Media transparency in Austria

Markus Bader* David Kirchsteiger† Milan Peschl‡

1 Motivation

Based on the media law (MedKF-TG) the Austrian Regulatory Authority for Broadcasting and Telecommunications (RTR) is obliged to publish the data provided by organizations. If the fund is under a certain limit the amount is left empty. The law aims to make the media cooperation (public tenders) and funding transparent. The organizations and companies which are monitored by the audit court (Rechnungshof) are obliged to publish this data.

The common citizen very often struggles with processing such high quantities of data, which he/she is not familiar with, which leads to lack of understanding or even a feeling of not-involvement in such matters (even from an observer's point of view). Therefore, our greatest motivation was to transform those sets of thousands over thousands of data into a visually appealing, easy-to-understand and easy-to-control form, which can be valuable for a broad variety of people, ranging from most simple citizens, to journalists, analytics or even politicians.

We also have opted for elaborating on an another RTR dataset, which provided us with the state the corresponding corporations are set in. Those datasets have been merged to provide us with a better possibility to drill-down in the data and hence expanding the capabilities of our project.

Task

Providing easy to understand visualisations to portray the most important information in the data, which is the flow of money: "Which media outlet receives how much money from which governmental institution? How much money has the government spend in the x-th quarter of a given year? How much money went to Viennese cooperations?" It was important for us to portray as much information as possible in the simplest way possible.

Our goal was to visualize each part of the dataset as overseeable and comprehensible as possible and in relation to the other attributes. In order to enable the exploration of the data for the user, meaning the user can specifically search for specific corporations. We also included the expenses by paragraphs over time.

The matrix possibly portrays the most information in a very compact way. It shows distribution of payments from every governmental corporation to every media outlet. The two search boxes offer many opportunities: It can show how one specific corporation distributes their fundings. It can also be used to compare media outlets about how much they receive from each and every donor. Furthermore, the distribution type graph offers the possibility to further filter in the matrix and compare which media outlets are the strongest in different categories (e.g. internet vs. cable).

Users

Our focus is to provide a tool for a vast majority of user groups, ranging from a most common, partially interested citizen, to field-related persons, who might use the tool as a resource for professional purposes. We have chosen to present 3 different personas, for depicting possible scenarios and their iterations:

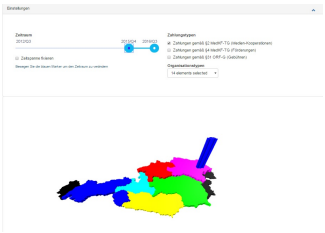


Figure 1: Map

The first user is a journalist seeking resources for her field-related article. They have a good understanding of visualisations, data and the topic it deals with and hence require extended options, when working with the visualised data. The journalist needs to be able to drill-down in the data to view more specific or meaningful information, for proper elaboration.

User 2 is a political activist, who is looking for information and data they can use in an upcoming talk they are about to hold. Much like the journalist, this user has some decent knowledge concerning visualisations, since he did this kind of stuff before many times and is experienced in reading graphs and charts and getting the most out of it. So even working across different dashboards is no problem for this user.

Our third user is the "average Joe" who is not trained or especially educated in visualisation and wants to know about the money flow between governmental corporations and media outlets. This user wants a quick and easy-to-understand overview over the data, hence rather basic graph types are preferred.

2 Related Work

What we used most, was the knowledge gained in the class, together with the feedback we were given for our Lo-Fi and Hi-Fi prototypes.

We do have found a resource of related-work, which was `www.medien-transparenz.at`. This project could be consider a very-well extended version of ours, offering everything we do, plus many extra visualisations and possibilities, such as:

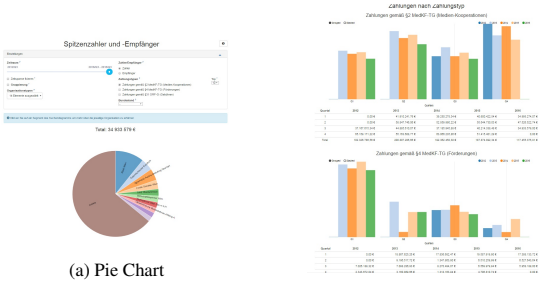
Interactive Map View PRO: catching, interesting, geographically explanatory, intuitive, easy to understand
CON: the user might not instantly see the difference between the federal states

Interactive Overview in Barcharts very similar to ours extended via ability to choose between Grouped / Stacked Barcharts and additional data in tables

Top Recipients and Spenders very interactive
CON: displayed via piecharts, which we were advised to not to use

