Using Visual Analytics for Decision Making

Lars Schütz[1,2], Susanne Raabe[1,3], Korinna Bade[1], Matthias Pietsch[3]

[1]Anhalt University of Applied Sciences, Department 5, Köthen/Germany · lars.schuetz@hs-anhalt.de
[2]Otto von Guericke University Magdeburg, Faculty of Computer Science, Magdeburg/Germany
[3]Anhalt University of Applied Sciences, Department 1, Bernburg/Germany

Abstract: The planning of built and natural environments incorporates participatory processes that are more and more supported by information and communications technology (ICT). On the one hand, this implies potentially more diverse knowledge of the masses in form of participatory contributions, e. g., opinions, ideas or formal statements. On the other hand, this leads to an increasing data complexity in terms of size and interconnectedness which in turn complicates making sense, retrieving implicit new information, and decision making. We describe the need for more sophisticated methods and present how the field of visual analytics can overcome these challenges by demonstrating the applicability of the visual analytics process model for one exemplary but fundamental use case that focuses on an analysis task in e-participation for planning purposes.

Keywords: Visual analytics, e-participation, decision making, planning process

1 Introduction

Awareness of participation methods in planning processes is still growing. In spatial planning, for instance, new infrastructure projects, caused by the energy transition, will change the environments of people. In this regard, the Institute for Advanced Sustainability Studies has pointed out that, although the majority of people supporting the energy transition basically agree with each other, many infrastructure projects are faced with protests by the public and also by local authorities (RICHTER et al. 2016). This aspect shows that planning without full acceptance of a project is unrewarding. The legitimate basis for decision makers to approve a project is grounded in the public's acceptance. This acceptance is mainly linked to justice and fairness of a planning process (RICHTER et al. 2016). Acceptance builds on normative expectation horizons, namely the possibility of approval or acceptability of a decision.

A participation process mainly involves two sides, the organizer or decision maker of the process, e. g., public administrations and planners, and the participants, e. g., citizens and public agencies. It is a "deliberative process by which interested or affected citizens, civil society organizations, and government actors are involved in policy making before a political decision is taken" (EIPP 2009). Each side, or more precisely each member of each side, tries to accomplish various tasks while following different needs, imaginations and goals. Therefore, it is difficult to find consent. In other words, it is difficult to make correct decisions.

In this context, new digitally supported methods of information and communications technology (ICT) gain increasing attention in planning and decision processes that play an important role in the e-participation domain. ICT decouples the process from time and space constraints, i. e., people can basically participate from anywhere at any time. This is needed because people want to have a say and shape their environment. In the end, this means that more and more people are fundamentally able to participate in planning and decision processes. But this immediately leads to at least three major challenges we have identified, namely

Journal of Digital Landscape Architecture, 2-2017, pp. 94-101. © Wichmann Verlag, VDE VERLAG GMBH · Berlin · Offenbach. ISBN 978-3-87907-629-1, ISSN 2367-4253, e-ISSN 2511-624X, doi:10.14627/537629010.

1) handling the complex data situation, 2) exploring the process data, and 3) discovering practical knowledge (see Section 2.1).

Visual analytics (WONG & THOMAS 2004, THOMAS & COOK 2005) provides ways for dealing with the previously mentioned challenges. It tightly couples semi-automated data analysis and interactive visualization instead of focusing on solely one of these approaches. Visual analytics methods can assist in revealing useful insights. In this paper, we describe its fundamental concept (see Section 2.2).

In order to bring e-participation and visual analytics together, we present and describe one typical use case. It derives from an exemplary analysis task that planners and decision makers try to accomplish, namely the exploration and evaluation of contributions of planning processes. Thereby, we demonstrate the applicability and effect of visual analytics (see Section 3). In other words, this paper reviews the visual analytics process for potential use in a real world planning situation.

Finally, we briefly conclude the paper and suggest future work (see Section 4).

2 E-participation and Visual Analytics

The following section is about digitally supported participatory processes and visual analytics. These topics are fundamental domains that shall be brought together.

