First publ. in: CROSSROAD Call for Contributions on Future Internet on ICT for Governance and Policy Modelling

Information Visualisation and Visual Analytics for Governance and Policy Modelling

Jörn Kohlhammer[1], Tobias Ruppert[1], James Davey[1], Florian Mansmann[2], Daniel Keim[2]

[1] Fraunhofer IGD, Fraunhoferstr. 5, 64283 Darmstadt, Germany
{joern.kohlhammer, tobias.ruppert, james.davey}@igd.fraunhofer.de
[2] University of Konstanz, Konstanz, Germany
{florian.mansmann, daniel.keim}@inf.uni-konstanz.de

Abstract. Professionals involved in governance and policy modelling have seen a dramatic increase in the volume of potentially relevant data. As in many other knowledge-based fields today, policy makers face the problem of an information overload. Putting data from sources as diverse as blogs, online opinion polls and government reports to effective use is a task which requires new tools for their analysis and visualisation. In this paper, we introduce the research field known as Visual Analytics and we argue that it can help policy makers integrate large amounts of heterogeneous data into their decision-making processes. In so doing, Visual Analytics has the potential to increase not only the relevance, but also the legitimacy of policy decisions. In addition, policy modelling could provide the Visual Analytics community with new research challenges.

Keywords: Information Visualisation, Visual Analytics, Policy Modelling

1 Introduction

We live in a world facing a rapidly increasing amount of data, which needs to be dealt with on a daily basis. Virtually every branch of industry or business, and many political or personal activities nowadays generate vast amounts of data. To make matters worse, the possibilities to collect and store data are increasing at a faster rate than our ability to use them for making decisions. In most applications raw data has no value in itself; instead we need to extract the information contained in it. This is the first step in the discovery of knowledge to aid in decision making.

The advent of the *Social Web* has given us a variety of participatory tools such as opinion systems, online social networking, blogs, wikis, and forums. These tools present government decision makers, governance bodies and civil-society actors with the possibility to bring about significant changes in the way future societies function. More precisely, the emerging technological environment allows for new processes of information and knowledge sharing across the interfaces of government and civil society. It allows the voicing of opinions, the expression of needs and the strengthening of participation in ways that can enhance participatory democracy. In this setting it is likely that the new information society will generate a massive amount of data detailing people's situations, what they think and what they believe. Policy makers, as well as civil-society groups, will have access to this data. They should be empowered to exploit them by extracting information from the data. The knowledge gained can be used by leaders and representatives to elaborate policies, to make decisions or to define rules of behaviour in a way that is well informed and legitimized by its fairness and transparency.

However, the exploitation of data generated within the Social Web is problematic if we consider the complex, fragmented nature of its sources, and its huge quantity. An additional challenge here is that public policies have to operate in complex, shifting environments, structured by changes at global, EU, national, and regional levels. Appropriate decision-making models, process flows or analytical and forecasting tools are needed to properly understand, interpret, visualise the information contained in data from the Social Web. Policy makers could then use feedback about their initiatives in order to align public policies with emerging civil society needs, requirements and expectations. Information and communication technology (ICT) can, thus, support policy makers in modelling their policy design and implementation.

In this context, *Visual Analytics* can help provide answers to two pressing questions in the quest for the optimal utilisation of ICT for governance and policy modelling. Firstly, how can we avoid being overwhelmed by this flood of data, knowing that its quantity will only increase in the future? Secondly, how can we make the best use of the information contained in the data and, in particular, how could it be presented in a way that could be used to inform the policy elaboration process in a more targeted fashion?

2 Information Visualisation and Visual Analytics

Information visualisation is a relatively young and interdisciplinary field that emerged from scientific visualisation during the last two decades. Information visualisation has its roots in the work of the French cartographer Jacques Bertin [1] who described a framework for the design of the basic elements of diagrams in 1967. Some years later Edward Tufte [8] developed a theory of data graphics, maximizing the density of useful information within the bounds of a diagram.

