

Exploring Interpersonal Relationships in Historical Voting Records

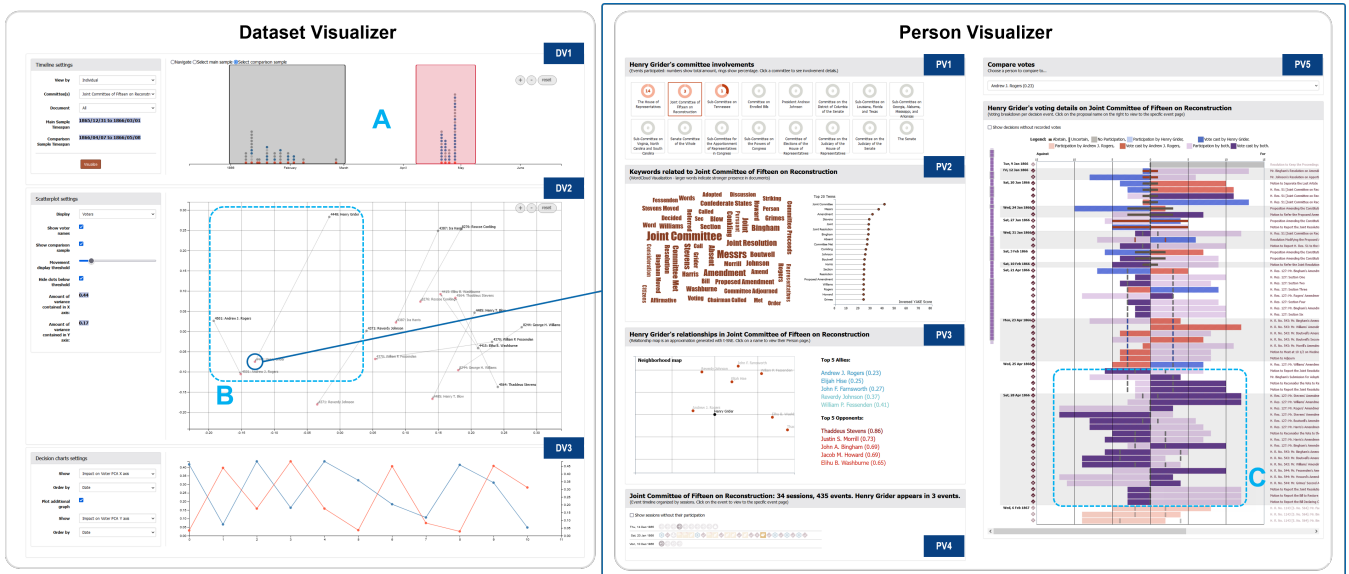
G. D. Cantareira¹ , Y. Xing¹ , N. Cole² , R. Borgo¹ , and A. Abdul-Rahman¹ ¹King's College London, United Kingdom²University of Oxford, United Kingdom

Figure 1: In the dataset visualizer interface (left, panels DV1-3), voting records from a constitutional convention dataset can be explored, and voting patterns from people can be visualized as projections. Data can be filtered and sampled, and information about the quality and composition of the projection itself can be visualized. Records from individual people can be investigated using the person visualizer interface (right, panels PV1-5), which contains committee involvement, keywords, relationships, event participation, and voting breakdown. The blue highlights show an example of workflow: parts of the timeline are chosen to be projected (A), a point of interest is identified in the projection as two people became closer (B) and the reason for that behavior - a large section of votes in agreement - is seen on the vote breakdown (C).

Abstract

Historical records from democratic processes and negotiation of constitutional texts are a complex type of data to navigate due to the many different elements that are constantly interacting with one another: people, timelines, different proposed documents, changes to such documents, and voting to approve or reject those changes. In particular, voting records can offer various insights about relationships between people of note in that historical context, such as alliances that can form and dissolve over time and people with unusual behavior. In this paper, we present a toolset developed to aid users in exploring relationships in voting records from a particular domain of constitutional conventions. The toolset consists of two elements: a dataset visualizer, which shows the entire timeline of a convention and allows users to investigate relationships at different moments in time via dimensionality reduction, and a person visualizer, which shows details of a given person's activity in that convention to aid in understanding the behavior observed in the dataset visualizer. We discuss our design choices and how each tool in those elements works towards our goals, and how they were perceived in an evaluation conducted with domain experts.