2.1 E-participation – Uncertainty in Practice

E-participation or online participation is, besides providing information and opportunities for consultation, about opening a channel for discourse between all participants that is at least partly supported or even fully carried out by ICT. This sounds promising but the European Institute for Public Participation stated a dilemma: "The handbooks are predominantly focused on face-to-face participation. Even when they talk about the opportunities offered by new technologies, their presentation is geared toward offline usage. They do not cover the opportunities offered by the internet and its interactivity" (EIPP 2009). Today, this divergence between theory and practice still exists. On the one hand, there is clear awareness of online tools, and there is also the desire for such systems that support the implementation and moderation of participatory processes, but on the other hand, especially organizers of participatory processes as well as involved decision makers see missing technical requirements, huge financial costs, and too complicated and time-consuming workflows as predominant reasons against the integration of ICT (HELBIG et al. 2016). The opportunities and benefits don't seem to be clear enough.

The growing number of potential participants and the integration of computer-aided methods as well as related tools into participation processes also lead to several key challenges that need to be solved for successful decision making:

1) *Handling the complex data situation:* Contributions of e-participation processes like ideas, comments, formal statements or documents comprise of many different data facets like natural language text, geospatial and time-oriented data, images, and rating or voting information. In an even more complex manner, contributions might additionally relate

to each other. An interconnected network of contributions evolves over time. It is challenging to store and represent the data in a proper way for future interpretation.

2) *Exploring process data:* The effective and efficient exploration of the accumulated data and making sense are a burden. Understanding the provided information is time-consuming and affords high cognitive demands due to the previously described data complexity.

3) *Discovering practical knowledge:* Discovering complex knowledge, such as similar or contrary contributions in the raw data, for supporting the planning process or decision making is currently based on manual analysis requiring substantial efforts. In the end, potential knowledge that might help in decision making remains hidden.

This means that opening a channel for sharing knowledge, opinions and ideas is simply not enough. In other words, what can all the data and their interconnections provide if we are not able to handle it in its entirety? Organizers need guidelines and support for moderation / facilitation and analysis. The goal is acceptance of the result by the participants. A comprehensible and clear decision is needed to convince more people. Overall, computer-aided methods and related tool support for data exploration and knowledge discovery are relevant in the e-participation domain. Properly combined methods for data analysis, visualization and human computer interaction are needed. These challenges are interwoven and need to be addressed in conjunction with each other. Here, visual analytics can provide a more sophisticated approach.

2.2 Visual Analytics

Visual analytics focuses on deriving knowledge and gaining insight from huge and complex datasets (WONG & THOMAS 2004). In other words, it is about analytical reasoning supported by interactive visual interfaces (THOMAS & COOK 2005). Visual analytics is an interdisciplinary research area that combines several fields including visualization, data mining, statistics, databases, machine learning and human computer interaction (KEIM et al. 2010). Therefore, it profits from the knowledge and methods of these single fields.

The visual analytics process model (KEIM et al. 2010, KOHLHAMMER et al. 2011) is a fundamental concept that integrates automatic and visual analysis methods as well as human interaction. The model is illustrated in Figure 1. It describes several stages (data, models, visualization and knowledge) and the following stage transitions (KEIM et al. 2010):

- *Data transformation:* Pre-process and transform data
- *Data mining:* Apply (semi-)automatic analysis methods to discover useful patterns
- *Data mapping:* Apply methods for visual data exploration
- *Model building:* Guide model building with visual findings
- *Model visualization:* Visualize model for evaluation
- *Parameter refinement:* Evaluate and refine the model
- *Interaction:* Visually reveal insightful information
- *Feedback loop:* Store knowledge to support future analyses

