
Interactive Visual Exploration of the Regesta Imperii

Markus John

markus.john@vis.uni-stuttgart.de
University of Stuttgart, Germany

Christian Richter

christian.richter.rr@gmail.com
University of Stuttgart, Germany

Steffen Koch

steffen.koch@vis.uni-stuttgart.de
University of Stuttgart, Germany

Andreas Kuczera

kuczera@uni-mainz.de
University of Mainz, Germany

Thomas Ertl

thomas.ertl@vis.uni-stuttgart.de
University of Stuttgart, Germany

Introduction

The Academy of Sciences and Literature in Mainz provides online access to the Regesta Imperii - a very extensive historical dataset based on documentary sources of German-Roman kings. About 125,000 regestae of emperors and popes are searchable and viewable in this online portal. The current user interface offers direct access to single documents through different form-based search facilities, as well as through a catalogue that directly reflects the structure of the regestae volumes as they have been created in this long-term project. In order to further improve access to this large data volume, we suggest an additional approach based on coordinated views. The usage of coordinated view approaches is very common in many domains (Stasko et al, 2008, Vuillemot et al, 2009, Koch et al, 2011). However, there is no publicly system available, which would provide a suitable access to this historical dataset. The motivation for this new approach is twofold. On the one hand, we improve the support for search and exploration tasks in this historical data set that are

based on imprecise information needs or on a less deep understanding of the available information. In practice, such imprecise queries can quickly lead to an overwhelmingly large number of search results. Allowing users to create and refine queries in a visual way, while offering immediate feedback on the number of entries requested, can help to cope with underspecified queries and help refining them iteratively (Jänicke et al, 2012). On the other hand, we offer a powerful means for visually analyzing the available information and understanding complex relationships by providing different linked perspectives on subsets of the collections. These perspectives include views on historic persons and entities as well as temporal and spatial information contained in the regestae. A usage scenario shows successful application of the approach.

Visual approach

We implemented a web-based visualization that is easily available to humanities scholars, since it does not require users to install software. The web-based visualization uses the library Data-Driven Documents (D3) (Bostock et al, 2011) and runs with a web browser supporting HTML5, SVG, CSS, and JavaScript.

Data processing

In the Regesta Imperii database all documentary sources of German-Roman kings are available in full text (xml-files). This data set consists of regestae volumes and register information. Regestae volumes are short summaries of a text and contain important metadata of a document, such as the title of the document, an ID for the unique assignment, date of issue, and place of issue as name and coordinates.

Since places, persons, and institutions may be known by different names, entries can be overlooked within the full-text search. Therefore, place and personal registers provide a central resource for work with the regestae and contain a list of places, persons, institutions, and additional information, such as the numbers of the regestae in which the entity is mentioned, a reference to another entity entry (if available), and a unique id.

Since the regestae have been manually digitized, the xml-files can contain syntax or other transmission errors, such as different date formats, geo-coordinates or tags. Therefore, the data must be parsed and well prepared in order to use them for a visual analysis.

Visual approach

After the regestae and register volumes have been successfully parsed and loaded into our system, users can start their exploration in the main view as depicted in Figure 1. The main view consists of the five coordinated views: (A) timeline view, (B) map view, (C) register view, (D) overview filter view, and (E) results view.

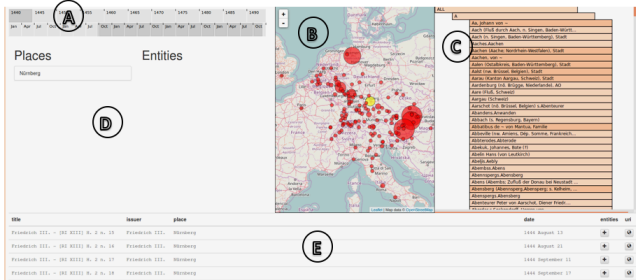


Figure 1. The interactive visualization approach for exploring and analyzing regestae of royal and papal records.

We initially depict all available data and enable users to create and refine queries in a visual way iteratively, to reduce the overall set. For example, users can select certain time periods and places in which the regestae were published or persons and places that are mentioned in the regestae.

The timeline view consists of two stacked timelines and enables users to select a time period by clicking and dragging, as depicted in Figure 6, comparable to the Simile approach (Huynh, 2008). The first timeline allows a coarse filtering and represents the whole-time range of the regestae. Once a timespan is selected, the second timeline is updated and permits a finer selection of the upper selected time range.

