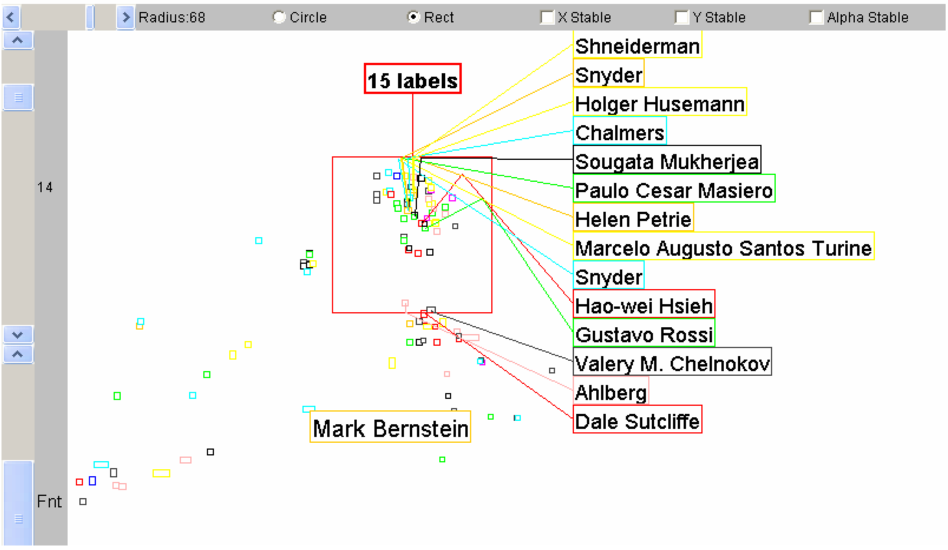

Figure 30: Example of the excentric technique to display many and long labels

3.4.2. Rearrange view

Rearranging views relates to keeping the same fundamental representation and what data is being shown, but rearranging the elements in order to carry out tasks such as to alter positioning or to sort the data.

3.4.3. Changing representation

There are visualisations where a change in the representation is desirable, e.g. when one is looking for a new perspective within the data or when there is limited available space (or “real estate”), which forces change. In these cases, the entire data representation may interactively change (Figure 31).

3.4.4. Highlighting connections

When the user of a visualisation may wish to examine different attributes of a data case simultaneously, then the technique of highlighting connections can be used. Alternatively, the same technique can be used when a user wishes to view the data case under different perspectives or representations but still requires to keep straight where the data case is (Figure 32).

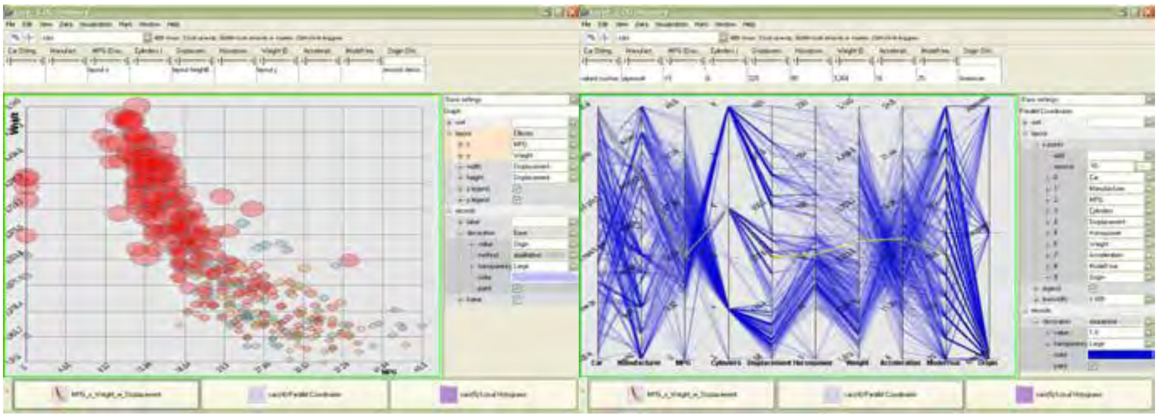

Figure 31: Selecting different representation from options at bottom

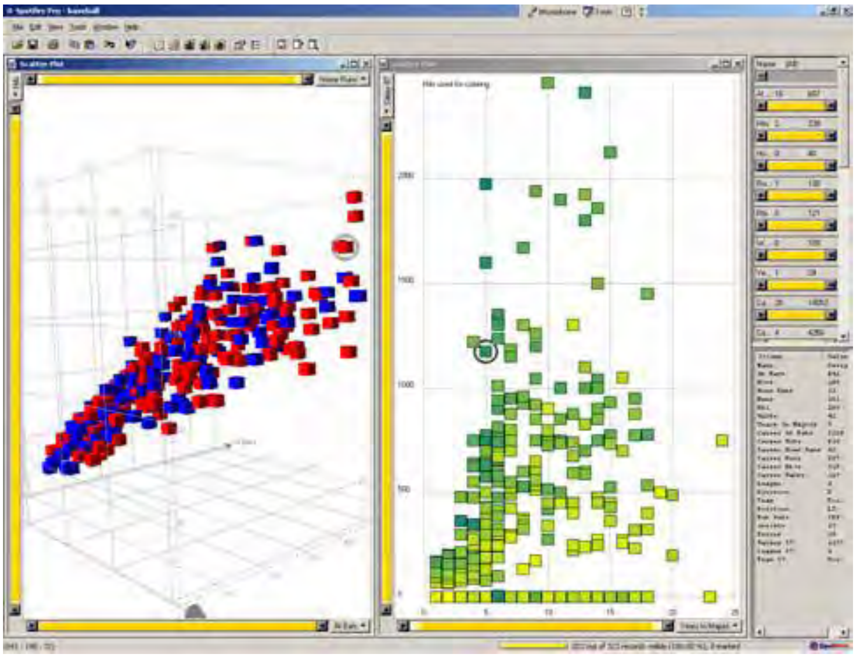

Figure 32: Two scatterplots depicting the same item in two different visual structures, side-by-side