1. Introduction

Voting records from legislative assemblies are a valuable source of data for understanding the political landscape and overall legislative process from democratic societies, and the analysis of roll call voting data is a common subject of study by political scientists [NMdSSDSF18]. In a historical context, the complexity and size of such datasets further increase when they are enriched with records and timelines from assemblies in which voting occurred, the actual text being voted upon, transcriptions from speeches and

debates, and so on; understanding how these elements interact with roll call data is essential in building a mental picture of events unfolding in voting sessions.

The Quill Project [CARM17] is a research initiative that investigates text analysis and data visualization with the aim of helping humanities researchers understand modern foundational legal texts and constitutional conventions. Quill contains multiple well-organized datasets built on an event-based structure, representing the negotiation of legal texts during legislative drafting for consti-

tutional conventions – processes that can last for several months or years. As each discussed text is the result of hundreds of amendments proposed by several actors over an extended period of time, visual presentation and navigation can become quite difficult. Analyzing political alignment and uncovering information about certain decisions or people can be particularly challenging due to the large timelines involving several people and events that may or may not have relevant voting records.

In this paper, we present an ensemble of tools focused on visually exploring relationships between people in voting records from Quill. The ensemble consists of two interfaces, a *dataset visualizer* and a *person visualizer*. The former uses multidimensional projections to show the evolution of voting behavior over different time periods in the dataset, while the latter contains detailed information on individual voters. We discuss our approach to designing each tool, including different iterations and communications with experts, present how the ensemble tackles our design goals, and finally discuss its evaluation. In short, our main contributions are: (i) A toolset for exploratory analysis of roll call vote data and providing context to the individuals involved in voting; and (ii) A case study on designing and developing such a tool, navigating domain expert expectations and constraints incurred from the need of integration to a larger platform.

2. Related Work

2.1. Visualization and Digital Humanities

Although visualization research for the digital humanities (DH) has existed as a field for quite some time, it has experienced a steady increase in interest from the visualization research community over the last few years [BSS20]. With the increasing computational capacity of modern systems and the evolution of information visualization techniques, new possibilities arise to extract new secrets from data, sometimes centuries old [BEAC*18]. Under the umbrella of *Visualization for Digital Humanities*, we can cite three applications in the history domain that are relevant to our work: visualization of documents, timelines, and relationships.

Once historical documents are digitized, any text visualization technique [KK15] can be directly applied to them. However, DH experts often have specific needs or a workflow that would benefit from specialized applications. Janicke et al. [JFCS17] discuss different uses for text visualization tools in humanities, and offer a taxonomy of tasks that are performed with them. Fields such as history are often concerned with relationships between texts, and how they can be summarized [HAC*15, KOTM13]. Other fields such as linguistics or literature may be concerned with structure or the sound characteristics of text in poetry [MLCM16, ARLC*13].

Approaches for timeline visualization for DH are often event-based. Event visualization methods focus on describing or summarizing events, and relationships between them [GGJ*21, FBM16]. Datasets from historical records often contain graphs of several interconnected elements that interact with one another over time, and visualizing these interactions in a concise and informative manner can be a challenge [FSR*21].

Visualizing relationships between historical figures can often be

a complex task due to the types of data available and their limitations. While some applications may work directly with historical records [IA12, KOTM13] such as voting, speeches, or minutes, others rely on indirect data, such as presence locations [OIK*11] to build a picture of events and relationships. Finally, it is important to note that certain applications may need to connect multiple approaches – for instance, investigating how events can affect relationships over time [NMM*14, FSR*21].

2.2. Visualizing Voting Records

There is an extensive history of research in interpreting voting records (*roll call data*) and understanding relationships between voters in historical legislation data [BL16]. The use of visualization techniques to aid in such tasks is not a new concept, and many different approaches for this application have been proposed and discussed [MCLL17, MMM22] over the years, from charts and line plots to sophisticated pattern recognition algorithms. Particularly, other approaches have been proposed for analyzing records similar to the ones stored in the Quill platform, such as United States constitutional conventions [ALH*15, SM07].