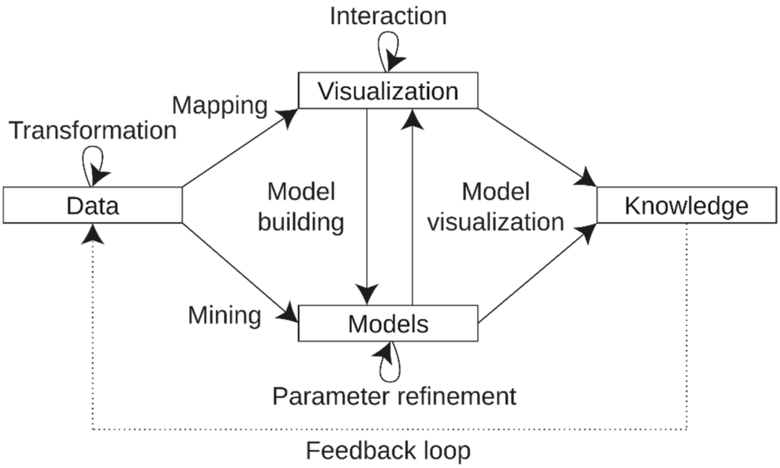

Fig. 1: Visual analytics process model

At first sight and starting with raw data, knowledge is obtainable either via automatic analysis or visualization. But this is just a limited view and not the whole truth. We emphasize that visual analytics is especially about the repeating combination of those approaches with a focus on interaction and further refinement, as well as the proper feedback or representation of constantly resulting adaptions of the involved models and visualizations.

The potential and idea of applying visual analytics to the e-government domain which is a superordinate field of e-participation has already been recognized. It has been stated that proper interactive visualizations can support governance and policy modeling including participation (KOHLHAMMER et al. 2010). Generally, visual analytics provides a huge range for the integration of sophisticated methods in participatory processes and spatial planning (e. g, cf. DE AMICIS 2009). But many applications only focus on information visualization and forget to heavily integrate intelligent semi-automated data analysis methods.

3 Case Study – Visual Analytics for Analyzing Contributions

In participatory processes, all involved parties try to accomplish several distinct tasks. The specific scope and especially the way of proceeding may differ because of different focuses and sizes of each single participatory process, e. g., spatial planning on regional level possibly involves more participants than spatial planning on local level that, in the end, might lead to a need for more sophisticated approaches and tools in order to deal with a wider range of complex interests and opinions. Also, there are differences between formal and informal participatory processes, e. g., in terms of legality, flexibility and liability. That enables different tools like voting and open or public discussions depending on the process type. We briefly show the applicability of the visual analytics process model by describing one possible use case in the sense of a practical scenario. In this case study, we focus on one major task category: the analysis of collected contributions. The following paragraphs describe the use case's s*ituation, goals* and *workflow*.

The use case's *situation* is about the analysis of incoming contributions for a planning and decision process in order to plan or develop a specific area of interest for a site design like a public green space. The process allows public participation. People can propose their ideas, interests and thoughts related to the future design of the area. Hence, we are in the beginning phase of the planning process, where no concrete plan has been developed yet, and participants are asked to provide ideas to guide the initial development of a plan. Contributions can be submitted online. An online GIS client provides a map of the area for exploration and positioning of each contribution. Therefore, each contribution comprises of written natural language text that expresses the proposal or idea for the area of interest, and geospatial information, i. e., a contribution is related to a spatial position. Exemplary contributions with notional data are listed in Table 1. In practice, there might be further information available like ratings or votes for or against favorite contributions, which are omitted here for the sake of simplicity. The list of contributions forms the raw dataset that, in the following, is processed by the visual analytics process.

Table 1: Contributions where each row represents one example with notional content

Id	Author	Content	XY-plane coordinates
1	John Doe	"The old tree definitely needs to stay. I like the shadows cast by it."	(341.1298, 123.8414)
2	Jana Doe	"Please add some benches at this location! Families and their children will use it."	(330.7743, 140.5791)
…	…	…	…

A planner wants to analyze the collected contributions, and thereby follows several goals. The development of a draft version of a planning concept and support in decision making between alternatives are the main focus. Based on the available data, this challenge involves many sub-tasks, e. g., the determination of a generalized overview of the proposed contents, the identification of similar or controversial contents, the examination of the most common proposals, or the retrieval of all participants and their contributions per aggregated idea. In other words, the implicit knowledge hidden in the data must be uncovered. Then, decisions like choosing the most prominent and approved ideas can be made.

