Research Challenge on Visualization*

David Osimo[1] and Francesco Mureddu[2]

Draft

* The research activity leading to this paper has been funded by the European Commission under the activity ICT-7-5.6 – "ICT Solutions for governance and policy modeling" within the Coordination and Support Action (FP7-ICT-2011-7, No. 288828) CROSSOVER project "Bridging Communities for Next Generation Policy-Making". The views expressed in this paper are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission

[1] Tech4i2 ltd., UK - david.osimo@tech4i2.com

[2] Tech4i2 ltd., UK - francesco.mureddu@tech4i2.com

Background

In this short paper we present an outline for discussion upon a new Research Challenge on Visualization. This research challenge has been developed in the scope of project CROSSOVER "Bridging Communities for Next Generation Policy-Making" in the view of the definition of a new Research Roadmap on ICT Tools for Governance and Policy Making, building on the model and the research roadmap developed within the scope of the CROSSROAD project[3], but with a stronger focus on governance and policy modeling. To this aim CROSSOVER focuses on amending two Grand Challenges, already part of the CROSSROAD roadmap: GC1 - Model-based Collaborative Governance and GC2 - Data-powered Collective Intelligence and Action. Each Grand Challenge consists in a number of research challenges. In particular the Grand Challenge 2 embeds the research challenge "Intuitive, collaborative visual analytics of data for policy making", which we aim to amend, update, improve and validate during the workshop.

Introduction and definition

As the Google CEO Eric Schmidt pointed out in 2010, currently in two days is created in the world as much information as it was from the appearance of man till 2003. This is due to the explosion in computing techniques, which led to the generation of a tremendous amount of data which are stored in the internet and processed in the IT systems all over the world. In fact as predicted by CISCO[4], by 2015 the annual global IP traffic will reach 966 Exabytes (10^{18} bytes) (nearly a Zettabyte (10^{21} bytes)), increasing fourfold from about 900 Petabytes (10^{15} bytes) back in 2000 and around 2,500 Petabytes in 2010[5]. But data are not only stored in the internet, rather in an exponentially increasing number of IT infrastructures. Some

[3] Osimo, D. et al., 2010. The CROSSROAD Roadmap on ICT for Governance and Policy Modeling.

[4] CISCO (2011). VNi Forecast. Available at: http://www.cisco.com/en/US/netsol/ns827/networking_solutions_sub_solution.html#~forecast

[5] Wikipedia (2011). Internet Traffic, Available at: http://en.wikipedia.org/wiki/Internet_traffic

examples of new technologies for data collections[6] are: web logs; RFID; sensor networks; social networks; social data (due to the Social data revolution), Internet text and documents; Internet search indexing; call detail records; astronomy, atmospheric science, genomics, biogeochemical, biological; military surveillance; medical records; photography archives; video archives; large-scale eCommerce. In fact, in order to manage this huge amount of data, when it comes to human-computer interaction there is a need to distil the most important information to be presented it in a humanly understandable and comprehensive way. Here it comes visualisation, which is a way to interpret and translate data from computer understandable formats to human ones by employing graphical models, charts, graphs and other images that are conventional for humans[7]. In a sense we can define visualisation as any technique for creating images, diagrams, or animations to communicate a message or an idea. Since from the beginning of human history, visualisation has been an effective way to communicate both abstract and concrete ideas. The appearance of digital visualisation led to the development of graphic hardware as well as to a wide array of technique used to visualize data in a number of ways[8]. Often visualisation is needed to enable interaction[9] and to demonstrate how an operation works and which results are generated and derivable. In this view visualisation is massively used to provide the results of a simulation to users as well as to receive feedback and promote interaction. The connections between simulation and visualization appears even more clear when dealing with user interface, which enable the visualisation to take user commands. Another field related to visualisation is visual computing, which refers to computing which allows interacting and working with visual images or objects' representations. The visual computing field, which is described as "the entire field of acquiring, analysing and synthesizing visual data by means of computers"[10], includes photographs, 3-D objects and scenes, video sequences, block diagrams.