3.4.5. Filtering/limiting

A fundamental interactive operation in information visualisation is changing the set of data cases being presented, either by displaying a larger set or reducing to a smaller subset. In other words, focusing or narrowing/widening. In addition to a new layout, it adds some interactivity, like a bounding box search interface and multiple views.

4. Reference case-studies on information visualisation

All of the aforementioned techniques for representing data in time and space, and interacting with the visualisations are implemented in web realisations. Out of the numerous examples presented in D2.1, the following reference cases have been selected in order to showcase the possibilities of web interactivity:

- The Human Development Report, United Nations Development Programme [32].
- The European Cluster Observatory [33].
- The Data Visualizer and the World at a Glance, World Bank [34],[35].

The above reference cases were selected either because they deploy many of the representation and/or interaction techniques presented in the previous chapters or because of their unique use of visual structures. Moreover, they also exhibit substantial relevance to the **e·nventory** concept, thus they are reviewed in sufficient detail in order to also provide input to future project activities.

The cases are judged in terms of the information visualisation approaches and tools on three axes:

- Representation (choice of visual structures)
- Interaction (choice of views)
- Usability (efficiency-speed, effectiveness and satisfaction)

Table 2 and Table 3 summarise the representation methods and interaction techniques presented in the previous sections that are used by these examples.

Table 2: Visual representation methods employed by the examined systems

	HDR/UNDP	ECO	Data Visualizer – WB	The World at a Glance - WB
Geographical maps	√	√		√
Bar charts	√	√		
Pie chart		√		
Line Chart	√	√		√
Scatterplot	√	√		
Time-series plots				
Bubble chart	√	√	√	√

Table 3: Interaction methods employed by the examined systems

	HDR/UNDP	ECO	Data Visualizer – WB	The World at a Glance – WB
Dynamic Queries	√	√	√	
Details on demand	√	√	√	√
Brushing	√			
Direct manipulation	√	√	√	
Direct selection	√	√	√	√
Overview & detail	√	√	√	√
Zooming & Panning	√	√	√ (only zoom)	√
Rearrange View	√	√		
Changing Representation	√	√		
Filtering / Limiting	√		√	

4.1. Human Development Report - UNDP

The Human Development Report (HDR) of the United Nations Development Programme [32] has been created with the aim to stimulate global, regional and national policy discussions on issues that are relevant to human development. To be of relevance, the data in the report requires the highest standards of data quality, consistency, transparency and accountability.

The HDR offers the possibility to browse a very rich data set and explore it through several views. Data is calculated mainly for 182 economies, and where possible the tables include data for the 192 members. All countries or areas are classified into one of four categories of achievement in human development (very high, high, medium and low). The country groupings also include the world and at least one geographic grouping -either the continents or UNDP Regional Bureaux groups. The HDR includes four composite indices: the Human Development Index (HDI), the Gender-related Development Index (GDI), the Gender Empowerment Measure (GEM) and the Human Poverty Index (HPI).

The UNDP has produced an interactive world map[23] to visualise the international HDR indicators and indices (Figure 33). It is based on the *StatPlanet World Map* software, a web-based interactive data visualisation and mapping application with an improved and easier map legend customization. StatPlanet has a wide range of features to cater to the requirements of different target audiences.

The visual exploration tool offered by UNDP uses multi-windowed views of different attributes of the data, some aggregated, some hierarchically structured, and it utilises progressive refinement based on direct manipulation of the visualised query results. The user has several options available as to viewing the data available which include “data by country” and “data by indicator”.

