Interactive analysis of the Viennese media transparency network

Paul Binder
University of Vienna

Ulrich Schweinitzer
University of Vienna

Philipp Dangl
University of Vienna

ABSTRACT

Interested and concerned citizens want to investigate money flows of governmental organization to media organizations. This is open data for everyone because of media transparency. This dataset will give you the opportunity to access the data in an interactive and user friendly way to explore and gain insight into the money flows of these organizations. The visualization is based on four different views with each showing a different aspect of the data.

Index Terms: information visualization, interactive data exploration, user-centred design, viennese media transparency, Tableau, D3

1 INTRODUCTION

This paper is about a design study of the Viennese media transparency network. Which means it shows the different money flows of governmental- to media organisations. When we started this project we wanted to make the government data, which is openly accessible, easier to understand and more interactive. We wanted to give the user the option to explore the dataset better and make it more user friendly than just a simple excel sheet.
This is why we decided to make this particular visualization to address the user needs. Furthermore, this whole visualization is for anyone who is political interested or just concerned about the money flows of governmental- to media organizations.

2 RELATED WORK

Basically there is not much Related Work for this particular kind of visualization. We also did not use any kind of these to build our own visualization.
To get the source data for our visualization visit:
https://www.rtr.at/

- This particular visualization is from students of the TU Wien and has some similar views to our own visualization. For example the bundled-edge graph and also the quarters used in a graph although it is differently utilized in our graph. They also addressed pretty much the same task we do in our visualization.
  Dashboard
- There is also another paper from the "Forschungsforum der Oesterreichischen Fachhochschulen" which also addresses the same problem with handling the data and making it more accessible to the user.
  Paper

3 APPROACH

3.1 Description of our design

The first thing we did, when we started with this project is that we put the data into Tableau and explored it a little, after that we each made a unique paper prototype which you can see in Figure 1.
Based on our three paper prototypes we agreed on the graphs our final Dashboard should contain. We discussed the different types

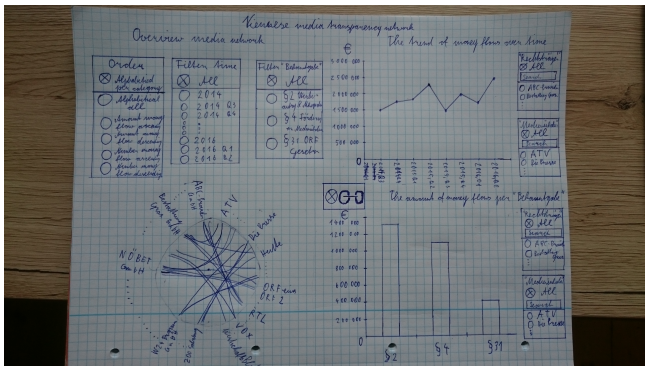


Figure 1: A simple paper prototype we created during Milestone 2.

of graphs each prototype had and listed all the advantages and disadvantages of the distinct views.
We also created a simple use case scenario which we later broadened because it did not make much sense at the beginning.

In Milestone 3 we reiterated on the use case scenario of Milestone 2 to make sure our graphs contain all functionality needed to achieve the purpose of the use case scenarios. For Milestone 4 we tried to integrate Tableau and D3 into a single dashboard and tried to make it as interactive as possible.

3.2 Reasons for design choices

We used 4 different charts and here we list a few reasons of our distinct graphs.
Bundled Edge Graph:
The bundled edge graph gives a good detailed view after selecting parameters. It displays the moneyflow between multiple entities better than any other chart and also lists every single transaction between them over the selected quarters. There are many different filter possibilities such as sort per category and alphabet, time, specific Rechträger or Medieninhaber" and money range. Another advantage is the highly reduced edge crossing.

Packed Bubble Chart:
The graph provides a good first impression of the highest transactions of the selected data and it is good for an average amount of categories in the same value range. If you select to specific Medieninhaber all the transcation are nicely distinguished between those. It also looks really fancy and colorful.