The disciplines in the domain of visualization are: Human-Computer Interaction (HCI), Usability Engineering, Cognitive and Perceptual Science, Decision Science, Information Visualisation, Scientific Visualisation, Databases, Data Mining, Statistics, Knowledge Discovery, Data Management & Knowledge Representation, Presentation, Production and Dissemination, Statistics, Interaction, Geospatial Analytics, Graphics and Rendering, Cognition, Perception, and Interaction.

Why it matters in governance

Today's governments face the challenge of understanding an increasingly complex and interdependent world, and the fast pace of change and increased instability in all the areas of regulation requires rapid decision making able to draw on the wider amount of available evidence in real-time. Moreover especially during crisis citizens take a proactive role in contributing to the policy-making process. There is a huge amount of data the governments can use in supporting their decisions, and visualization tools are increasingly being used to make sense of these very same data, turning information overload into an opportunity. In the policy context, it allows for more meaningful, evidence-based policy debate. It proves particularly relevant to make public data such as government spending meaningful for non-experts, but it also allows experts to dig into data in order to extract patterns and validate models. It is above all effective when dealing with complex and non-predictable patterns, such as those related to assessing and anticipating public policy impact. Policy advisors especially need advanced tools to quickly and iteratively analyse data and their implication

---

[6] Source: wikipedia

[7] B. Bederson, B., Shneiderman (2003). The Craft of Information Visualization: Readings and Reflections, Morgan Kaufmann

[8] J. van Wijk. (2005). The Value of Visualization. In IEEE Visualization Conference. pp. 79-86

[9] B. Thomas, P. Ahrweiler (Ed.). (2005). Internationale partizipatorische Kommunikationspolitik - Strukturen und Visionen. Münster: LIT. Available at: http://www.ucd.ie/management/petra-ahrweiler.htm

[10] Saarland University (2011). Visual Computing. International Master Programme of Study Available at: http://www.master-visual-computing.de

without necessarily specific knowledge about underlying algorithms. Visualization is one of the most advanced fields in policy modeling, being able to foster the design of more effective and efficient policies, as well as to make sense of large datasets, such as those provided as open government data. In fact the huge increase in data availability is also due to the so called "open data" movement, characterized by the fact that all across Europe and the US, governments are increasingly publishing their data repositories for other people to access and use it. In this view visual analytics allows turning the information overload into an opportunity, allowing for more meaningful, evidence-based policy debate. Multiple visualisation enable illustrating multi-faceted problems, such as policy-related issues, thereby facilitating informed debate. Arguments and graphical representations of reality-like structures and computer assisted visualization policies are able to stimulate policy modelling by increasing participation from the public, thereby ensuring an appropriate visualisation can therefore be considered a key component of a mature democracy.
In particular there are three principles according to which visualization can help with open government: statistical graphics ground debate in reality; new visualizations aimed at words rather than numbers hold out the hope of providing unfiltered insight into the minds of politicians and citizens alike; citizens are becoming visually literate and able to understand the message of complex visualizations. On the other hand visualization is of a pivotal importance when assessing the potential benefits stemming from the use of ICT for Governance and Policy Modelling. In fact when evidence is supported by advanced visualisation ambiguity and prejudices related to the decision are ruled out. This allows for concentrating on real issues as well as carrying out wider more easily and timely, contributing to a significant reduction of the amount of time needed for the decisions-making process. Moreover tools for scenario design, simulation and forecasting, when assisted by models and visualisation techniques allow a better understanding of the effects and consequences of policy decisions, and enable timely evaluation of alternatives.

Recent trends

Traditionally the first examples of visualization date back to the 19th century with the drawings[11] by Charles Joseph Minard (1781-1870), who developed a format to show data tied to a timescale with a landscape background. In 1869 Minard applied its drawings to show the march of Napolean's army towards Moscow, starting with 422,000 and ending with 10,000 men, and Hannibal's crossing of the Alps, starting with 97,000 and ending with 6,000 men. The modern visualization field, making use of computer graphics, originated in the late 1980s with the studies on scientific visualization applied to fluid dynamics, volume visualization, molecular modeling, imaging remote-sensing data, and medical imaging[12]. Some more recent areas, such as information visualization, mobile visualization, location-aware computing and visual analytics arose around 2000. According to Rosenblum[13] the new trends in the field involve the integration of visualization techniques with areas such as machine vision, data mining and data bases to promote broad-based advances. Another trend deals with the combination of algorithms with usability studies.
Nowadays the hottest topic in the field is information visualization, which is becoming a mainstream tool due to the increase in the information availability. In fact a massive quantity of tools, critical for analysts and researchers, but also for common people, is now available online.