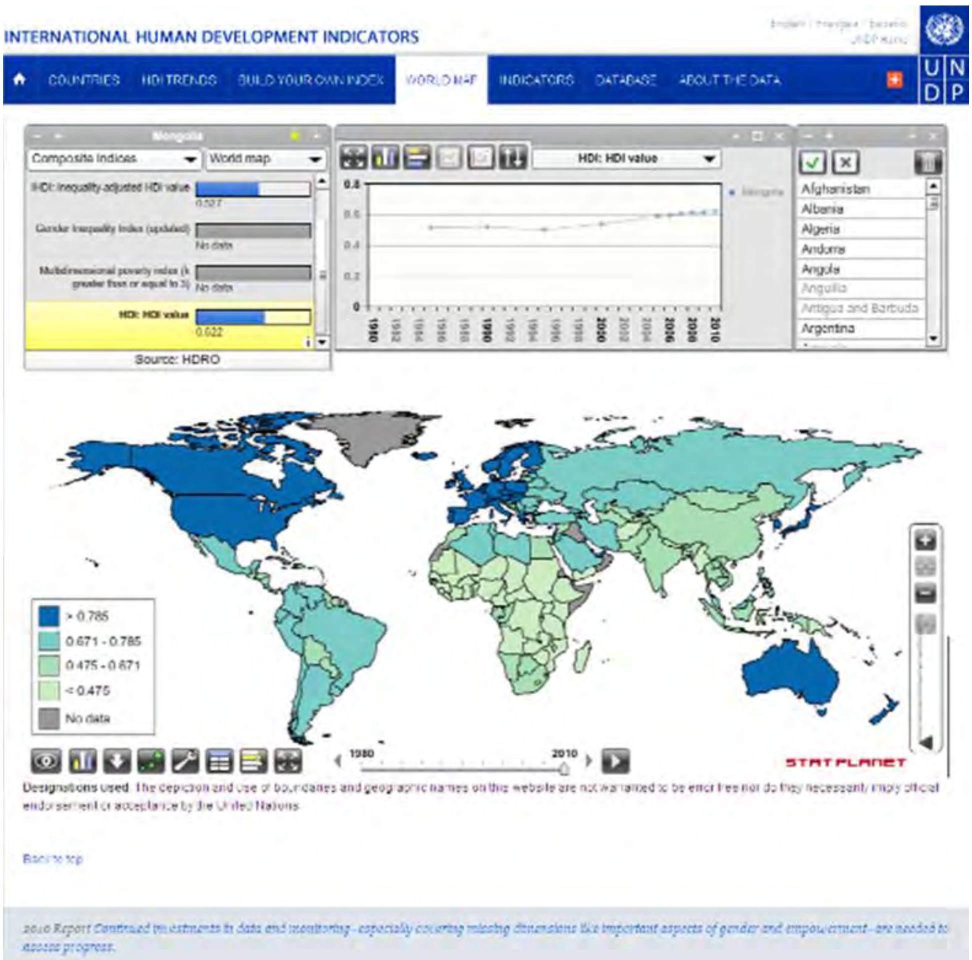

Figure 33: Human Development Report Observatory

[23] http://hdr.undp.org/en/data/map

Overall, the HDR/UNDP is an impressive, highly interactive, and visually appealing tool. It combines successfully a map visualisation with charts in providing overview and details on demand using direct selection and manipulation.

However, there is a lot of information into a single environment thus compromising the usability and understandability of the tool and there are also some minor usability issues in the functionality and arrangement of the World Map elements. Such shortcomings by no means lessen the added-value, and the aforementioned observations are only meant to identify possible areas for improvement to be taken into account by the **e·nventory** project.

The following sections focus on the information visualisation aspects of the HDR/UNDP interactive World Map tool.

4.1.1. Representation

The HDR/UNDP World Map tool is based on map visualisation. The lower part of the screen is a representation of the world map. The upper part is occupied by tables used to refine the visualisation and provide detail graphs as well. At the lower left part of the screen there is a toolbar.

The tool successfully manages to represent a wealth of data on one screen in a visually attractive and, at the same time, clear and concise manner. It is an appealing visualisation which invites the user to explore the data, even though the user may not have an immediate interest or knowledge of their domain.

The top left box allows the user to select an index (health, education, etc.) and a region (the whole map, Europe and Central Asia, etc.) of interest (Figure 34). Selection of region activates on the map only the countries that belong to the region. The user may further refine the selected regions by using the box on the right to select particular countries of interest.

The selected indices appear on the box and, while the user explores the map (using on mouse over to select a country), the indices value in the box is updated to reflect the selected country. In the middle box, a graph representation is available for a visual representation of the selected data (Figure 35).

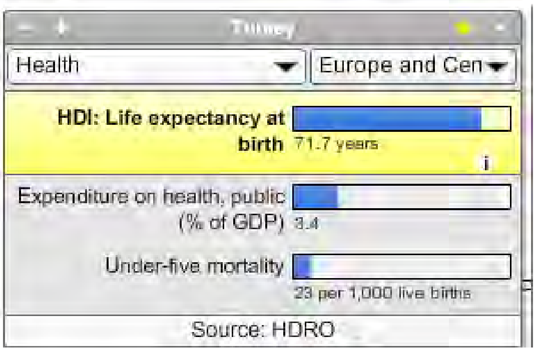

Figure 34: The HDR/UNDP tool's top left box with the selected indices and values and the region selection

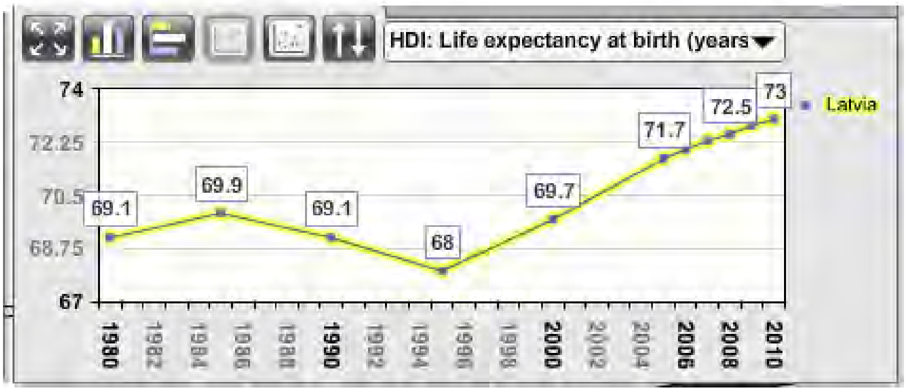

Figure 35: The HDR/UNDP tool's middle box with the visualisation of a specific index for a specific country

While moving over the map, pop-ups appear with the index values for the specific countries. Through the options bar on lower part of the screen, the user may select to represent the index value for the pop-up with a bar chart (Figure 36).