Sankey diagram Fulfills the same task, as our matrix PRO: users instantly sees the comparison of sizes between notes
CON: nodes only show particular data when hovered, it's hard to

*MtrNr: 1404709
†MtrNr: 1402259
‡MatrNr: 1109587



(a) Pie Chart

(b) Temporal Bar Chart

Figure 2

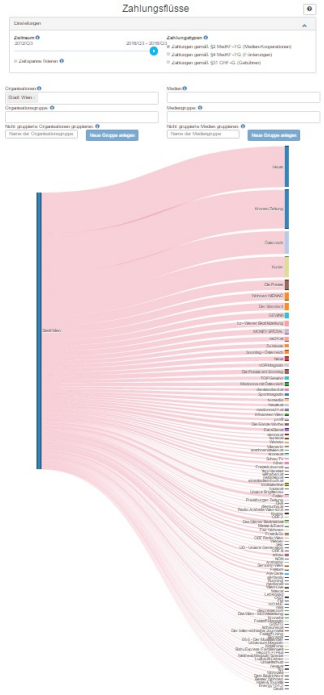


Figure 3: Sankey Diagram

select and navigate between the smaller ones

Filtering and selecting data Slidebars for selecting time periods (quarter-to-quarter)
Others like search-bars and checkboxes fulfill the same purpose as our filtering options, but in a more extended and fashionable way

Overall, this source was a great source of reflecting on our project, because when comparing the similarities and differences, we have found that there was much in common (in the purpose and some visualisation and data drill-down methods), assuring us we were heading in the right direction. The differences on the other hand have shown us, that our project might not be as valuable as theirs, but can pose as a valuable basis to be built on.

As an addition, we wanted to bring some real-life experience into the project, hence we consulted a real world finance expert, Alexander Pis, who works as a Junior Consultant at Cassis Ltd and is a friend of one of ours. He provided us with valuable information about what people usually want to see / understand / have visualised and how to properly do it, further helping us with our graph type choice and structure.

From our previous ideas, we only stuck with using barcharts (displaying data chronologically quarter by quarter, top spenders and top recipients) and expanded them a bit further.

3 Approach

Description

The data provided consists over a very large set of rows (114000) with six columns: RECHTSTRaeGER: The legal entity which has to publish their expenses on media or media funds to the Austrian Regulatory Authority for Broadcasting and Telecommunications (RTR). These entities can be Governmental Organizations (GO), ministries, Non-Governmental Organizations (NGOs), private corporations. QUARTAL: The quarter of the year the transaction took place. BEKANNTGABE: The publication law (paragraphs 2, 4, 31) as described above. LEERMELDUNG: An indicator that the EURO column is empty. MEDIUM MEDIENINHABER: The media name which receives the money. EURO: The amount of money (in Euro).

As you can imagine the possibilities to drill down further are quite limited since it lacks of ordinal information about the media or the legal entities. Since filtering by law and time are the only ways to categorize, we tried to join our data with other datasets from `www.rtr.at`. This enabled us to gain information about the state the media is located and the distribution type (cable, internet, analog, digital, satellite) of the media. Unfortunately, this information is only provided for very few media companies.

At first, we considered multiple option for visualising our data, such as JavaScript libraries (d3, plottable), but then decided to opt for Tableau, since it is already a very powerful visualisation tool.

Our visualisation changed greatly from our first Lo-Fi prototype to now. In our Lo-Fi Prototype we used more and more complicated visualisations (including bar charts, line charts, pie charts, heatmaps etc.), but due to the feedback, we decided to follow a different path, with less graph types, but more in-depth elaboration, including drill-down possibilities.

Design Choices

In order to meet our goals we split the vis into two dashboards: Time and Top overview and detail search view. The first dashboard aims to give the user an impression of the top spenders and recipients (S&R) and the temporal development of the expenses in each paragraph. The view consists of 3 individual graphs: At the top there are two stacked barcharts showing the S&R with highest amount of money spent, which is shown on the y-axis. The color channel indicates the paragraph, on all three charts. The bottom barchart shows the sums of money spent in one quarter of a year, horizontally divided by the paragraph. These barcharts enable an comprehensible and accurate vis of the data. In order to observe the partial sum of one entity in one quarter of a year, the user can hover over the bars of the time barchart, which highlights the slice on the top two barcharts, representing the money spent in that quarter of a year. Also by hovering over the top bars, a tooltip shows the money spent in a quarter of a year. Thus, the user can explore the temporal development of the top S&R and of the paragraphs and compare the paragraphs by their sum. The ”Top-view” allows the user to get to know the most important entities covered by the data. The paragraph 31 (ORF-G), is only shown in the temporal view, since the only entity is the austrian broadcast (ORF). The second dashboard tries to give the user the ability to explore the data in more detail. The central matrix view shows the amount of money transacted between one S&R pair, which are represented by the columns and rows. The value of the money is also coded in sequential colors to ease the comparison between cells. The columns are sorted descending by their total sum, afterwards the rows are sorted in the same manner. By sorting the cells the user won’t lose the overview. The user can now search for any entity in two search inputs, which will alter the matrix. This gives the user the ability to compare any entities and observe their total range of