The following *workflow* demonstrates how the described endeavor can be addressed by using a software application that integrates the visual analytics process model for data processing, exploration and analysis. The workflow's steps correspond to the stage transitions of the process model (the responsible transition is mentioned in brackets):

1) *Bin counting (data transformation):* The planner decides to get an overview and starts the evaluation by exploring "hot" regions of the provided map, i. e., subareas with a relatively high number of contributions. Therefore, the planner triggers the application-internal creation of a heatmap. Thus, a uniform grid is applied to the map and the number of contributions that spatially correspond to each cell is computed by counting. This leads to an internal representation (see left of Figure 2).

2) *Heatmap visualization (data mapping):* In the application, each cell's value is mapped to a color according to a color map, e. g., by using greyscale. This leads to a visual encoding of the "hot" regions (see right of Figure 2).

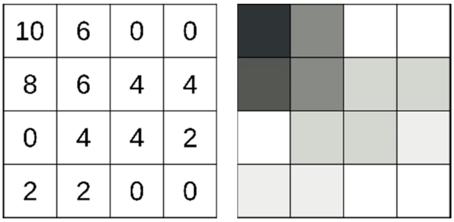

Fig. 2:
Internal or raw heatmap representation (left) and final visual representation (right) encoding the number of contributions per subarea

3) *Heatmap exploration (interaction):* On the one hand, contributions relating to the same spatial region are likely to relate to each other. Although they might be controversial, processing them together is a good idea. On the other hand, "hot" regions are a good starting point for processing the contributions as they are the main focus of interest from the contributor's point of view. Regarding these aspects, the planner gets support in exploring the visual heatmap representation by using panning and zooming interaction techniques to focus the attention on certain regions. For every cell in the viewport, zooming reveals the individual contributions at its pinned location. Let's assume the planner decides to investigate the "hottest" area with ten contributions. This yields the result in the left of Figure 3.

4) *Contribution selection (interaction):* For further investigation purposes, the planner selects a single contribution, #10 in the right of Figure 3, to see its actual content. Once the content has been analyzed by the planner, he wants to continue with this contribution. As noted earlier, analyzing related contributions one after the other makes the analyzing task easier for the planner. As spatial closeness is only one indication for similar contributions, the system can provide a ranking of contributions in order of decreasing similarity to the selected one, taking into account several criteria like geographic closeness or text similarities.

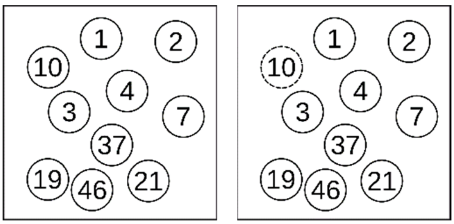

Fig. 3:
Revealed contributions of one enlarged cell (left) and selection of a single contribution (#10) (right)

5) *Ranking (model building):* Hence, the selected contribution acts as a query for building a ranking model similar to search engine rankings (see Figure 4). The goal is to retrieve all contributions of the current viewport and sort them by their similarity to the queried contribution.

```
Model = {
  "name": "ranking",
  "query_contribution_id": 10,
  "result_list": [
    {"contribution_id": 37, "score": 0.91},
    {"contribution_id": 2, "score": 0.74},
    ...,
    {"contribution_id": 46, "score": 0.0}
  ]
}
```

Fig. 4:
Excerpt of a simple abstract ranking model based on text similarity using computed scores ranging from 0.0 (not similar) to 1.0 (similar)

6) *Score encoding (model visualization):* In order to be able to examine the ranking model, the ranking needs to be visualized. Again, an option is to map the achieved score of each contribution to a color and apply this to the map visualization. Simultaneously or alternatively, the ranking order can be visualized as a descending list like internet search engines do (see Figure 5).

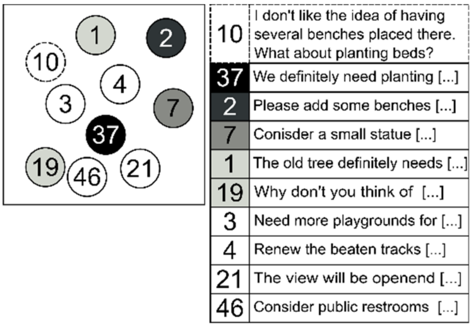

Fig. 5:
Ranking model visualization incorporating a visual (left, right) and positional (right) encoding of the score related to the selected contribution (#10).