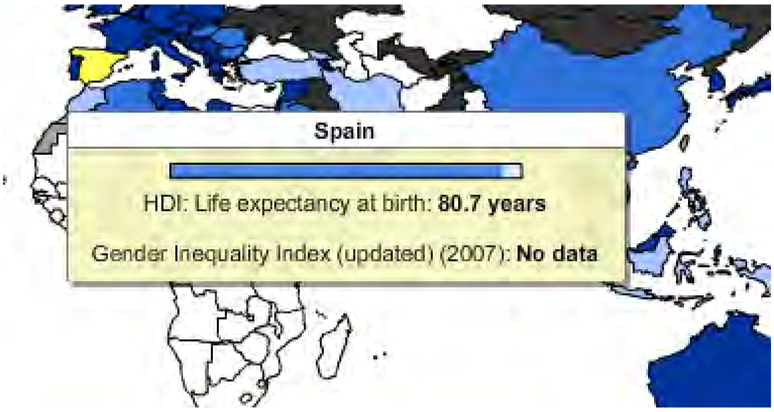

Figure 36: The HDR/UNDP tool's pop-up window on mouse over a specific country of the selected region

Colour coding on the map is used to provide an overview of the index's value for each country (Figure 37). A legend for the colour coding is provided on the lower left side. Time selection is accomplished through a year selector at the lower part of the map. A "play" option by the time selector allows an automatic successive change of the selected year in order to provide a better sense of the change of the index through time.

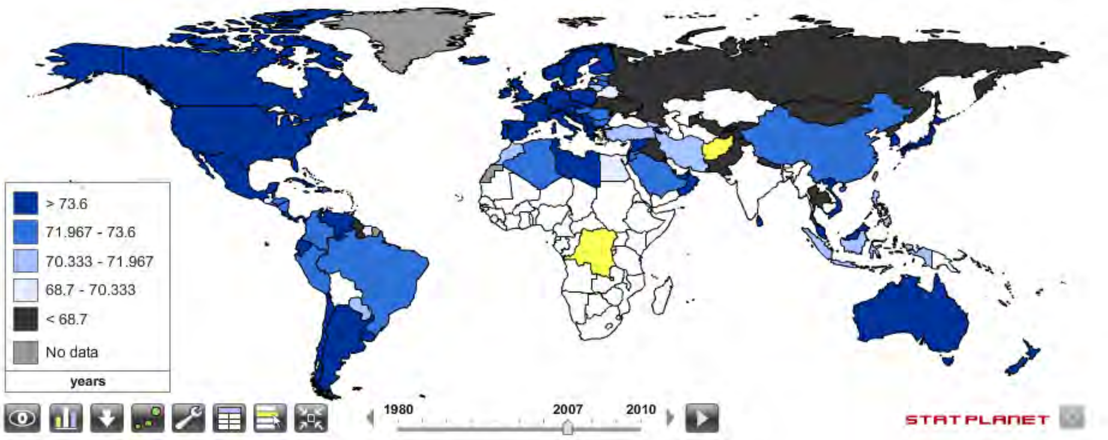

Figure 37: The HDR/UNDP tool's world map with colour coding according to the index values

The toolbar at the bottom offers also a bubble chart option, however the legend is not very clear.

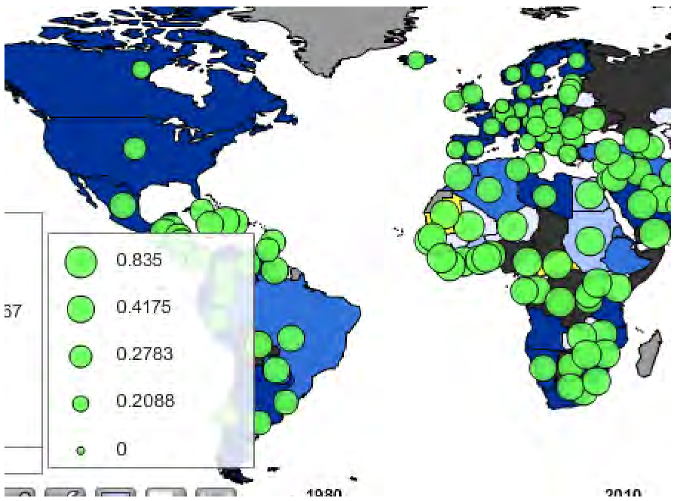

Figure 38: Bubble chart option

A useful option is to export and save the map and the graphs as images, even though the information given on the exported images is not sufficient, especially the graph.

The visualisation options and features are intuitive; however the user needs to explore the screen for a while before realizing its full potential. This "overloaded" screen may be considered as a weakness by specific users, especially those with a tendency to get easily distracted by visual features.

On the whole, the World map successfully conveys visually the information that the user is exploring. It has many options available for customization of the screen and seamlessly combines the visual information with numeric data keeping always the visualisation clear and "readable".

4.1.2. Interactivity

The HDR/UNDP World Map comes with a rich set of highly interactive and customizable tools which enable developers to;

- create interactive maps very efficiently, and
- customize them extensively.