To get an overview where the regestae were published, users can discover the map view as depicted in Figure 2. The map view uses the JavaScript library Leaflet (Agafonkin, 2014) which supports interactive features such as zooming and panning. The red circles in the map represent places where the regestae were published and the circles size is scaled proportionally to the places occurrences, similar to the approach (DARIAH-DE, 2015). This helps to get a quick overview of important places. When hovering over a circle, a tooltip shows the corresponding place name. By selecting one or more circles (highlighted in yellow), the places are added to the search query.

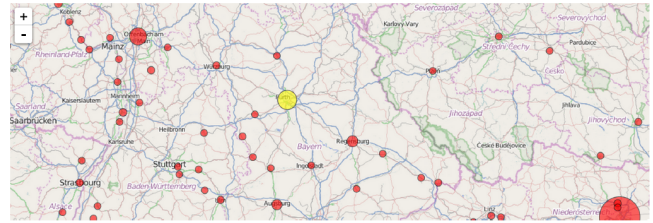


Figure 2. The map view gives an overview where the regestae were published.

The register view represents the entities from the register volumes in an alphabetically sorted hierarchical structure as depicted in Figure 3. Users can explore the different entries by clicking on the different hierarchies. Nodes, which contain further entities are displayed in darker color. Furthermore, users can select one or more entities to adapt the search query.

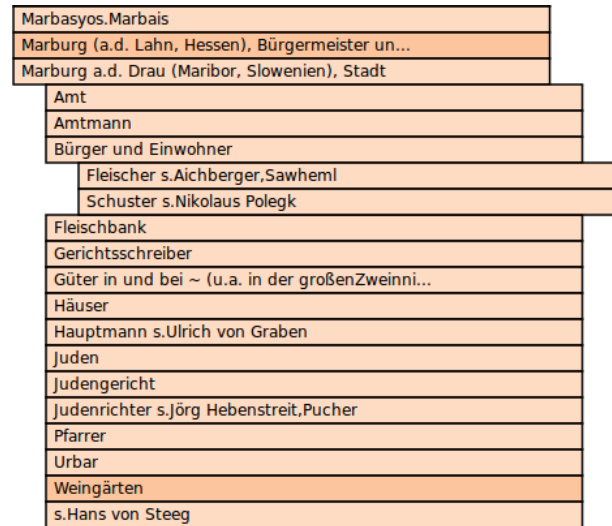


Figure 3. The register view represents all persons, places, and institutions which are mentioned in the regestae.

From the overview filter view, users can get a summary about all the selected places and entities which determine the search results. In addition, users can deselect places and entities from the list to adapt the search query.

Places

Linz
Nürnberg
Waldshut

Entities

Aachen (Aache; Nordrhein-Westfalen), Stadt
Einwohner und Bürger
Fischmarkt (Farwisch)
Kopist s.K., F. Meyer

Figure 4. The overview filter view gives an overview of all selected places and entities.

Based on the combined search query, the result view lists all regestae entries, which are included in the search results as depicted in Figure 5. For each entry, the list displays the following metadata: title of the regesta, issuer, place and date of issue. By clicking on the icon in the column entities, users get the information of all entities occurring in the regesta in a separate list view. Furthermore, users can select the icon in the last column uri to jump directly to the corresponding regesta entry on the web page of the Regesta Imperii. This enables users to analyze the regesta in detail.

title	issuer	place	date	entities	uri
Friedrich III. - Chmel n. 1758	Friedrich III.	Nürnberg	1444 September 25	+	🔗
Friedrich III. - Chmel n. 1776	Friedrich III.	Nürnberg	1444 Oktober 4	+	🔗
Friedrich III. - [RI XIII] H. 14 n. 281	Friedrich III.	Nürnberg	1444 Oktober 4	+	🔗

Figure 5. The result view displays all regestae entries which are included in the result set.

Usage scenario

In the following section, we present a usage scenario that occurred during one of the joint workshops of users from the Academy of Sciences and Literature in Mainz and the developers of the approach. Therefore, our usage scenario represents instead the insights and lessons learned through several sessions with two experts from the Regesta Imperii.