Tree map:
Easy to determine the high and low values going from top left to bottom right A major advantage is that it doesn't matter how many categories or values there are, they are always displayed on the same page.

Line chart: This graph provides a really good view to see the trends over the quarters for the moneyflows. Because of this it enables the user to make predictions based on the trends.

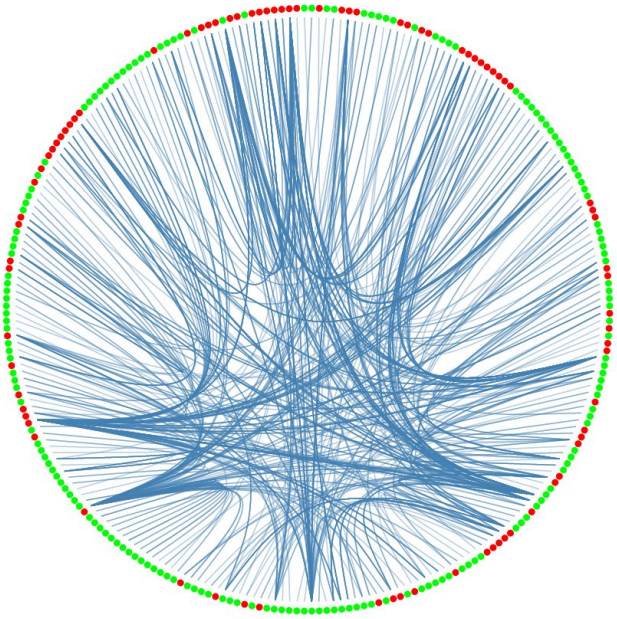


Figure 2: The node view of the bundled edge graph.

4 Implementation

4.1 Implementation Methods

Languages: Javascript, HTML, CSS
Toolkits and libraries: Tableau, D3, JQuery (jquery-3.1.1.js), noUiSlider (nouislider.js, nouislider.css), Searchable Option List (sol.js, sol.css), wNumb (wNumb.js)

We implemented three of the four graphs with Tableau and the bundled edge graph with D3 and the other toolkits. We used noUISlider for the money range slider and the Searchable Option List was necessary for the functions and the display of the Choose a Rechträger / Medieninhaber filters. Moreover, we formatted the numbers with the wNumb library. Furthermore, we used some css options to improve the representation of the bundled edge graph and the filters. We connected all these libraries and toolkits in one html file due to the length of the code and because there was only one graph and to make the processing of work easier.

4.2 Implementation Challenges

We implemented the representation of the nodes in the bundled edge graph as a textual description if less or equal than 250 nodes are displayed and color separated circles without description above 250 nodes was quite a challenge see Figure2. Additional information is shown in a tooltip in both cases.
One major problem we noticed while creating the bundled edge graph was that transactions in our data set are all one directional from a particular Rechtsträger to a Medieninhaber, which leads to an unusual form of bundling.
In addition, the D3 bundle functions limited our choice of order of the nodes and the sorting functions. Another problem that the sorting of the nodes is currently limited to a category separated order which also adds to the unusual form of bundling.
Another problem was that html has no approach to realize a range slider with two inputs and it also has no good representation of multiple values inside drop down lists. So we had to use some libraries and toolkits for these cases.
Moreover, the representation of a high amount of nodes in circular form was not the best solution but we found no better solution for the bundled edge graph which would be easy to implement.
Another challenge was to copy all the CSVs to one CSV and reduce the loops to get a better performance and to avoid calculation overhead due to multiple invocations. We also had to filter the CSV to the correct values and format them into an array that fits the format we needed for the bundling process.
In the high fidelity prototype we used for a better distinguishability blue as default color and red for all Rechträger and green for all Medieninhaber nodes and links. We decided to use red for the Rechtsträger and green for the Medieninhaber.
We checked for colorblindness on this site Vischeck.com and it looked fine.