As a matter of fact, the user can:

- *choose the visualisation* between different types of visualisations (data can be viewed as maps, bar charts, line graphs and scatter plots),
- *select countries* by clicking on them in the world map or country selection window. Statistics and details for the country selected are displayed inside pop-ups and visualisation components. Moving the mouse over a country on the world map brings up a pop-up containing information about that particular country for the selected indicator (as well as the bookmarked indicator, if there is one). In addition, the country data for all indicators in the current category is shown in the indicators panel,
- *animate changes over time* (maps, bar charts and scatter plots can all be animated over time),
- *zoom into map regions* such as Europe or Africa, or define a custom region (maps and visualisations will narrow down to the selected region and provide additional details),
- *hide or reveal details* by hiding or revealing interface components, country names or statistics,
- *customize* the colours and value ranges of maps and graphs,
- *export* maps for use in other applications or publications.

Mouse-over and mouse-click may be used to interact with the map, the graphs and the selection boxes. Pop-ups are used to show data values and definitions (in the case of indices).

The tool allows for effortless direct selection and manipulation of the screen features. The boxes are movable with drag and drop, countries may be selected with a click and on mouse over their information is made visible.

The top left box along with the top right country selection tool successfully implements a dynamic querying mechanism where the user can customize the visualised information and directly see the outcomes on screen. Brushing is also applied as a change in the selected parameters is reflected both on the map and on the top middle graph.

The world map plays the role of the overview visualisation with details on demand available through the top centre graph, the data table option, the pop-ups, etc.

Zooming and panning is also available as an option. For the map, there is a zooming tool on the left whereas the map may be panned with drag and drop. There is a very useful zooming feature for the top boxes as well, which changes the size of the boxes.

Filtering is also possible for countries especially, through the country selection tool.

Rearranging and changing the representation is also an interaction technique used in World Map. Although there is not a big set of representations available, they are suitable for the specific application and highly customizable.

The bookmark indicator is a nice extra feature.

On the whole the interactivity of the World Map tool is one of its strongest features, especially as it comes without delay and effortlessly on the part of the user.

4.1.3. Usability

As already mentioned, the World Map is a highly interactive and visually rich tool which, at the same time, may be used without particular effort or training.

However, the environment seems and is quite complex and, as is the case, there are some usability issues present.

- The windows are inside the map area; this creates an issue with the way the windows (views) appear and disappear. Furthermore, it is not obvious how the separate windows interact with each other. The default size of these windows compared to the one of the map is a little bit unbalanced, as the map may be seen initially as less important or secondary to the top windows. In the case of the bookmarked indicator, it even obstructs the map, while the user must hover versus click to get data into all the windows.
- There is a problematic persistence of the data of the last choice made: when you hover over a country, for example, and then take your mouse pointer off, the memory keeps the last choice. Furthermore, hover and click behaviours are inconsistent between views.
- The time bar is not very obvious on the representation, although it is a very central feature.
- It is not obvious where the cursor is, since scrolling with the use of the keyboard arrows makes *two* of the windows respond (indices & countries).

There are also some interface inconsistencies:

- The column chart icon in the graph window shows column chart view of data whereas in the lower left of the application the same icon means toggle the entire graph window on or off.
- The left and right windows have a useful "- +" button which resizes them. The middle one does not have this button.

4.2. European cluster observatory

The European Cluster Observatory (ECO) [33] is a knowledge platform for regional competitiveness and clusters. It provides a single access point to information and analysis of clusters and cluster policy in Europe. It produces analysis and reports on regional competitiveness conditions, transnational cluster networks, clusters in emerging industries and studies on better practices in cluster organisations. The purpose of the European Cluster Observatory is to inform policy makers, cluster practitioners and researchers throughout the world about European clusters, cluster policies and cluster initiatives. ECO is managed by a consortium led by the Centre for Strategy and Competitiveness (CSC) at the Stockholm School of Economics. It is financed by the European Commission, DG Enterprise and Industry, under the Europe INNOVA initiative, and the Competitiveness and Innovation Framework Programme (CIP).

ECO provides two main visualisation sections:

- The Cluster Mapping of Regional Clusters (Figure 39) is based on 38 cluster categories and 259 NUTS-II regions. It provides statistical information on the geographic concentration of various industries and indicators of economic performance as well as employment and performance statistics. Users can access data for a selection of indicators, sectors and regions.
- The Organisation Mapping (Figure 40) lists regional and local private-public partnerships focused on cluster improvements.

Its visualisation software is based on Google Maps and Google Chart Tools. Motion Chart in particular is a dynamic flash based chart to explore several indicators over time. It requires bubble name, time and 2 columns of numeric values. Optional columns may be numeric values or categories.

On the whole, ECO offers a clear and simple interface that offers several visualisation options. However, there are several issues such as the screen clutter when the data are many and the selection tool issues. Such shortcomings by no means lessen ECO's added-value, and the aforementioned observations are only meant to identify possible areas for improvement to be taken into account by the **e·nventory** project.

The following sections focus on the information visualisation aspects of the ECO visualisation tools.

4.2.1. *Representation*

The Cluster Mapping tool gives a wide selection of statistical information on clusters and regions. There are many options for viewing the data and comparing countries or regions across sectors for the available indicators for a selected year or years.

The main screen of the tool is clear and concise with the map on the left occupying the main part of the screen and the selection of regions and indicators on the right.