7) *Knowledge processing:* Based on the visualized ranking model, many different possibilities for further actions arise while details depend on specific needs and concrete software implementations. In the case of aggregating ideas to build a concept plan as described in the use case's goals, one important step is the aggregation of contributions. Near duplicates or similar ideas are first candidates to be clustered in one idea. Controversial contributions might be guiding the development of different alternative concept plans that can be later presented to be voted on. Therefore, the system shall support the iterative reduction of contributions left to analyze as well as their direct transfer to concept development. In this regard, the planner could select and group contributions that are ranked near to each other. At the same time, the planner might add written notes to the grouping for future work or reporting. From the planner's point of view, this can also change the individual processing state for the contribution, e. g., from "unseen" to "merged".

4 Conclusion and Future Work

Visual analytics tightly couples semi-automated data analysis and visualization in an interactive fashion. It is beneficial for ICT-supported participatory processes, because it integrates individual complex data analysis and exploration tasks into one workflow or session. This leads to new interactive possibilities for knowledge discovery and decision making which is crucial in the e-participation domain. But the reviewed model is a general framework. Its merit depends on specific realizations of data stages and transitions. In this regard, the possible spectrum of definite realizations is too large for a general evaluation.

It is still necessary to evaluate specific visual analytics methods for concrete analysis tasks in the e-participation domain in order to proof the applicability and their mentioned benefits. We are currently implementing visual analytics methods in a Web-based application. In slightly more detail, this part concerns the ranking, clustering and visual exploration of participatory contributions in order to analyze similar or contrary opinions or ideas.

This work has been funded by the German Federal Ministry of Education and Research (BMBF), funding line FHprofUnt, grant identifier 03FH011PX4. The responsibility for the content of this publication rests with the authors.

References

De Amicis, R., Stojanovic, R. & Conti, G. (Eds.) (2009), GeoSpatial Visual Analytics. Geographical Information Processing and Visual Analytics for Environmental Security.

EIPP – European Institute for Public Participation (2009), Public Participation in Europe. An international perspective. www.partizipation.at/fileadmin/media_data/Downloads/Zukunftsdiskurse-Studien/pp_in_e_report_03_06.pdf, last access: 16-09-30.

Helbig, D., Pietsch, M., Schütz, L., Bade, K., Richter, A. & Nürnberger, A. (2016), Online-Beteiligung in Entscheidungs- und Planungsprozessen – Anforderungen aus der Praxis. AGIT – Journal für Angewandte Geoinformatik, 2-2016, 508-517.

Keim, D. A., Kohlhammer, J., Ellis, G. & Mansmann, F. (Eds.) (2010), Mastering the Information Age – Solving Problems with Visual Analytics. Eurographics Association.

Kohlhammer, J., Ruppert, T., Mansmann, F. & Keim, D. A. (2010), Information Visualisation and Visual Analytics for Governance and Policy Modelling. In: 2nd CROSSROAD Workshop Gap Analyses & Roadmap Validation: Side event of the International High-Level Symposium on ICT for Governance and Policy Modelling.

Kohlhammer, J., Keim, D. A., Pohl, M., Santucci, G. & Andrienko, G. (2011), Solving Problems with Visual Analytics. Procedia Computer Science, 7, 117-120.

Richter, I., Danelzik, M., Molinego, G., Nanz, P. & Rost, D. (2016), Bürgerbeteiligung in der Energiewende – Zehn Thesen zur gegenwärtigen Etablierung, zu Herausforderungen und geeigneten Gestaltungsansätzen, IASS Working Paper, Februar 2016.

Thomas, J. J. & Cook, K. A. (Eds.) (2005), Illuminating the Path: The Research and Development Agenda for Visual Analytics. IEEE Computer Society Press.

Wong, P. C. & Thomas, J. J. (2004), Visual Analytics. IEEE Computer Graphics and Application, 24 (5), 20-21.