DELIVERABLE 7.2: EVALUATION OF VISUAL ANALYTICS PROTOTYPES (VERSION 1)

ABSTRACT
This document describes the evaluation of the first version visual analytics prototype implemented within Task 7.3 of Work Package 7 “Visual Analytics Techniques for Supporting the Policy Making Process and eParticipation”. The evaluation was conducted within Task 7.4 “Visual Analytics Prototype Evaluation” of Work Package 4.
This page has been intentionally left blank.
1 Introduction

This document describes Deliverable 7.2 “Evaluation of Visual Analytics Prototype (Version 1)”. The evaluation was conducted within Task 7.4 “Visual analytics prototype evaluation” of WP7 “Visual analytics techniques for supporting the policy making process and e-participation”.

Goal of the visualization component is to provide an intuitive access to the technical components of the ePolicy project. The three visual interfaces - one for the global optimizer, one for the social simulator, and one for the opinion mining component - will be evaluated regarding their usefulness and their usability. These evaluation criteria can be divided into functional (does the visual interface fulfill the requirements of the user?) and non-functional (is the interaction design and the visual encodings usable for the user?) requirements respectively.

1.1 Functional Requirements

The functional requirements of the visualization component are directly connected to the functional requirements of the three mentioned technical components. Mainly, the visual interfaces have to provide access to these technical components in two ways. Firstly, the users should be enabled to visually interactively define input parameters for each component. Secondly, the users should be enabled to view the output data of the technical components and analyze this data. In order to evaluate these criteria first of all a requirement analysis for the visual interfaces is conducted (cf. Deliverable D7.1 “Prototypes of visual analytics techniques V1”). This is realized via a user questionnaire. Based on these requirements a task completion test is designed. In these tests the users had to solve specified tasks with the implemented visual interfaces that cover the functional requirements on the visual interfaces. If these tasks are completed successfully with an average rate of over 70%, the functional requirements are met.

1.2 Non-Functional Requirements

Non-functional requirements on the visualization component concern the visual encodings and the interaction designs. In order to test these features, a questionnaire for each visual interface is designed. In these questionnaires the users are asked whether they understand the visual designs and whether they intuitively know how to use the interfaces. Moreover, the users are enabled to give informal feedback on the designs regarding a possible improvement of the visualization component. If more than an average rate of over 70% approve the usability of the visualization component regarding its visual encodings and interaction design the non-functional requirements of the visualization component are met.

2 Evaluation Methodology

We conducted a case study to evaluate the visualization prototype delivered in D7.1. This prototype assures the visual-interactive access to the global optimizer developed in WP3. In the following, we describe the evaluation methodology.
2.1 Experimental Design

The system evaluation was hampered by the fact that, to the best of our knowledge, no comparable visual interfaces for Policy planning and Strategic Environment Assessment (SEA) could serve as a baseline to evaluate against. Neither in the targeted policy analysis domain, nor in the optimization domain computational models for solving multidimensional optimization problems have been presented in a comparable manner. For this reason, the experiment was designed towards an evaluation of the targeted purpose of this approach. We considered a) the validation of the functional user requirements and b) the verification of the visual encodings of the system as the two important factors to prove the success of the system.

We conducted an informal summative case study based on task-completion tests and a user questionnaire. The field-based case study was conducted with our web-based application to enable a real-world setup. The unsupervised environment enabled the participants to perform the tests without being influenced, in order to give credible feedback. Model parameterizations for the respective tasks were carefully selected and tested in a previously applied test-run to control the resulting data values and their visual representations. The dependent variable of the case study was the task-completion. Additionally, qualitative feedback based on user preference was gathered within the user questionnaire.

2.1.1 Task-Completion Test

We designed the tasks to validate the functional user requirements (cf. Section 3 in Deliverable D7.1 “Prototype of Visual Analytics Techniques (Version 1)”). The requirements were derived from a requirement analysis described in Section 3 of Deliverable D7.1 “Prototype of Visual Analytics Techniques (Version 1)”. They can be summarized to the following:

Functional:

1) Visual definition of target function and constraints (FR1)
2) Visualization of calculated optimized energy plan (FR2)
3) Comparison of energy plans (FR3)
4) Consideration of environmental, economic and social impacts (FR4)

Technical:

5) Provide application as web-application (TR1)
6) Consider common visual and interaction styles of web-applications (TR2)
7) Connect service to optimization module of WP3 (TR3)