transaction partners. The bottom view shows the distribution types and their sum in a barchart. As mentioned above most rows don't hold a distribution type field. The reason why we still decided to keep this view is that for a few entities (especially of paragraph 4) the distribution types are covered entirely, so one can examine the distribution over the distribution types well, which is an very worthwhile information in our opinion.

General speaking, we focused on eliminating everything that was either inefficient, or redundant and rather replaced it with simpler, but not less effective visualisations. Sticking to fewer graph types allows the project to be more consistent and generate less cognitive load. Additionally, the drill-down capabilities of the dashboards not only compensate the "simpler graph type selection," but also expand the options the user has in end-effect, which leads improved user experience and widening the possible target audience.

4 Implementation

As was already mentioned above, we used Tableau for creating all our implementations and published them via Tableau Public. This provided us with the feature to easily implement them on our website using HTML5 and JavaScript. What is definitely an added feature, is the reusability of these dashboards (whether we are talking about implementation into other projects or the ability to share / download the dashboards), which is certainly a valuable addition. Although Tableau is very intuitive and simple, the flexibility is limited to a few graph types. Also more complex interactions and dynamic filtering techniques (like double color coding onclick) aren't supported. Especially in our case the plain data could be displayed more attractively (e.g. instead of matrix view, a sankey diagram). Concerning the simplicity of the data we first struggled to build an attractive and interactive visual representation, which doesn't only consist of barcharts. Although the most chart are still of barcharts we managed to enable a drill-down. Also the raw data contains inconsistent names of media and organizations (e.g. Antenne Steiermark, Antenne Steiermark Regionalradio GmbH, Antenne Steiermark Regionalradio GmbH & Co KG).

5 Results

Performance

The performance of the project is pretty decent, everything works as expected. However, there is a noticeable Tableau dashboard loading time occurring in the first run of the website.

Additionally, the implementation of a search box (as we were instructed in the previous feedback) is a valuable tool, but brings a small space for errors to the game, since, as mentioned in Problems and Challenges above, name inconsistency is present.

Scenarios

A journalist wants to write an detailed article about distribution of money towards media outlets based on the way they are distributed

Julia Wilhelm is a Journalist for the Vice Magazine and wants to write an article about how much money well known media outlets receive from governmental institution sorted by the distribution type of the media. Go to the 2nd Dashboard Choose the distribution type by clicking on the bar.

She is presented with the overview in Fig. 4: If she wants to see only results for chosen outlets, she can do this by typing their name into the search bar below "Medium Medieninhaber" Fig. 5.

A political activist wants to hold a talk at a TEDx event about potential influence of the government over the political stance of different media outlets and if the money those outlets receive may result in a bias on how they report about different topics.

Peter Mller is a political activist who is pretty famous for his various talks about the influence of the government in news. He recently stumbled upon our website and decided to use the information he found to prepare for a talk at an upcoming TEDx event in Vienna. In

Figure 4

Figure 5

the first Dashboard he gets an overview over major donors/recipients and a timeline showing the total expenses sorted chronologically by quarter. Fig. 6

He can now use this information in the second dashboard to find out from whom exactly the mayor recipients receive the money. Mller does this by typing their names into the search boxes Fig. 7 and investigate further if those tend to whitewash news about their main contributors.

The average Joe wants to find out which media cooperations get the most money from the government

Herbert Groening saw a documentary on TV about potential government influence in the kind of news big media corporations report about. In that documentary he has heard about the RTR and the kind of data they publish. He decides to further look into it and stumbles upon our website and wants to get an overview and find out who the main donors and recipients are. He has to use the first dashboard. Here he is presented with an interesting overview over the whole dataset Fig. 6.