The Integration of both, Tableau charts and D3 graph into one final interactive dashboard, was one of our finishing tasks for milestone 4. Examples or related work on integrating Tableau charts into D3 or vice versa integrating a D3 graph into a Tableau dashboard, are very limited to non-existent. The only possible way we found was to integrate our D3 graph, uploaded to our team projects webspace, as a webpage element into our Tableau Dashboard and then upload and publish it on Tableau Public. The major drawback of this solution is that we had to use two sets of filters on the final dashboard, since cross-filters werent possible with the D3 graph as a webpage element. This took away some of the valuable space on the dashboard, which we could have used to enlarge the D3 graph or one of the Tableau charts. Another problem was the scaling of the D3 graph inside the Tableau dashboard. We could have adapted the sizes of the Filter elements and the bundled edge graph to only fit inside the webpage element in the dashboard but we tried to make it scale dynamically to the width and height of the webpage so you can still work with the bundled edge graph alone in a browser. Moreover, Tableau was kind of limiting the possibilities of these graphs, for example we wanted to use a specific bar-chart view, but we could not realise it the way we wanted in Tableau to display the data correctly and useful for the user. The problem was it could not properly display multiple Rechträger to multiple Medieninhaber. So we decided to replace it with a bubble chart.
Later on we were thinking about replacing the bubble chart with a different diagram but the limitations of the data set itself restricted the possibilities of making distinct views. Because of this we decided to keep it for the final submission.

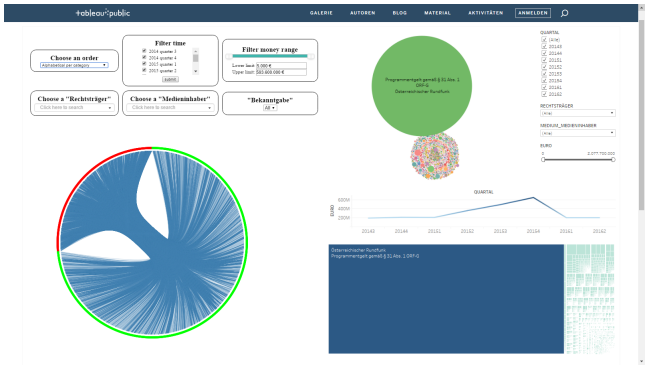


Figure 3: Problem of default view.

5 Results

5.1 Scenarios of use

Here are three possible use cases on which we decided in our Milestone 3:

Use Case 1:
Description:
An Employee of the Heute newspaper wants to compare the received amount of money flow of the 2nd quarter of 2015 with the competitor Kronen Zeitung in the same quarter.
Execution:
First deselect ALL. After that the Employee has to select both Heute and Kronen Zeitung in the filter of Medieninhaber. He has to leave Rechtsträgerand amount of money on default and change the filter of the year to only 2015 2nd quarter.

Use Case 2:
Description:
A student has to analyze the different transactions of the universities of Austria in the year 2015 for a student project.
Execution:
First deselect ALL. After that the student has to select the filter for the time, in this case he selects all 4 quarters of 2015, after that he selects all universities in Austria and leaves the filters for Medieninhaber and amount of money on default.

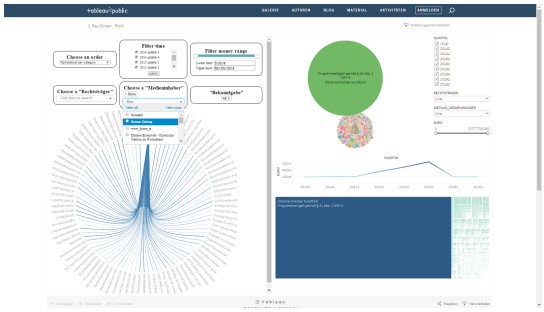
Use Case 3:
Description:
An Employee of the Bundeskanzleramt wants to see where the most amount of money went to in the 2nd quarter of 2016, as well as see all the separate transactions including the amount in euro.
Execution:
First deselect ALL. After that Employee has to select Bundeskanzleramt as Rechtsträger, the 2nd quarter of 2016 in the time filter and leave the Medieninhaber and amount in euro as default.