Among these methods, multidimensional projections can be highlighted for their ability to display complex relationships between data instances in high-dimensionality spaces. Humanities researchers have been refining techniques to mathematically model voting ideologies in a manageable number of dimensions, such as the NOMINATE model and its developments [BL16, PR99], for decades, and new metrics to describe political alignment in multidimensional space are still being developed [PRDC20]. Brigadir et al. [BGCC16] discuss different techniques and applications for dimensionality reduction in voting records. In 2007, Slez et al. [SM07] employed traditional multidimensional scaling (MDS) to visualize changes in voting alignment between states from the 1787 US constitutional convention voting records, over multiple predefined periods. More recently, da Silva et al. [NMMSSDSF18] presented CivisAnalysis, a framework composed of multiple visual tools that explores roll call voting data, including Sankey diagram-based timelines and multiple projection techniques: Principal Component Analysis (PCA), t-distributed stochastic neighbor embedding (t-SNE) [VdMH08], and MDS.

A common consideration with the aforementioned approaches is that time windows to visualize and compare voting from are often considered to be already defined in the dataset by historical context, and therefore exploratory sampling of the timeline takes a secondary role. Additionally, they are not expected to be connected to detailed representations for other entities in the domain, such as people or other timeline events in the case of Quill, which poses a few further challenges.

3. Design Requirements

3.1. Quill

The Quill platform [CARM17] is a system designed to store and present historical records related to the negotiation of legislation texts. These negotiations may be described using different types of actors and events, depending on the historical context and political

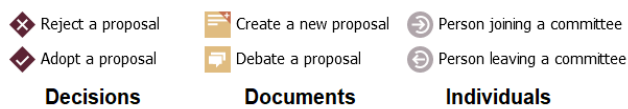


Figure 2: Examples of different event icons in the Quill platform.

structure of the negotiation in question. Quill's main intended audience is humanities academics, such as researchers and teachers.

Negotiation processes stored in Quill are modeled after official and unofficial journals and minutes, recording proposal texts, details of decisions, and descriptions of speeches made during sessions. Data is structured as events, linked to relevant people and event sequences, building a representation of debates and discussions. Therefore, an analytical tool for Quill may require pulling data from multiple entities in its relational database to build data instances containing all information relevant to the task at hand.

By default, Quill organizes its datasets as timelines consisting of sequences of *events*. The data entities from Quill that are relevant to our toolset are listed as follows:

Event. Moment-to-moment events that happen during a Session: proposals, motions, speeches, and so on. Each type of event can have different data associated with it.

Session. Meeting of representatives occurring on a given date; sessions happen under a Committee.

Committee. A group of representatives formed to discuss a certain topic; part of a Convention.

Convention. Name given to a Quill dataset. Encompasses a set of committees, people, documents, sessions, and events.

Person. Part of one or more committees. Can also appear in individual events as a participant.

Document. Text with a history of changes by different events.

Decision. Particular type of event that indicates change or amendment that requires voting by attending participants. Includes vote counts and outcomes, and may include records of individual votes cast by each participant.

Please refer to [CARM17] for a more detailed explanation of the data structure in the platform.

Quill also uses a common icon system for events on all of its interfaces, which users are already familiar with. Therefore, this system *was to be* employed when different event types are presented in visualization tools. An event taxonomy divides all event types into four categories (*individual*, *decision*, *document*, *procedural*) that contain distinct visual features. Figure 2 shows a few examples of the different icons on the platform.

3.2. Domain Requirements

We have worked closely with domain experts and compiled the following set of requirements based on our interactions. While requirements were mostly defined during initial meetings and discussions, refinements were made over the course of our preliminary user studies and surveys. More details on our discussions with experts throughout development can be seen in Section 6.

R1 Insights on voting interactions. The visualization tools should

enable the exploration of patterns and points of interest in votes over time to help identify trends or groups.

R2 Insights on people's roles and behavior. The visualization tools should provide information on people's actions, involvement, and voting patterns as to provide a picture of their role in a given convention.

R3 Access to direct information. Users may not be familiar with certain abstract representations or analytical models. There should be direct links between visualization and data.

R4 Adherence to existing platform conventions. New visualization tools should respect Quill's pre-existing visual language and layout elements that are already familiar to the users.

3.3. Design Tasks

The set of design tasks devised to meet the domain requirements is as follows:

T1 Select and filter through data. For **R1**, **R2**, **R3**. The Quill platform has a complex structure of relationships between elements such as people, delegations, committees, and events. Providing ways to subdivide them for navigation is essential.

T2 Comparison between people. For **R1**, **R2**. The tools should provide ways to compare the behavior of different people, observing similarities and dissimilarities.

T3 Establish timelines. For **R1**, **R2**. Users should be able to view the sequence of events in a convention or committee, either for voting or a single person's actions.

T4 Provide person context. For **R2**, **R4**. The tools should provide information on how a person fits in the context of a convention or committee, and what it entails.