The tasks of the task-completion test were chosen by means of covering each functional component. By defining the task setup, attention was paid to ensure that each task covers an ‘atomic’ functional unit of the analysis workflow. This enabled the identification of bottlenecks in the analytical workflow and a target-oriented enhancement thereof. We were interested in what way non-expert users would be able to comprehend the analytical context of the domain-specific model. In this way, statements regarding the ‘generalizability’ of the case study became possible.
First of all, the user had to calculate an energy plan with the default input parameters. Then, she had to solve task 1 and task 2 with the Optimized Plan View (FR2). The first task should test, whether she is able to find the relevant information displayed. The second task should test whether the user discovered the additional information in the heatmap by using the mouse-over tooltip. Next, the user had to apply the detailed impact view to solve task 3 (FR4). Task 4 could be solved by understanding the remaining visualization of the Optimized Plan View (FR2). Afterwards, the user had to specify an alternative plan in the Input Interface by changing some of the default constraints (FR1). The two calculated plans were compared by the user in the Compare View to solve tasks 5 and 6 (FR3). The six tasks to cover the functional requirements were:

1.) Which of the energy sources in the energy plan produces the lowest amount of energy?
2.) On which environmental receptor the most negative impact is induced? What is its value?
3.) Which activity has the greatest negative impact on “Quality of sensitive landscapes”?
4.) Which energy source needs the highest amount of “Aerial power line supports”?
5.) Which energy plan has a more positive impact on the air quality?
6.) Which of the two plans produces more energy?

2.1.2 Questionnaire
In addition to the task-completion tests, we provided a questionnaire that enabled the participants to give qualitative feedback. It targeted the usability of the system. Questions of concern were the verification of the visual encodings and interactive capabilities. First, questions about the participants’ profession, their domain of expertise and their common analysis tasks, were asked. Additionally, the questionnaire incorporated eleven closed questions about the visual encodings and the usefulness of each view. Finally, we provided open questions to gather informal feedback of individual user preferences.

2.1.3 Participants
Ten non-domain experts agreed to participate voluntarily in the task-completion test. All of them had a profound background in information visualization. None of them had expertise in the energy domain. Thus, the results of the task-completion test were not influenced by domain knowledge which allows the generalization of the results. The questionnaire was answered by all non-experts, and two additionally recruited experts from the energy domain with no expertise in visualization. Most of the participants were male. None of them reported color blindness. The age of the participants reached from 22 years to 38 years with a median of 28 years. All of the participants had an academic degree.

3 Task-Completion Test Results
Figure 1 shows the results of the task-completion tests. Tasks one, three and four were completed correctly by most of the participants while tasks two, five and six had higher error rates. The tasks with the low error rates were completed by using the Input Interface, the Optimized Plan View and the Impacts View. For the tasks with the higher error rates the users had to make use of an additional Overview and the Compare View. The Overview consisted of a scatterplot visualizing
all computed plans with respect to their overall energy produced and overall costs. In the Overview the user had to select a set of plans for comparison.

The usability test shows that on average 90% of the first four tasks were completed correctly but only 65% of the last two tasks. This concludes that the users were able to use the Input Interface and calculate optimized energy plans. Thus requirement FR1 is fulfilled. The Optimized Plan View supports the analysis of energy plans and the Impacts View displays the environmental impacts. The tasks 1-4 related to these views were completed correctly by most of the users. Thus, requirements FR2 and FR4 are met. The results of the last two tasks indicated that the exploration of multiple energy plans and the following comparison is not as easy as the previous tasks. To solve the task the Overview and the Compare View had to be used. The error rates of the last two tasks showed that the visual-interactive design of these views did not support the user as well as the other views. Thus the requirements FR3 and FR4 are met but the solutions have to be improved.

![Figure 1: Task-completion test results. On the x-axis the six questions are categorized and the y-axis shows the percentage of correct, wrong or missing answers.](image)

4 Questionnaire Results

The questionnaire was designed to test the usability of the system with respect to the visual encodings and interactive capabilities. For each view the users were asked regarding its intuitiveness (see Question 1 for each view) and usefulness (see Question 2 for each view). For the Impacts View, we additionally asked the participants whether the information should be grouped to provide a better overview (see Question 3 in Impacts View). In the following, we present the results of the questionnaire.
4.1 Input Interface

The Input Interface (see Figure 2) was designed to enable the visual access to all possible degrees of freedom of the optimization module.