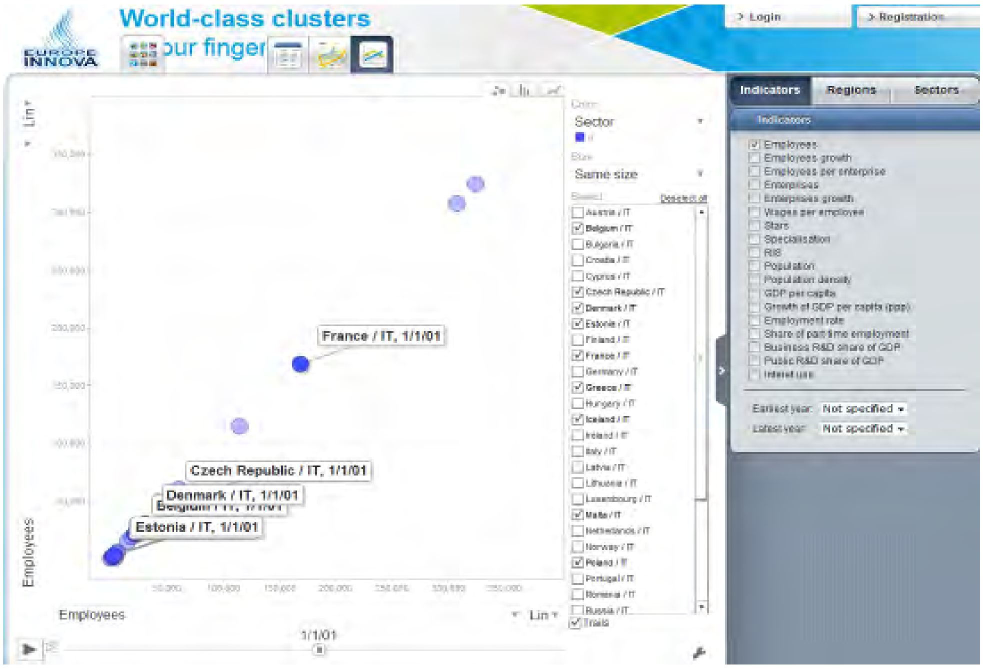

Figure 39: The European Cluster Observatory Cluster Mapping [33]

Several representation options are available:

- Map: the user can visualise on a Google Earth map the volume of a selected indicator for all the countries that belong to the selected region. The country circle represents a volume measure. Each colour in the circle represents a different service of the selected sector. In this view only one indicator and one year can be depicted.
- Moreover, for certain indicators (e.g. GDP per capita) the selection of sectors should be removed for the volumes to be visualised in the map.
- Table: the user can select indicators to be presented in a table with the countries in alphabetical order.
- Graphs: the data can be visualised in a plot, bar or line graph. This option allows for easy comparison between all countries of a region or for a selection on two-variables indicators. The position of the bubbles is determined by the indicators selected for the horizontal and vertical axes. For the size and the colour of the country circle there are three options: Same size/same colour or the size/colour to represent the volume of one of the two indicators.

The bubble visualisation used in the case of the map is clear, however, it is not as informative as it could be. In order to see the value that a specific bubble represents the user has to click on the bubble. There is, however, the possibility for mouse over, but it displays the country name only.

On the other hand, the graph option is quite informative, clear and concise, but only in the case where not all the data (e.g. all the sectors) are selected. When such a selection takes place there is the chance that the visualisation becomes unreadable. There are again, as in the case of the map visualisation, interactivity issues that will be discussed later in this section.

In the Organisation Mapping application, a user can choose to display cluster organisations (more than 1,100) on a map. No other visualisation options apart from the map are available. The issue that arises with this visual representation is that, with this number of data, it is rather cluttered.

A nice feature is the layout of the map and the selection window on the right. There is the possibility to minimize this window on the side in order to make more space for the map. However, there is also a legend for the colours on the left, which is not movable or retractable.

In general the representation is not overloaded with tools and many options; it is clear and not cluttered. On the other hand, it is not always immediately informative.

Figure 40: The European Cluster Observatory Organisation Mapping[33]

4.2.2. *Interactivity*

All the ECO visualisation windows offer a high degree of interactivity. It offers dynamic query interactivity, using the selection window on the right. Zooming and panning options are also available. One problem identified with zooming is that after a certain zooming level the bubbles/charts disappear.

Direct selection and manipulation are implemented, maybe a little bit more successfully in the graph visualisation, as the map offers direct selection on the map bubbles or graphs, but the user has to click on the bubble to get the indicator value whereas hovering over the bubble shows the country name. This is also the only point where the map overview is combined with details on the visualised elements. Rearranging of the view is accomplished to a basic degree in the case of the selection window on the right side. Lastly, the method implements the changing of representations as it offers map and graph options.

A minor problem that affects the interaction is the slight delay in certain cases when the visualisation should be refreshed.

4.2.3. *Usability*

At first glance the visualisation seems easy to use and straightforward. However, there are some usability issues which affect its smooth operation. It is not clear at every time what are the user selections, indicators, regions and sectors. In several cases when the user makes a combination of data that cannot be visualised, an error message is produced, explaining the problem. Alternatively, the system should make sure that this selection is not possible instead of allowing the user to make it and then say that it is wrong. Maybe a more efficient way to implement the dynamic querying options is needed. Last but not least, ECO lacks tools to customize visualisation options (colours, etc).