T5 Links to existing systems. For **R3**, **R4**. Whenever an element is shown or referred to, connections should be made to existing pages in the platform that explain it.

T6 Explainable visualizations. For **R3**. Visualizations should be directly linked to the data, whenever possible.

4. Toolset Description

The toolset presented in this paper consists of two interfaces, presented as webpages, designed mainly with d3.js [D3]. The first is a *dataset visualizer*, showing a timeline of voting events for an entire convention dataset. The visualizer allows users to sample the dataset and explore samples using 2D projections. The second is a *person visualizer*, which shows details about a person's behavior in a convention.

The workflow between the two interfaces is shown in Figure 3. The components in the *dataset visualizer* have direct links to both the *person visualizer* page and external existing systems in the Quill platform, such as description pages for events, sessions, and committees. They also have indirect links to components in the *person visualizer* page, such as event lists, relationships, and vote breakdown panels that use similar data. While examining vote records in the *dataset visualizer*, people of interest can be selected and examined in further detail in the *person visualizer* interface.

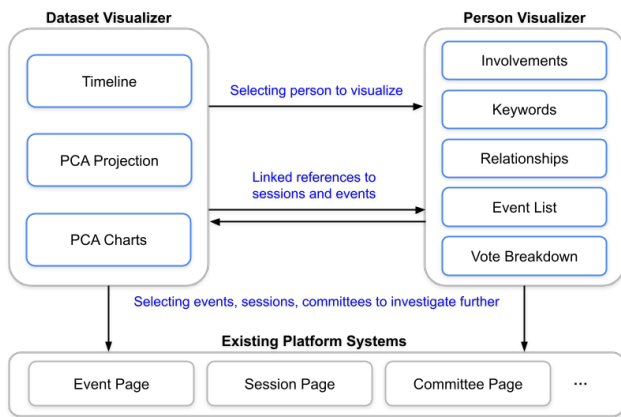


Figure 3: Interactions between the two interfaces in the toolset and other existing systems in the Quill platform. Both interfaces connect to existing systems, such as description pages for events, sessions, and committees, while also sharing information and linking one another as people of interest are identified and examined further.

4.1. Dataset Visualizer

The *dataset visualizer* consists of three panels, as shown in Figure 1: a timeline (DV1), where samples of voting events in a time interval can be selected for projection, a PCA scatterplot (DV2) generated using the selected sample, and a set of charts (DV3) containing information about the PCA projection.

For a given convention dataset, the *dataset visualizer* gathers metadata from all *people*, *sessions*, and *decision events* in the convention. From decision voting records, we generate a $N \times M$ vote matrix V , where N is the number of people and M is the number of decisions in the dataset. Each entry V_{ij} is given by the vote from voter i in decision j : -1 for a vote against, +1 for a vote in favor, and 0 otherwise (abstentions or otherwise unrecorded votes).

4.1.1. Timeline

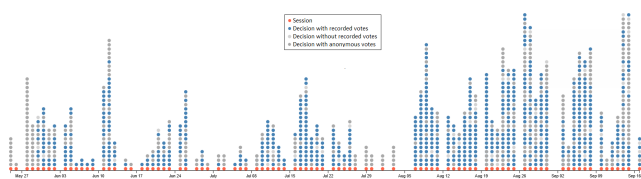


Figure 4: Timeline representation for U.S. Constitutional Convention 1787 dataset. Each orange dot is a session, and the other dots above it are decisions, chronologically ordered from bottom to top. Decisions with individually cast votes available are shown in blue, while decisions with only vote counts or results are shown in gray.

The timeline is drawn as a series of dots positioned over a horizontal line containing the total timespan of the convention. Each session is represented as an orange dot. Each session is described by a sequence of events, and events that may include voting (decisions) are shown as secondary dots on top of the session dot in chronological order. A decision that has a record of individually cast votes (either by people or by delegation) is shown in blue,

while decisions without those records (only total votes and/or outcome) are colored in gray. A decision may not have individual voting records for several reasons, such as anonymous voting or having an outcome decided by the chairman. Anonymous votes are colored in a darker gray tone. Figure 4 shows the timeline.

This timeline representation serves two purposes: the first is to allow users to have an overview of the event distribution in the dataset (**T2**), showing how sessions are organized over time and their varying amounts of decisions, serving as a guide to selecting samples. The second is to provide insights into data availability and uncertainty, since the user is instantly made aware if blank or gray-filled areas do not correspond to their mental model of the convention (**T6**).