---

11 http://www.math.yorku.ca/SCS/Gallery/minbib/index.htm

12 Lawrence J. Rosenblum (1994). "Research issues in scientific visualization". In: *Computer Graphics and Applications, IEEE*. Volume 14, Issue 2, March 1994 Page(s):61 - 63.

13 http://www.cs.stonybrook.edu/~cgi05/cgi_main_files/funding_panel.html

Inspiring cases

Decision Theather's applications: Watersim[14], Decision centre for a desert city[15] and Biodiversim[16]
InstantAtlas's applications[17]
Information Aesthetics[18]
DebateGraph[19]

Policy applications of visualization tools

With regard to the governance context, some visualization tools, such as Decision Theater[20], can be applicable to a wide array of issues and situation (education, environment, public health, urban growth). Other examples of visualization tools applications are:

- Demographics visualizations, allowing stakeholders and decision makers to have a clear picture of the data and of their trends over time. Visualisation of demographic data make easier the design and evaluation of various policies, as there is no need to dig through acres of numbers. In fact advanced algorithms are able to create figures and illustrations easy to interpret. Typical examples of such tools include the GapMinder[21] (which embeds visualizations of various demographic data at global level), as well as Dynamic Choropleth Maps[22], DataPlace[23], Hive Group[24], Name Voyager[25], State Cancer Profiles[26]Legal Arguments visualisation: text analysis, argumentation mappings and visualisation algorithms can be applied to legal documents in order to simplify legislation making it more accessible and comprehensible to the general public, or in order to visually represent corroborative evidence (e.g. the tools Carneades[27] and Deflog[28])
- Discussion Arguments visualisation, making use of visualisation techniques for visualizing the flow of a discussion that include various arguments, in order to instantly get awareness of the topics discussed, as well as of the arguments and the support such arguments gain. In this view visualisation supports all interested stakeholders to understand the flow of a discussion, which is presented to them in a structured and interactive format, avoiding numerous discussion threads. Example of such visualisation tools include DebateGraph[29], which is intensively used for building argumentation maps, as well as Araucaria[30], Compendium[31], Argublogging[32] and Rationale[33]

---

[14] http://watersim.asu.edu
[15] http://dcdc.asu.edu
[16] www.tradeoffs.org
[17] http://www.instantatlas.com/clients.xhtml#government
[18] http://infosthetics.com/
[19] http://www.debategraph.org
[20] http://dt.asu.edu/
[21] http://www.gapminder.org
[22] http://www.turboperl.com/dcmaps.html
[23] http://www.knowledgeplex.org/dataplace.html
[24] http://www.hivegroup.com/gallery/worldpop/
[25] http://www.babynamewizard.com/voyager#
[26] http://statecancerprofiles.cancer.gov/micromaps/
[27] http://carneades.berlios.de/downloads/
[28] http://www.ai.rug.nl/~verheij/aaa/
[29] http://www.debategraph.org
[30] http://araucaria.computing.dundee.ac.uk/
[31] http://compendium.open.ac.uk/institute/
[32] http://www.arg.dundee.ac.uk/?p=624
[33] http://rationale.austhink.com/

- Geovisualization, which is based on the provision of theory, tools and methods for visual analysis, synthesis, exploration and representation of geographical data and information in order to derive problem specific models and design task specific maps for incorporating geographical knowledge into planning and decision making. Some examples of such tools include ESTAT[34], GeoViz Toolkit[35], the geovisualization tools at the US National Cancer Institute[36], some applications of InstantAtlas[37]
- Advanced visualization applications used for security and national defense. In this fields, software advances are being led both on the military and on the corporate front. In fact business organizations also have urgent information visualization requirements that support their business intelligence and situational awareness capability, data mining and reporting requirements. In this view many of the software innovations are being targeted at financial and corporate requirements, but are also applicable to the defense domain due to common data mining and information visualization challenges. Examples of such tools are: DataMontage[38], HoneyComb[39], Oculus GeoTime[40] and Starlight[41]
- Visualization applications adopted for financial markets monitoring and visualizing in real time. An example of such tool is SmartMoney[42]

Tools on the market

There is a massive quantity of visualization tools in the market, both freely available and enterprise level, let us shed some light.