5.2 Performance

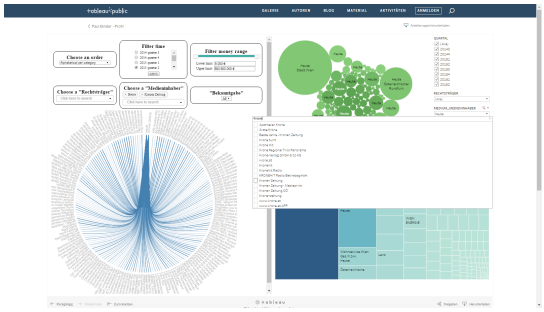
For the given time period of the project, we are rather satisfied with the performance of the final dashboard. With the provided amount of datasets we had for this project, the Tableau charts react very smoothly and fast while loading the full data overview and reloading after selecting specific filters. The D3 graph has a slight delay loading the overview with the full data, coming with the higher complexity compared to the Tableau charts, but also loads fast after selecting filters for the graph and therefore reducing the amount of data. Our dashboard, published on Tableau Public, is optimized for desktop pcs and laptops with moderate screen resolution, but not for smartphones and smaller tablets. For even better performance, a single set of filters in the dashboard applying the selected options to the Tableau charts as well as the D3 graph would be optimal, but as mentioned in 4.2 we didnt find a solution for that.

5.3 Feedback

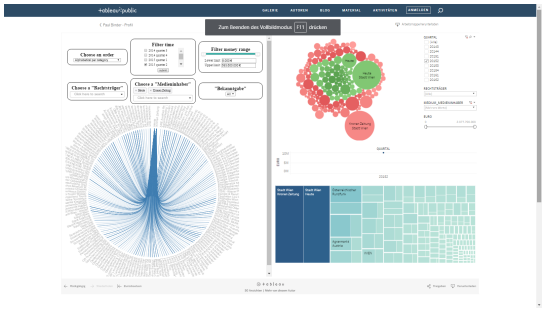
We read the feedback on our submissions of the previous milestones and tried to implement most of the given suggestions.
First of all, we tried to include all important information for our project in this final written paper, especially going into detail in the approach and implementation. We also describe the different scenarios of use of milestone 3 in more detail, including screenshots of single steps and the expected results of the scenario.
Furthermore, we gave a better reasoning for our design choices, specifically the different chart types we used and why we used them. It is possible that there would be a better alternative to the D3 graph for the data of this project but we decided to keep it in the dashboard because it provides detailed insight and interactive exploration of the data.