Now, he decides to check out the timeline to figure out how much the above really spent/received during "times of interest", like the

Figure 6

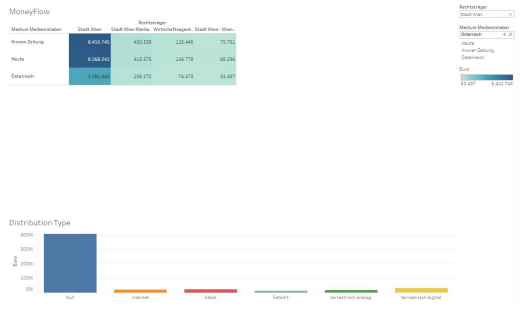


Figure 7

Figure 8

presidential election 2016 of Austria since he notices a spike in quarter 1 and 2 of that year. Fig. 8

Feedback

We asked two laymen, who haven't pre experience in vis or the topic of media cooperation. The process of interrogation was structured in first explaining them the context and the let them independently without intervention of the supervisor explore the vis. The interviewees were asked to think aloud, which was written down by the supervisor. The first Interviewee mentioned that the highlighting in the first dashboard can be misleading to think that the money flow is from the top spenders to the top recipients, although their independent. Also the stacks and their representation aren't obvious and clear at first sight. At the second dashboard it wasn't clear whether a the clicking a row, column or cell is triggering an alteration of the vis. The second interviewee (Mr. Bacher) also mention the last point and furthermore suggested to alter the distribution type view also after searching one specific media or organization. Moreover he was overwhelmed by the tooltips with exclusion and hide options from tableau which actually don't belong to the vis. He also wondered where to deselect the current selection. In order to overcome the points mentioned above we tried to hide all misleading information and describe the procedure accurately in this paper.

6 Discussion

The biggest strength of our visualisation is that it transforms a huge data set into a really simple and meaningful visualisation, which, when combined with additional drill-down options, directly tackles the given task. At a glance a user sees the most important information like the major donors or recipients. If the user wants more specific data they have the possibility to drill down on it via the matrix, where he/she just has to type in the desired company into the searchbox and receives exact amounts of money, separated into columns, showing which spender spent how much. In comparison with the visualization from `www.medien-transparenz.at` our matrix gives on the one hand a more compact view of the moneyflow and is able to show more media corporations and organizations at the same time with the absolute values. On the other hand it lacks of time filtering. Furthermore comparing our top S&R view from dashboard 1 with the counterpart from `medien-transparenz.at`, a bar chart enables an easier comparison between columns in contrary to a pie chart. But the proportion is missing in the barchart.

The simplicity and quantity of the graphs might be considered also a weakness. While it serves the given purpose (providing a easy-to-understand and control visualisation) and generates less cognitive load, if we would've implemented more complex methods of visualisation (for example in different dashboards, which would be present only to satisfy various needs of analysts and other professionals). There are a few disadvantages: On the first dashboard the comparison on one bar between the different quarter-of-years (stacks) is can be quite hard, since one has to hover over to see the stacks. There could have been chosen a color coding for each quarter-of-year instead of the binary paragraph value. Also the time view at the bottom could be grouped by quarter-of-year instead of grouping by paragraph. Like our interviewees told us, the second dashboard could be simplified by displaying the searched parameter immediately instead of having the user to click on a row or column. However, it wouldn't be possible to display entities with similar names (e.g. If one searches for "ministerium", it displays all ministries). Of course, the bottom distribution type view could be roughly criticised, since there exist so many null values. The reason for this decision is that there are entities (e.g. Red Bull Media House GmbH) which don't contain null values. This information might be very important to users, who want explore the distribution of a single entity instead of the overall distribution. Furthermore we couldn't find any information or dataset to enrich the organizations, though it would have been nice to categorize them.

Lessons learnt

The lesson we learnt is, that instead of focusing on variety and visual fanciness, we should always go with what is proven to be functional and efficient for the given task. The major problem we had, was that the prototypes didn't match with the data and the technology we planned to use. Thus, we missed to iterate the prototypes and improve our ideas. We encountered that vis is not only about static graphs, but about enabling exploring and/or build models on top of the data through visualisation. This makes it clear that drilling down the data to its very details is an important but not trivial task. The visual mapping should be identified as early as possible and should be designed consistently without redundancies.

7 Work Separation

- David: M4 Report: Implementation, Task, Approach, Discussion; Final product: Dashboard Design (teamwork), Dashboard Construction in Tableau
- Markus: M4 Report: Tasks, Users, Implementation, Results, Scenarios, Discussion; Final Product: Dashboard Design (teamwork)
- Milan: M4 Report: Users, Related work, Implementation, Approach, Discussion; Final Product: Website, Dashboard Design (teamwork)

8 References

- `http://www.unet.univie.ac.at/~a1109587/final_product.html`
- Alexander Pis, Junior Consultant, Cassis Ltd. (for M2 and M3)
- M2 & M3 Feedback

- www.medien-transparenz.at
- ww.rtr.at