In a first step, the expert explores and analyze the map view. That way, she gets a quick overview of the places where regestae were published. During the exploration, the place Nürnberg (Nuremberg) has aroused her interest, since she has already examined these entries a long time ago. To find out more about the regestae entries, she starts a search query by selecting Nürnberg (Figure 6A) in the map view and discovers the search results in result view. However, the result set is too large for a further deeper analysis. Therefore, she refines the search query by selecting the time period from 1440 to 1450 in the timeline view (Figure 6B), because she is especially interested in the early regestae entries. As the next step, she searches for entities that are mentioned in the regestae with the aid of the results view. She finds out that there are many connections to French entities. To further analyze that, she selects the hierarchy “Frankreich, Königreich” (France, Kingdom) in the register view, as depicted in Figure 6C. By adjusting the search query, she received a specified subset of the collection for a further analysis. This way, she finds that primarily French kings are mentioned in the regestae entries.

During the analysis, she learns from the map view that Neustadt near Bremen (Figure 6D) has many regestae entries which she did not expect. To inspect this in detail, she selects Neustadt and explores the result list. By analyzing the different entries in the list and web page of the Regesta Imperii, she finds out that the geo-coordinates were manually digitized incorrectly. Consequently, the expert corrects the entries in the database by assigning these entries to the actual place Neustadt near Vienna.

While these sessions, we received a lot of feedback that showed that our approach improves access to the large data volumes of the Regesta Imperii and facilitates search and exploration tasks, as well as assisting in understanding complex relationships and gaining new insights.

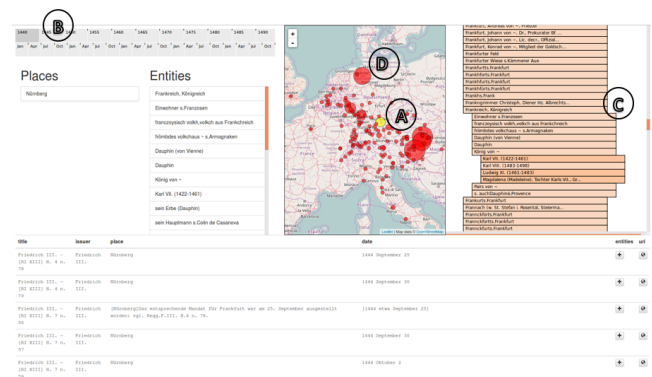


Figure 6. An example search query for the place Nürnberg, French entities, and the time period from 1400 to 1440.

Conclusion and future work

The presented interactive web-based approach has been evaluated through expert feedback that recommends it as an effective method for exploration analysis.

We are planning to extend the different linked views to support users with additional information. Concerning this issue, we have implemented the relative distribution of the regestae volumes in the timeline view, as depicted in Figure 6, and we are currently working on the co-highlighting between the views. We will also ensure that experts from the Regesta Imperii are able to correct errors that arise during the digitization process interactively.

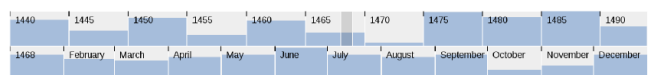


Figure 7. Timeline filter view of the selected year 1468 with a relative distribution of the regestae over time.

Bibliography

- Stasko, J., Görg, C., and Liu, Z.** (2008). Jigsaw: Supporting investigative analysis through interactive visualization. *Information Visualization*, 7(2):118–132.
- Vuillemot, R., Clement, T., Plaisant, C., and Kumar, A.** (2009) What's being said near "martha"? exploring name entities in literary text collections. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology*, 2009. VAST 2009., pages 107–114.
- Koch, S., Bosch, Giereth, H. M., and T. Ertl, T.** (2011). "Iterative Integration of Visual Insights during Scalable Patent Search and Analysis," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 5, pp. 557-569.
- Jänicke, S., Heine, C., Stockmann, R., and Scheuermann, G.** (2012). Comparative visualization of geospatialtemporal data. In *Proceedings of the 3rd International Conference on Information Visualization Theory and Applications*, IVAPP, pages 613–625.
- Bostock, M., Ogievetsky, V., and Heer, J.** (2011). "D³ data-driven documents." *IEEE transactions on visualization and computer graphics*: 2301-2309.
- Huynh, D.** (2008). SIMILE—Timeline. Simile Projects, <http://www.simile-widgets.org/timeline/> (accessed 28.10.2016).
- Agafonkin, V.** (2014). "Leaflet. an open-source javascript library for mobile-friendly interactive maps." (accessed 28.10.2016).
- DARIAH-DE.** (2015). Geo-Browser: <https://geobrowser.de.dariah.eu/>(accessed 28.10.2016).