![Input Interface](image)

**Figure 2: Input Interface: Definition of target function (a), and constraints (b) and labeling of the energy plan (c)**

The results presented in Figure 3 show that most of the users were satisfied with the Input Interface. Nevertheless, it was mentioned that the constraints should be organized into groups to get a better overview (e.g. budget constraints, constraints on energy sources, constraints on environmental impacts). Other users stated that the interface has too much information which should be explained.

![Questionnaire results for the Input Interface](image)

**Figure 3: Questionnaire results for the Input Interface**
4.2 Optimized Plan View

The Optimized Plan View shown in Figure 4 presents the calculated energy plan.

![Optimized Plan View](image)

**Figure 4: Optimized Plan View**

(a) Overview of costs and produced energy (b) Normalized distribution of secondary activities (c) Absolute distribution for one selected secondary activity (d) Impacts on the environment.

The evaluation results for the Optimized Plan View (see Figure 5) are similar to the Input Interface. Most of the users were satisfied with the design of this view but stated that it took a while to get used to it. The headlines of the visualization are not precise and 50% of the users said that this view offers too much information. It was also suggested to allow customization of the visualization especially for the colors so that users are able to highlight energy sources themselves.

![Questionnaire results](image)

**Figure 5: Questionnaire results for the Optimized Plan View**
4.3 Impacts View

The Impacts View presented in Figure 6 gives a deeper look into the relations between activities and the receptors.

![Impacts View](image)

**Figure 6: Impacts View: Activities (rows) and receptors (columns) sorted by name.**

The evaluation results (see Figure 7) show that the Impacts View is easy-to-understand for 92% of the users but only 75% think that it is useful. Mainly this is because of the small labeled columns and rows. The usefulness of the sorting function was questioned and the sorting speed is too slow. A further suggestion was to remove unimportant rows with only white cells from the visualization. Finally all users agreed that the activities should be grouped by energy source to get a better overview.

![Questionnaire results for the Impacts View](image)

**Figure 7: Questionnaire results for the Impacts View**
4.4 Overview

The Overview (see Figure 8) displays all calculated plans to recognize main differences of the plans.

Figure 8: Overview: Left: Scatter plot with six precomputed energy plans. The energy plan represented by the dot at the top is selected. Right: The glyph which shows all information that defines the selected plan.

The least successful view in the questionnaire is the Overview (see questionnaire results in Figure 9). Only 50% of the users intuitively understood the visualization and considered it useful. A few users were not able to recognize the selection functionality which is a crucial criterion. Some users mentioned that the dots in the scatter plot should be labeled for a better recognition. The glyph in the sidebar may be too cryptic and the relation between the dot and the glyph is not easy-to-understand.

Figure 9: Questionnaire results for the Overview
4.5 Compare View

The Compare View (see Figure 10) has the target to compare a smaller set of energy plans that the user is interested in.

![Compare View](image)

**Figure 10**: Compare View: (a) Comparison of three energy plans in total produced energy and overall cost. (b) Energy and cost split up on the different energy sources used by the energy plan. (c) Comparison of environmental impacts caused by the energy plan.

The Compare View had the greatest approval with 92% for being intuitive and 100% usefulness (see Figure 11 for the results of the questionnaire). The only suggestion was to add the sum of all impacts on the environment caused by one energy plan.

![Questionnaire results](image)

**Figure 11**: Questionnaire results for the Compare View
4.6 Summary of Questionnaire Results

The questionnaire was designed to test the usability of the system with respect to the visual encodings and interactive capabilities. For each view the users were asked regarding its intuitiveness and usefulness. For all views except for the Overview above 80% of the users stated that they are easy to understand. Above 70% of the users argued that the views are useful, the Compare View even gained 100% of the users approval. The Overview was only intuitively understood by 50% of the users. Its usefulness was only affirmed by half of the users. Hence, we assume that the lower task-completion rates of tasks 5 and 6 are caused by the inappropriate design of the Overview, not the Compare View. As a result the Overview was excluded from the application after the test. Calculated energy plans are now directly added to the Compare View.

4.7 Further User Suggestions

It was highly recommended by multiple participants to integrate the browser functionalities “Back” and “Forward”. Further the domain experts suggested providing a download of calculated solutions to use them for further work. Another important recommendation was to display strengths and weaknesses of alternative solutions. Many of the participants suggested to explain the information and interaction with a “FAQ” or “How-to” page.