The scientific research field of information visualisation arose when the theories developed by Bertin and Tufte were first used in interactive computer applications by the user interface community, namely Robertson, Card and Mackinlay in the 1980s. By then the objectives of interacting with large amounts of information and dynamic querying had been identified. In the following decade, the number of visualisations increased, which led Ben Shneiderman to encourage the development of taxonomies and frameworks to describe visualisation technologies in 1996 [7]. One of the most famous works in information visualisation followed in 1999 from Card, Mackinlay and Shneiderman [2], who described the analysis of multidimensional and novel kinds of data, as well as formalising the information visualisation process. The multidisciplinarity of the field was shown by Ware in 2004 [9]. He emphasised the importance of human perception in information visualisation.

New problems emerged from the ability of new technologies to store data at ever growing rates. The field of data mining, which developed parallel to that of information visualisation, took a different approach to finding information in massive amounts of data. The goal in this community was not to support human perception with interactive visualisations, but to allow computers to do the work of an analyst. Of particular relevance here were innovations in the fields of text mining and knowledge discovery in databases (KDD), which was formalised by Fayyad, Piatetsky-Shapiro and Smyth [4].

Visual Analytics as a new interdisciplinary research field seeks to combine the advantages of human perceptual abilities and the processing power of computers. The purpose of this new scientific field is to solve the information-overload problem. Visual Analytics brings data-mining and text-mining technologies, which can pre-process massive amounts of data, and information visualisation, which can be used to visualise important and neglect unimportant information, together. Visualisation becomes the medium of a semi-automated analytical process, where humans and machines cooperate for the most effective results. The user has the ultimate authority in specifying the direction of the analysis relating to a particular task. At the same time, the system provides effective means of interaction, allowing the user to concentrate on this task. Furthermore, a visual representation, sketching the path from data to decision, provides a reference for the collaboration of groups of users across different tasks and abstraction levels.

A great variety of issues have to be addressed when designing and implementing Visual Analytics applications. Visual Analytics research is, thus, highly interdisciplinary. In addition to information visualisation and data mining, it draws on the related research areas of data management, data fusion, statistics, and cognitive science (among others). All of these research areas produce rigorous scientific work, each in a vibrant research community. Visual Analytics research shows that bringing together these research efforts can lead to novel, highly effective analysis tools, contributing solutions to the information-overload problem in many important domains.

Visual Analytics is defined as “the science of analytical reasoning supported by interactive visual interfaces”. The term was coined by Jim Thomas in the research and development agenda “Illuminating the Path” [3], which is widely considered the intellectual fingerprint of the discipline. The need for a new discipline concerned with data analytics stems from the need to address hard analytical problems where neither the machine nor the human alone can efficiently and effectively find a solution. On the one hand, disciplines like data mining, KDD, artificial intelligence, statistics, etc., with their emphasis on automatic computation, do not provide enough support to the open-ended nature of modern data analysis problems. Human expert knowledge plays a key role in this domain and cannot be removed from the process. On the other hand, pure data visualisation solutions, like those offered by information visualisation, simply don’t scale to the complexity and the size of the data analysis problems we face today.

Visual Analytics Process

Visual Analytics can be explained by the visual analytics process, a series of steps that combine human and computational steps in an integrated fashion to meet an analytical goal. Fig. 1 shows the main elements involved (ovals) and the data transformations between them (arrows). Preprocessing and transformation are often the preliminary steps necessary to extract the data of interest and to format them in a way that fits the shape of the problem at hand. The analyst can then select between visual or automatic analysis methods. Mapping the data to a visual representation may directly lead to the desired knowledge, but it is more likely an initial visualisation is insufficient and further user interaction is needed.

Several iterations of data visualisation and interaction may lead to the construction of a model able to describe the process or phenomenon of interest. In its simplest form a model could simply be one or more hypotheses. The construction of a model using data-mining methods could also be the first step after data pre-processing. Once a model is created the analyst has the ability to interact with the automatic methods by modifying parameters or selecting other types of analysis algorithms. Model visualisation can then be used to verify the findings, to refine the model itself or to engage other analysts in the process. Alternating between visual and automatic methods is characteristic of the visual analytics process and leads to a continuous refinement and verification of preliminary results.