Users can interact with the timeline panel in several ways: navigation (zooming and panning), investigating individual dots (hovering the cursor shows details about the session/decision, and clicking opens its description page in the Quill platform, **T5**), filtering (by committee or related document), and selecting a set of decisions for projecting (**T1**).

4.1.2. PCA Projection

The timeline panel allows for the selection of either a single sample or a *main* sample and a *comparison* sample. A sample consists of a series of decisions filtered by time, committee, and related documents. Once samples have been selected, they are projected as a 2D scatterplot in the projection panel, with the aim of providing insights to users regarding relationships in voting behavior (**R1**, **T2**).

From a sample S containing a set of decisions filtered by session date, committee, and document, we first filter it into a sample S^* consisting only of decisions valid for projection, i.e., decisions with at least one individually recorded vote. Then from V we produce a filtered $m \times n$ vote matrix v , where m is the number of voters with participation in at least one of the sessions from decisions in S^* and n is the number of decisions in S^* . We then project v into a $m \times 2$ matrix v^p using PCA.

Although there are more sophisticated techniques for displaying neighborhood relationships [EMK*19], we chose PCA for this application due to our requirements – domain experts wanted to be able to obtain explanations for the spatial positioning of voters, and the ability to correlate it with individual decisions (**R3**, **T6**). As PCA is a linear projection technique, its coefficients show the exact influence of each decision in the process, and how correlated they are to one another in the voting records. The matrix v^p is then displayed as a scatterplot, assigning a black dot to each voter. Alongside the projection, PCA covariance ratios are shown for each of the two dimensions, informing how much of the sample's behavior in the original space is actually represented by the PCA projection.

If a comparison sample was also selected, users can choose to display its projection alongside the first one (**R1**, **T2**, **T3**). PCA projections from different sets of features have incompatible spaces and cannot be directly compared, as they are generated from different transformations over different sets of dimensions. We use Procrustes analysis to apply a second linear transformation to the comparison sample to match the main sample as best as possible, and display the projections from both samples. The comparison sample

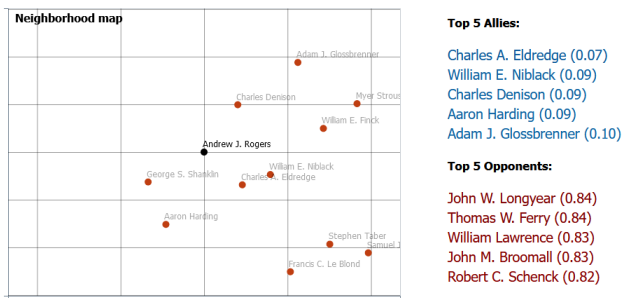


Figure 8: Relationship map and allies/opponents list. The relationship map is centered on the selected person and shows people with similar voting behavior. The allies/opponents list shows the five people with the most similar/dissimilar voting patterns.

many conventions do not currently have keywords implemented. To fill gaps when necessary, we use Yake [CMP*20] to detect the most commonly occurring words in document descriptions from a committee. A dictionary of stop-words was defined alongside domain experts to generate more relevant results. Figure 7 shows the keyword cloud.

Alongside the keyword cloud, we can also display a lollipop chart of the inverted Yake scores of the top 20 words, as an additional aid to illustrate the impact and frequency of each word in the committee's documents.

4.2.3. Relationships

This panel shows an overview of P 's relationships with other people in the committee (T2, T4). It consists of a relationship map (a 2D projection of the voting matrix v with all decisions from the selected committee, centered on P) and a list of their most frequent allies and opponents in voting. Differently from the projection in the dataset visualizer, we use a non-linear projection technique for the relationship map: feedback received during our preliminary studies indicated that, while not always possible, having the neighboring people in the projection match the list of allies was important to users. To prioritize this aspect, we generated the relationship map using t-SNE. Empirical tests were also conducted using Uniform Manifold Approximation and Projection (UMAP) [MHM18] with perceptually similar results for local relationships.