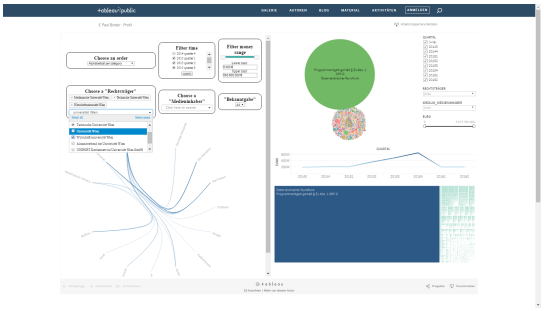
(a) Start of the Use Case 1



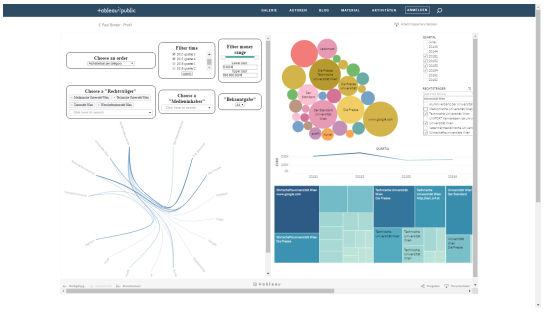
(b) Next step of the Use Case 1



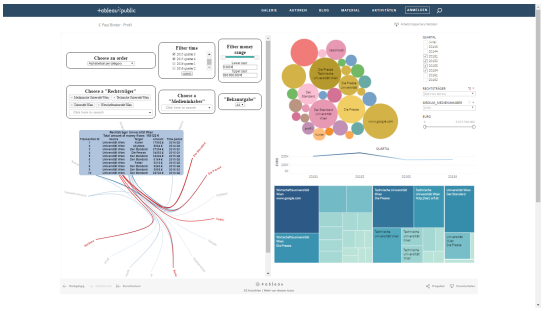
(c) Final step of the Use Case 1



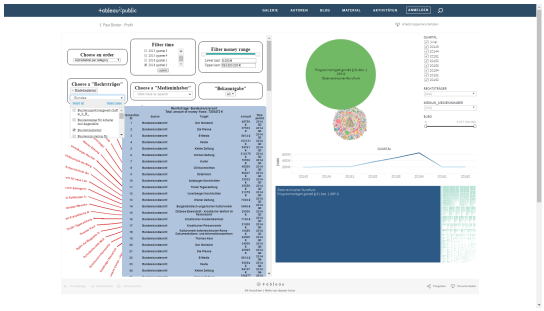
(a) Start of the Use Case 2



(b) Next step of the Use Case 2

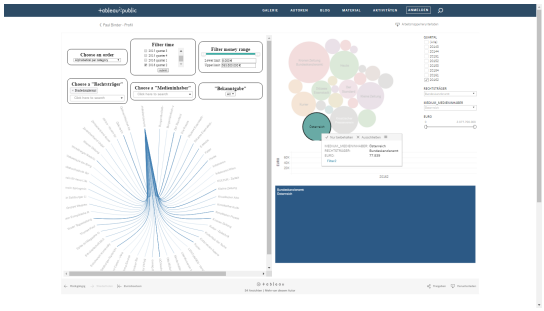


(c) Final step of the Use Case 2



(a) Start of the Use Case 3

(b) Next step of the Use Case 3



(c) Final step of the Use Case 3

6 DISCUSSION

6.1 Strengths and weaknesses of our approach and implementation

Strengths:
The dashboard provides a good overview about the Viennese media transparency network. The user can play with the filters and explore the data, for example in the bundled edge graph you can see the transactions as edges between Rechtsträger and Medieninhaber which are highlighted on mouse over. Furthermore, you also have tooltips for more detailed information, such as the time period and amount of money.
In the Tableau views, there is a line chart for trends, which displays the amount of money over time depending on your selected filters. The bubble chart and the tree map show the money transactions between Rechtsträger and Medieninhaber. If you select two or more Medieninhaber for example, each of them are displayed in good distinguishable colors. The tree map is more useful in 1:n relation but in a n:n relation the Medieninhaber in the bubble chart is easier distinguishable.

Weaknesses:
The default view of the graphs is unclear because of the high amount of data sets and a few data sets include a high amount of transactions and total money, for example the Österreichische Rundfunk. There are also no cross filters. So if you want to select a specific Medieninhaber or Rechtsträger you have to select them twice. You have to select them one time in the bundled edge graph and one time in the Tableau graphs. The performance of loading the data suffer by the large number of data sets. For a really professional implementation and for the economy, we would have to store some often used calculations or use a more efficient way of data accessing.

6.2 Lessons we learned

We learned to work more intensive with D3, Javascript and HTML and how the syntax works. Moreover, now we know how to combine them usefully and also how to integrate HTML and Javascript into Tableau.
We also learned some little but important syntax differences between Javascript and other programming languages and how to access arrays and object arrays and nesting versions of them. Furthermore, it is important to be aware about choosing colors, especially for color blind people. In addition, we were also confronted to deal with a large amount of data and to find a good or at least an acceptable representation of it.
We also learned more about Tableau and also its limitations, that it is nice and easy to get some basic graphs but if you want to do a little bit more, it will get hard to realise those things, for example making cross filters between our D3 bundled edge graph and our Tableau graphs.

6.3 Task separation

Paul Binder:
Tableau, D3 implementation into Tableau, Final Report and written reports.

Ulrich Schweinitzer:
D3 implementation, D3 implementation into Tableau, written reports and Final Report.

Philipp Dangl:
Did the majority of the D3 implementation of the bundled edge graph, Final Report