In the visual analytics process, knowledge can be gained from visualisation, automatic analysis, as well as the preceding interactions between visualisations, models, and the human analysts. The feedback loop stores this knowledge of insightful analyses in the system and contributes to enable the analyst to draw faster and better conclusions in the future.

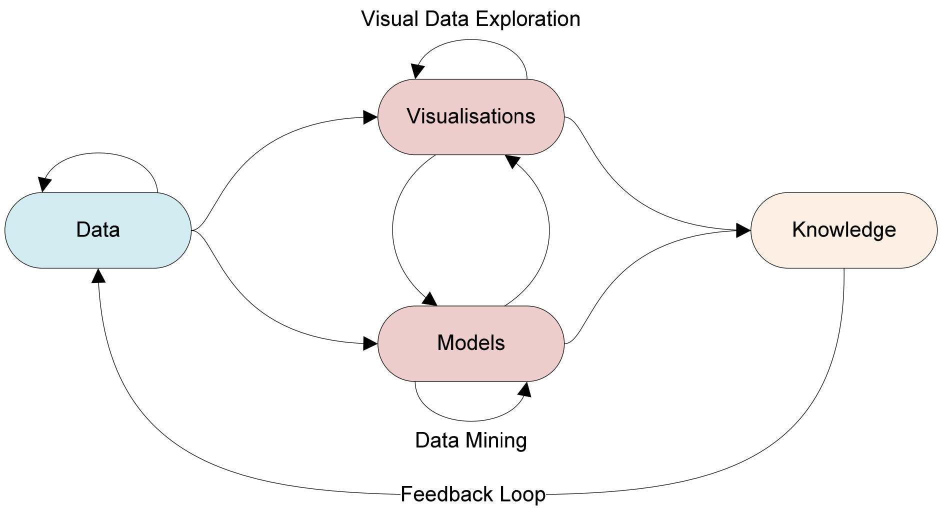

Fig. 1. The Visual Analytics Process is characterized through interaction between data, visualisations, models about the data, and the users in order to discover knowledge.

VisMaster

Visual Analytics is a field involving several individual disciplines. A current FET-Open Coordination Action, called VisMaster CA, unites research excellence from the fields of data management, data analysis, spatio-temporal data, and human visual perception research with the wider visualisation research community. The goal of VisMaster CA is to create awareness for Visual Analytics and to create a community of academic and industrial R&D excellence in Europe. Its major deliverable is a research roadmap for visual analytics in the EU, which is open to link to other similar activities like CROSSROAD. More information on VisMaster can be found at www.vismaster.eu.

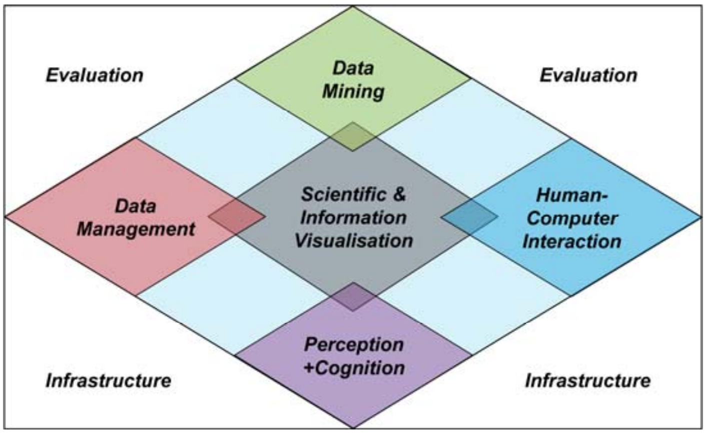

Fig. 2. Visual Analytics integrates scientific and information visualisation with core adjacent disciplines: data management and analysis, spatio-temporal data, and human perception and cognition.