Finally, we want to highlight the facts that differed between the domain experts and the other participants. All in all, the impressions on the Overview are similar in both groups. The greatest difference is that the domain experts wanted more information in the views unlike the other participants which were overwhelmed by the amount of information. Since policy makers and public stakeholders are new to the process of energy plan making the same effect could occur if they use the application. Thus it is a challenge to combine these two opinions in one application.

5 Informal Evaluation

The goal of the informal evaluations is to approve early design decisions. During the development phase implementations were demonstrated to domain experts to determine if the direction of the development is correct. These evaluations had no prepared tasks or questions. The results were only communicated verbally. Multiple iterations were presented in that way and the feedback was incorporated. During these informal evaluations three benefits emerged.

1) The visualization of the optimized energy plans offered a better overview of the data generated by the optimization model. Thus, it was possible to evaluate the optimization model with analyzing the visualizations. We detected some irregular values and communicated them to the modeling experts from WP3. Because of that they improved the optimization model.

2) Due to the rapid prototyping of our application we showed early implementation to the project partners. They examined our concept and identified which visualizations had to be improved. These formative evaluation results improved our development and prevented bad design decisions.
3) Discussions with the modeling experts showed that the Input Interface offers an easy way to access the optimization model. Thus we decided to make more input parameters of the optimization model accessible.

6 Conclusion
In this document we have described deliverable D7.2 “Evaluation of Visual Analytics Prototype (Version 1)”. It consists of the evaluation of the visual interface to the Global Optimizer component developed in WP3. First, we described our evaluation methodology including a verification and a validation approach. The visualization prototype delivered in D7.1 was tested against its usefulness and its usability. We conducted task completion tests and questionnaires with both non-expert participants and domain experts. The usefulness of the system was proven and the usability of most of the presented visualization designs were considered positive. The remaining negative issues and comments by the evaluation participants will be targeted in the next implementation cycle of WP7 (cf. Task D7.3 in DoW).

The visualization system and its evaluation has been submitted for publication to the International Symposium on Visual Computing (ISVC 2013). We are proud to announce that the paper has been accepted and will be presented this year at the conference.

As next steps, the second implementation cycle will be started (see D7.3). Visual interfaces for the social simulation component and the opinion mining component will be designed and implemented. Furthermore, the existing presented visual interface to the Global Optimizer component will be refined, based on further upcoming requirements and the results of the evaluation.

---

7 Appendix – Task-Completion Test and Questionnaire

ePolicy – Engineering the policy making life-cycle

Questionnaire

Intro

The ePolicy project is a joint research project funded by the European commission with the goal to support policy makers in their decision process. One of the use cases in the project focuses on optimizing the regional energy plan of the region Emilia Romagna in Italy.

At the following Link (URL: http://pcgris-tib.igd.fraunhofer.de/epo/) you will find a first version of the visualization prototype.

The underlying questionnaire will be used to detect user requirements for the visual interface, and to evaluate the first version of the visualization prototype. Please fill in the first part of the questionnaire (“Requirement Analysis”), before exploring the visualization prototype. Please do not resize the window when using the prototype.

General Questions

Personal Information

Name (optional): ____________________________________________________________

Profession: ______________________________________________________________

Degree of education: ______________________________________________________

Domain of expertise: ______________________________________________________

Analysis tasks: ___________________________________________________________
Usability Test

Please follow the next steps and fill in the empty spaces with the answers.

Firstly, start the prototype under the following link:
...

Now, please choose “Global Interaction Optimizer” and then “Start”.

In this test we start with no pre-computed energy plans. So, please choose “Calculate Energy Plan”. In this view you can see the input interface. The constraints are predefined to match the regional requirements for the region Emilia Romagna in Italy and the target of the optimization is to maximize the produced energy. We proceed with calculating an optimal energy plan with the given constraints by clicking on “Compute”. Now you can see the details of the computed energy plan. In the top left the costs and produced energy is shown. On the right impact on the environment can be seen and the chart on the bottom shows a distribution of activities that have to be done to install the energy plan.

**Which of the energy sources in the energy plan produces the lowest amount of energy?**

**On which environmental receptor the most negative impact is induced? What is its value?**

Now have a detailed look at the impacts by clicking on “Detailed Impacts”. Every row shows an activity that has to be done when installing the energy plan. Every column stands for an environmental receptor that is influenced by these activities.