Freely available tools

First of all we have visualization websites useful for sharing and presenting data, provide clear context on important cultural, environmental, social and economic issue, build chart and share visualization and discoveries. Such examples are Data360, FlowingData, Hohli, IBM Many Eyes,

Then we have data visualization tools used for plotting data on maps, frameworks for creating charts, graphs and diagrams and tools to simplify the handling of data transforming them into spreadsheets, visual data mining and database exploration system, data visualization system for high-dimensional data, visualization framework for animating data. Some examples of those tools are: Data Wrangler, JavaScript InfoVis Toolkit, VisDB, Graphviz, IBM OpenDX, Gephi, GeoCommons, Miso Dataset, Polymaps, Processing, Protovis, Raphael, Tableau Public.

---

34 http://www.geovista.psu.edu/ESTAT/

35 http://www.geovista.psu.edu/geoviztoolkit/index.html

36 http://gis.cancer.gov/nci/geovisualization.html

37 http://www.instantatlas.com/clients.xhtml#government

38 http://www.stottlerhenke.com/datamontage/examples/madcap/Air_force_wargame_simulation.htm

39 http://www.hivegroup.com/solutions/demos/merit.html

40 http://www.oculusinfo.com/papers/GeoTime_Brochure_06.pdf

41 http://starlight.pnl.gov/

42 http://www.smartmoney.com/map-of-the-market/

Finally freely available visualization tools are very common in the aforementioned argument visualization field, where they are used to map arguments and discussions in order to instantly get awareness of the topics discussed, as well as of the arguments and the support such arguments gain. Examples include: Rationale, Araucaria, Truthmapping, Argumentations, Debategraph, Idebate

Enterprise-level software

Apart from free visualization tools, there are also many more advanced software which are used by firms in order to satisfy their information visualization requirements for business intelligence support and situational awareness capability, as well as data mining and reporting requirements. Other uses include enterprise knowledge visualization, linking knowledge to spatial data, online analytical processing and data mining, advanced social network analysis and visualization, data mining and interactive visualization, communication of location-based statistical data, on-line and batch environment for business graphics, information visualization tools for high dimensional non-linear data, visual analysis of data in spreadsheet format, analysis of high volumes of unstructured text, analysis of high-dimensional data in large complex data sets and of multivariate time-oriented data.

Some examples of such software are: CViz Cluster Visualization, IBM ILOG Visualization, Spotfire, Survey Visualizer, Infoscope, Inspire, Sentinel Visualizer, Grapheur 2.0, InstantAtlas™, Miner3D, VisuMap, Drillet, Eaagle, GraphInsight, Gsharp

Other examples of visualization software can be found in
http://groups.diigo.com/group/crossoverproject/content/tag/visualization

Key Challenges and Gaps

Policy modelling lacks visualisation techniques able to navigate through text, analyse large text repositories, analyse opinions, and to interactively simulate policy decisions (Kohlhammer et al., 2010[43]). New tools like the Word Tree[44], Treemap, Tag Cloud and Bubble Chart[45] are available but lack interactivity.
Another possible challenge is to adapt existing techniques to policy modeling:

- CirVis3D[46], which can visualize clustered opinion snippets as well as display time series in order to show the opinion trends over time
- RelaNet[47], which displays the network relations and thereby is able to show the connections and co-variances of the different opinions overtime

[43] Kohlhammer, Jörn, Ruppert, T., Davey, J., Mansmann, F., & Keim, D. (2010). Information Visualisation and Visual Analytics for Governance and Policy Modelling. In: *CROSSROAD Call for Contributions on Future Internet on ICT for Governance and Policy Modelling*