3 Visual Analytics for Governance and Policy Modelling

Today's governments are faced with the challenge of understanding an increasingly interdependent and complex world. Especially during crises (e.g., the financial crisis in 2008) citizens demand effective and transparent but also sustainable solutions. Governments hold large amounts of data (e.g., about citizens, taxes, budgets, economic developments, etc.) that can support this effort. Moreover, the rise of the Social Web opens new opportunities to involve citizens in the democratic process by stating their opinions and ideas on the Web. In conclusion, there is enough data available to support governments in their political decisions, but the information hidden in the data has to be extracted before it can be used in government and policy modelling scenarios.

Visual Analytics as a research field which addresses precisely this problem: The extraction of information hidden in large data sets. The analysis of stakeholder opinions can, for example, be supported by Visual Analytics techniques. So far, in politics common visualisation techniques like bar charts, line charts, pie charts etc. are applied for showing election or poll results. At a certain level of complexity of the data these techniques become ineffective. In addition they are not appropriate, for example, for the visualisation of opinions in the form of text snippets.

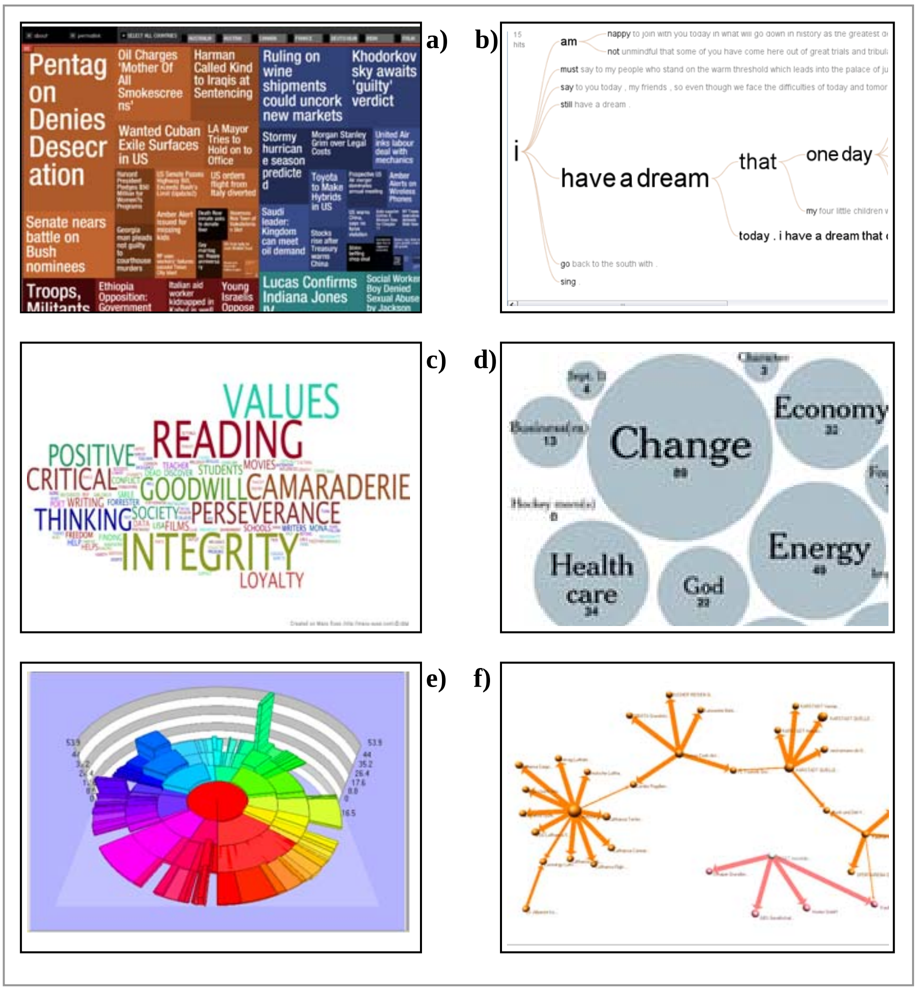

Fig. 3. Several visualisation techniques are shown that can be used for different stages in the policy modelling area: a) Treemap b) Word Tree c) Tag Cloud d) Bubble Chart e) CirVis3D f) RelaNet