4.3. The Data Visualizer and the World at a Glance - WBG

The World Bank Group (WBG) consists of five international organisations, owned by 187 member countries, with different roles in the mission towards inclusive and sustainable globalization and to fight poverty and enhance the quality of living in the developing world. It provides technical and financial assistance to developing world countries. It works on many issues including health, education, and economic policy with more than 2,000 indicators.

The World Bank coordinates statistical and data work and maintains a number of macro, financial and sector databases. The user may access the data and datasets in several ways and there are also several presentation methods to explore the data.

The Data Visualizer Bubble Chart is a World Bank tool which employs bubble charts to display data. In each chart, the size of the country circle represents a volume measure, such as population or GDP. The position of the bubbles is determined by the indicators selected for the horizontal and vertical axes. The time series used ranges from 1960 to 2008 (Figure 41).

The World at a Glance is another visualisation tool available in the Data section of the World Bank. It features graphs of certain indicators as well as an overview of an indicator on the map. Clicking on an element on the map (indicator for a country) returns the data available for this country along with graphical representations (Figure 42).

The World Bank visualisation tools are effective as visualisations and present only minor usability issues. Any shortcomings identified by no means lessen World Bank tools' added-value, and the aforementioned observations are only meant to identify possible areas for improvement to be taken into account by the **e·nventory** project.

4.3.1. Representation

Data Visualizer Bubble Chart: it is a very attractive and at the same time simple and effective implementation of a bubble chart. The colours and the design are relaxing and appealing. The screen options and features are arranged so as to provide easy access to the available options without making the display cluttered. The indicator selector tool on the left provides easy access to definitions with pop-ups. One of the most interesting features is the time animation which shows the gradual change of the selected indicators in time.

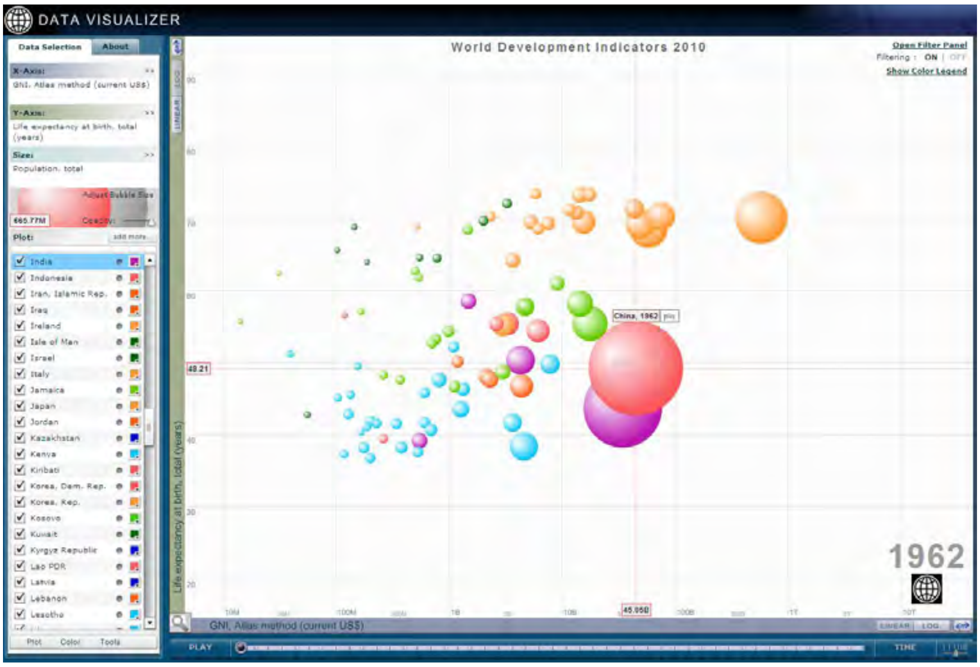

Figure 41: The World Bank's Data Visualiser tool[34]

A possible shortcoming is the country - colour selector on the left. The bubble colours are not really helpful to identify the countries, and the user has to actually move the mouse pointer on the bubble to see which country it represents.

The World at a Glance uses simple line graphs in combination with a map graph to provide a visual representation of the available indicator data. The visualisation is clear and concise; however the surrounding textual and image elements are a bit distracting.

4.3.2. *Interactivity*

Data Visualizer Bubble Chart is based on a dynamic query interface and it offers direct selection possibilities on mouse click and on mouse over. Details on demand are available on the bubble chart as when a bubble is selected values are presented along with the country name on the bubble. There is a zooming option available, but the possibility for panning is not available and the zooming needs to select the zoom-in or zoom-out tool every time the user wishes to change the zooming level.

The World at a Glance map visualisation with its surrounding supporting charts is simple and effective in its interactivity options. Zooming is available and clicking on a bubble retrieves the information on the selected country and presents it in a new page. Details on the selected bubble are also provided on mouse over.