**Which activity has the greatest negative impact on “Quality of sensitive landscapes”?**

In order to see which energy source needs this activity most we go back to the energy plan view by clicking on “Optimized Plan” in the navigation panel on top of the page.

**Which energy source needs the highest amount of “Aerial power line supports”?**

Now, please calculate an alternative energy plan by clicking on “Specify Plan”. Please edit the constraints so that the energy from “Thermoelectric Biomass Plants” is not allowed in the energy plan. Then, compute the optimal energy plan.

Now, you can compare the two calculated plans by selecting the “Overview” button. You can see the energy plans and by clicking on the points you get information about the plan in the sidebar. You can see the constraints, and the calculated optimal values of the plan.
Select both plans by using the selection box and click “Compare Selection”.

Which energy plan has a more positive impact on the air quality?

Which of the two plans produces more energy?

That concludes the test, thank you. Feel free to use the application again, and please answer the following questionnaire.
**Questionnaire**

Specify Plan

In the view ‘Specify Plan’ the user can specify the target function and the constraints.

![Specify Plan: visual interface to define objective function and constraints.](image)

Looking at the visual interface – do you intuitively know, how to use the interface?

- [ ] Yes
- [ ] No, why not: _________________________________________________________________

Do you find the organization and display of the interface useful for your work?

- [ ] Yes
- [ ] No

Comment: _________________________________________________________________________

Is any information missing in the interface?

__________________________________________________________________________________

Is any information provided that should be neglected in the interface?

__________________________________________________________________________________

Do you have any suggestions for improving the interface?

__________________________________________________________________________________
Optimized Plan

After the definition of target function and constraints the user can click ‘Compute’ to achieve an optimized plan calculated by the system. The results are visualized in a dashboard.

Looking at the visual interface – do you intuitively know, how to use the interface?

☐ Yes
☐ No, why not: _________________________________________________________________

Do you find the organization and display of the interface useful for your work?

☐ Yes
☐ No

Comment: _________________________________________________________________________

Is any information missing in the interface?

__________________________________________________________________________________

Is any information provided that should be neglected in the interface?

__________________________________________________________________________________

Do you have any suggestions for improving the interface?

__________________________________________________________________________________

__________________________________________________________________________________
Impacts on Environment

Besides analyzing the optimized solution for the energy plan the impact of this plan on the environment can be observed by clicking on ‘Detailed Impacts’ (see Figure 2d)).

Figure 3: Detailed Impact: detailed visualization of the impacts of activities (rows) on the environmental receptors (columns).

Looking at the visual interface – do you intuitively know, what is shown in the interface?

☐ Yes
☐ No, why not: _________________________________________________________________

The visualization shows positive or negative impacts. Does this visualization convey information which is useful for your work?

☐ Yes
☐ No

Comment: _________________________________________________________________________

Do you have any suggestions for improving the interface?

__________________________________________________________________________________

Which of the provided information is important for you, which can be neglected?

__________________________________________________________________________________

Should the impacts be grouped with respect to the energy source?

☐ Yes
☐ No, why not: ____________________________________________________________________
Overview

In the overview all plans already specified are visualized in a scatter plot. By clicking on one point, an overview of the plan is provided (right side of view). Several plans can be selected (by specifying rectangular), and compared (click on ‘Compare Plans’) in an additional view (see Figure 5).

Figure 4: Overview of energy plans: Visualization of plans in scatterplot showing costs (x-axis) and produces energy (y-axis). Right: compact visualization of plan.

Looking at the visual interface – do you intuitively know, how to use the interface?

☐ Yes

☐ No, why not: _________________________________________________________________

Do you find the organization and display of the interface useful for your work?

☐ Yes

☐ No

Comment: _________________________________________________________________________

Do you have any suggestions for improving the interface?

_______________________________________________________________________________
Comparison of Plans

Figure 5: Comparison of plans

Looking at the visual interface – do you intuitively know, what is shown in the interface?

☐ Yes
☐ No, why not: _________________________________________________________________

The visualization shows three selected plans. Do you find the organization and display of the interface useful for your work?

☐ Yes
☐ No

Comment: _________________________________________________________________________

Which of the provided information is important for you, which can be neglected?

________________________________________________________________________________

Do you have any suggestions for improving the interface?

________________________________________________________________________________
What is your overall impression to the system? Is it useful for your work?

__________________________________________________________________________________

Do you have any open questions, ideas, suggestions, ...?

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

Thank you very much for your support!

The ePolicy Project Team