The area of governance and policy modelling currently lacks visualisation techniques to properly display and navigate through text, to analyse large text repositories, to analyse opinions, and to interactively simulate policy decisions – to name but a few deficits. However, Novel visualisations like the word tree [10], treemap, tag cloud and bubble chart(see Figure a-d)[1] do exist. But they currently lack the desired degree of interactivity and some are still at an early stage of development.

The interesting challenge lies in adapting existing visualisation techniques to policy modelling and analysis. For example the visualisation technique CirVis3D (see Figure e) can visualize hierarchically clustered opinion snippets [6]. Moreover time series can be displayed to show the development of opinions over time. Another visualisation technique, RelaNet (see Figure f) [5] is able to display relations in networks. This can be adapted to the dependencies and connections between different arguments or opinions.

Besides the adaptation of existing techniques there is still the need for entirely new visualisation approaches and technologies to be developed in order to simplify the analytical process of text mining, for instance, or policy analysis, in general. Due to these intersections in the two research fields, the connectivity between the topics has to be enforced. The results of these efforts are, on the one hand, new solutions and tools for the governance and policy modelling processes. On the other hand, visual analytics will benefit from special challenges and use-case scenarios arising from governance and policy modelling.

4 Conclusion

Visual Analytics is an emerging field with lots of potential and opportunities to solve a wide range of information-overload problems. In particular, it involves the combination of human expertise and the computational capabilities of modern automated analysis methods through advanced visual interfaces. These enable decision makers to gain insight into the analysis process. This will potentially lead to more confidence in the analysis than if it were the simple numeric or tabular output of an automated analysis method.

In this position paper, we discussed what Visual Analytics research is, and how its current approaches and initiatives can be part of the future research agenda for governance and policy modelling. Existing visualisation and Visual Analytics techniques can be directly applied to certain steps in the policy modelling process. Furthermore, the field of governance and policy modelling has unique challenges that require a collaborative effort between various research disciplines including visualisation and Visual Analytics.

[1] http://manyeyes.alphaworks.ibm.com/manyeyes/

Acknowledgments. The project VisMaster CA acknowledges the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open grant number: 225924.

References

1. Bertin, J.: Sémiologie Graphique. Les diagrammes, les réseaux, les cartes, With Marc Barbut [et al.]. Gauthier-Villars, Paris (1967)
2. Card, S.K., Mackinlay, J.D., Shneiderman, B. (eds.): Readings in Information Visualization - Using Vision to Think, Morgan Kaufman Publishers, Inc., San Francisco, CA, USA (1999)
3. Cook, T., Cook, K.: Illuminating the Path: Research and Development Agenda for Visual Analytics, IEEE-Press, (2005).
4. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: The KDD Process for Extracting Useful Knowledge from Volumes of Data, Commun. ACM, Vol. 39, No. 11, New York, NY, USA (1996)
5. Landesberger, T. von, Goerner, M., Schreck, T.: Visual Analysis of Graphs with Multiple Connected Components. IEEE Symposium on Visual Analytics Science and Technology (2009).
6. 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)
7. Shneiderman, B.: The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations, IEEE Symposium on Visual Languages, Washington, USA (1996)
8. Tufte, E. R.: The Visual Display of Quantitative Information, Graphics Press, Cheshire, CT, USA (1983)
9. Ware, C.: Information Visualization - Perception for Design, Second Edition, Academic Press, San Diego, CA, USA (2004)
10. Wattenberg, M.: The Word Tree, an Interactive Visual Concordance, IEEE Transactions on Visualization and Computer Graphics, Vol. 14, No. 6, Ohio, Columbus, USA (2008)