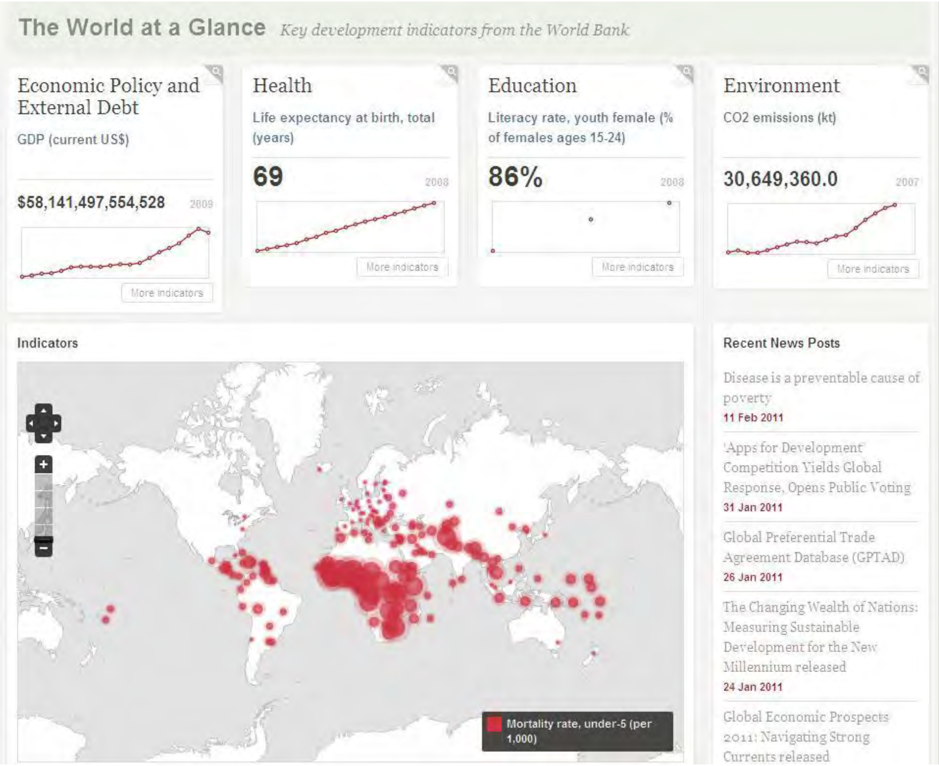

Figure 42: The World Bank's The World at a Glance tool for visualizing key development indicators[35]

4.3.3. *Usability*

Data Visualizer Bubble Chart interface is instinctive and clear. The axes selectors are easy to use, and well designed with the definition feature as a plus. The zooming tool is not so effective as after zooming in or out once, the user has to click again on the tool to activate it. There are several supporting tools available at the "Tools" option at the bottom left side, like exporting the visualisation in jpeg and also the animation. The main usability issue in this data representation tool is the Plot left bottom tool. It is not immediately clear how the user can take advantage of this tool and how it works, what is the purpose of the "Show trail" dot at the left of the colour and what exactly affect the check boxes next to the country names.

The World at a Glance is simple and does not present any particular usability issues.

5. Conclusions

The present document described the principle features of information visualisation, namely representation and interaction, by presenting examples of the various techniques used to represent and interact with data throughout a wide spectrum of applications. The goal of this has been to set a foundation for referencing case-studies of information visualisation tools available on-line, in order to reveal interactive visualisation features adopted by data observatories that are related to the **e·nventory** work.

Even though the interactive visualisation of eInfrastructures-related data, i.e. of data concerning grids and supercomputers, is sparse, several international data observatories have adopted information visualisation techniques that provide their global users with multiple visualisation tools, which, in many cases are highly interactive and visually appealing.

What is evident from the aforementioned case studies and critique is that these visualisation systems can enable people to look at large amounts of complex data and quickly find the information they want, to navigate and interact with data more easily, to recognize patterns and trends, and to obtain a better understanding of the information [13].

However, many of these tools are also too complex for a broader public and, consequently, suffer from usability problems. As Tufte [17] notes, what is to be sought in designs for the display of information is the clear portrayal of complexity. Not the complication of the simple; rather the task of the designer is to give visual access to the subtle and difficult – that is, the revelation of the complex.

In the framework of the **e·nventory** project, visualisation of information is central to the intuitive operation of the **European eInfrastructures Observatory** portal; it is imperative that the communication of impact assessment metrics and benchmarking indicators to the stakeholders takes place in a clear, unambiguous, and effective way, otherwise the service may not be utilised by the eInfrastructure community to the desired extent possible.

The variety of representation and interaction techniques that have been discussed herein, as well as the related reference case-studies, will constitute vital input to the **e·nventory** consultation activities that follow. In particular, it complements the analysis of international observatories, repositories and indicators as well as the proposal for the first set of **European eInfrastructures Observatory** indicators.

This document also helps shape the functional specifications, including its architecture, web engine, data collection, storing, and updating mechanisms, as well as its representation and interaction methodology, and usability design of the **European eInfrastructures Observatory**'s pilot web platform, the "Indicators Visualiser", by addressing the needs of different user groups and use cases, an aspect that seems to have been overlooked by the cases studied.

In conclusion, the review of techniques used in the selected reference cases informs the choice of representation and interaction features to be developed by the **e·nventory** project's on-line visualisation tool. In particular, the main outcome of this review emphasises the importance of clarity in the representation and usability in the interaction.

The representation of too much information into a single environment, as is the case with the impressive HDR/UNDP visualisation application, may compromise understandability, while the addition of too many data selection options, as is the case with the European Cluster Observatory interface, may result in clutter and inflexibility. Such shortcomings by no means lessen the added-value of these tools, but provide **e·nventory** with valuable suggestions and recommendations for the design and development of an effective visualisation platform.

-- Page intentionally left blank --

e·nventory
The European
eInfrastructures
Observatory
www.enventory.eu

e-infrastructure

CAPACITIES