[44] Wattenberg, M.: The Word Tree, an Interactive Visual Concordance, IEEE Transactions on Visualization and Computer Graphics, Vol. 14, No. 6, Ohio, Columbus, USA (2008)
[45] See http://manyeyes.alphaworks.ibm.com/manyeyes/, which can be also found in the Diigo collection……

[46] Landesberger, T. von, Goerner, M., Schreck, T.: Visual Analysis of Graphs with Multiple Connected Components. IEEE Symposium on Visual Analytics Science and Technology (2009)
[47] Landesberger, T. von, Knuth, M., Schreck, T., Kohlhammer, J.: Data Quality Visualization for Multivariate Hierarchic Data, IEEE Information Visualization Conference (INFOVIS), Columbus, OH, USA (2008)

But nevertheless what is missed is a better interaction of visualization approaches and analytical processes of text mining, as well as a better integration between new opportunities for data collection, such as open data and participatory sensing, policy modelling and visual analytics tools. Moreover visualisation is largely a demand- and design-driven research area, which struggles to be compatible with existing FP7 instruments. Additionally, a strong lightweight collaboration with designers and visual artists is difficult to fit in existing policy tools.

Current research

- Close the loop of information selection, preparation and visualisation
- Simultaneous multiple visualisation
- Integration of visualisation with comments / wiki / blogs
- Collaborative platform display
- Interaction between visualisation and models
- Mobile visual analytics tools
- Geo-visualisation of government data
- Integration with opinion mining and participatory sensing
- Evaluation framework for visualisation effectiveness
- Visualisation infrastructures for policy modelling issues

Future research: long term and short term issues

Short-term research

- Re-usable, mashable tools for visual analytics
- Tighter integration between automatic computation and interactive visualisation
- Bias identification and signalling in visualisation
- Perceptual, cognitive and graphical principles
- Efficiency of the visualisation techniques to enable interactive exploration interaction techniques such as focus & context
- Impact evaluation of visual analytics on policy choices

Long-term research

- Learning adaptive algorithm for users intent
- Advanced visual analytics interfaces
- Intuitive affordable visual analytics interface for citizens
- Development of novel interaction algorithms incorporating machine recognition of the actual user intent and appropriate adaptation of main display parameters such as the level of detail, data selection, etc. by which the data is presented

Summary Overview

Freely Available tools	Enterprise Level Software	Policy Applications	Current research	Short term future research	Long term future research
- Visualization websites useful for sharing and presenting data -Data visualization tools -Argument visualization tools	-Information visualization requirements for business intelligence and situational awareness -enterprise knowledge visualization linking -Online analytical processing and data mining - Advanced social network analysis and visualization -Data mining and interactive visualization communication of location-based statistical data -Information visualization tools for high dimensional non-linear data -Visual analysis of data in spreadsheet format	- Demographics visualizations, allowing stakeholders and decision makers to have a clear picture of the data and of their trends over time - Legal Arguments visualisation: text analysis, argumentation mappings and visualisation algorithms - Discussion Arguments visualisation, making use of visualisation techniques for visualizing a discussion's flow -Geographic visualization tools -Financial markets monitoring and visualizing in real time -Advanced applications for security and defense	-Close the loop of information selection, preparation and visualisation - Simultaneous multiple visualisation -Integration of visualisation with comments / wiki / blogs - Collaborative platform display Interaction between visualisation and models -Mobile visual analytics tools -Geo-visualisation of government data -Integration with opinion mining and participatory sensing -Evaluation framework for visualisation effectiveness - Visualisation	-Re-usable, mashable tools for visual analytics -Tighter integration between automatic computation and interactive visualisation -Bias identification and signalling in visualisation -Perceptual, cognitive and graphical principles -Efficiency of the visualisation techniques to enable interactive exploration interaction techniques such as focus & context -Impact evaluation of visual analytics on policy choices	-Learning adaptive algorithm for users intent -Advanced visual analytics interfaces -Intuitive affordable visual analytics interface for citizens -Development of novel interaction algorithms incorporating machine recognition of the actual user intent and appropriate adaptation of main display parameters such as the level of detail, data selection, etc. by which the data is presented

			infrastructures for policy modelling issues		