The allies and opponents list is obtained directly from ordering pairwise distance calculations between people in v . We use squared Euclidean as a distance metric, as to further emphasize distances between people with opposing votes while still keeping note of abstentions and uncertain votes. Each person X on the list has their distance to P normalized as

$$d_{px}^* = \frac{d_{px}}{d_{max} \cdot m_{part}} \quad (1)$$

where d_{px} is the distance between P and X , d_{max} is the maximum possible distance in a single vote (in this case, 4), and m_{part} is the total number of decisions P had a vote cast (including abstentions). Allies in this list are colored in tones of blue, while opponents are colored in tones of red. Clicking on a name will open a link to that person's *person visualizer* page. The Relationships panel is shown in Figure 8.

4.2.4. Events

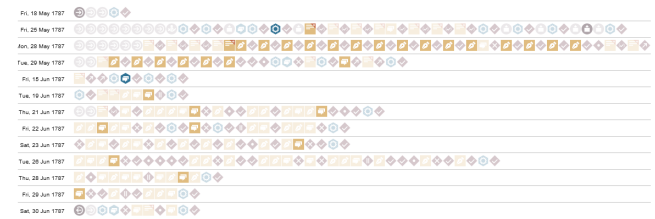


Figure 9: Event participation panel. Sessions are arranged vertically by date, and for each session, a sequence of event icons is shown. Each icon is directly linked to event description pages in the platform. Events with participation from the selected person P are drawn in full opacity, while all others are drawn in half opacity.

The Events panel offers a person-focused timeline of events in the selected committee (T3, T4), and consists of a vertical list of sessions; for every session, event icons are displayed horizontally indicating the sequence of events that occurred during that session, alongside the session date. Participation from P is given by opacity – events that include P are shown in full opacity, while other events are shown in half opacity. The list allows for hiding sessions without participation from P .

While this representation has certain scalability limitations, we opted for this approach as opposed to a more concise representation due to users' existing familiarity with following event icon sequences icons using other tools in the Quill platform (R4) and to its direct connection with events (T6). For interactivity, hovering the cursor over an icon will show event details, and clicking the icon will open its description page (T5). The Event panel is shown in Figure 9.

4.2.5. Votes

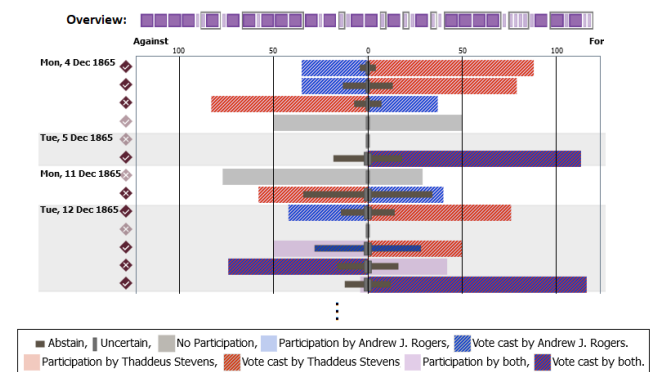


Figure 10: Vote breakdown panel. An overview shows all decisions in the committee. Each session displays its decisions as a centered bar chart, with different symbols for different vote types (see legend). The selected person P is shown in blue, while a comparison person X is shown in red. The image portrays the initial section of a longer chart. An overview representation is shown on the left side.

The voting panel contains a breakdown of P 's votes in all decisions in the committee (R1, R2, T3, T4, T6). This panel consists of an enhanced centered bar chart, with decisions ordered vertically

by time as rows and grouped into sessions with their date displayed. For all rows, there is a central vertical axis representing zero on a horizontal scale. The right side of this axis contains votes *for*, while the left side contains votes *against*. Bars are drawn to each side with a length corresponding to the total amount of votes for each option in that decision. A darker, thinner bar is laid on top of the previous one to represent abstention. This bar has an equal length on both sides, indicating the total amount of abstentions. Finally, a rectangle is used to indicate the number of uncertain votes, positioned at the appropriate point on both sides of the horizontal scale. *P*'s actions are shown in blue; if they were present in a decision, its vote bars are shown as light blue; otherwise, they are shown as gray. *P*'s taken action is shown in a saturated blue pattern; the side they voted for, abstention bar, or uncertainty rectangle is colored this way. For each row, an icon representing the decision outcome and its name, with a link to a decision description page (T5), is also displayed.

This view may become extensive depending on the number of decisions in a committee. To attenuate this issue, we employ two methods: the first is to allow users to hide all decisions that do not contain voting records, as some committees can contain a large number of them. The second is to provide a compact overview as a guide, showing all decisions as boxes contained in larger rectangles representing sessions. Decisions with participation are shown in a larger size. Clicking on a box will highlight the corresponding decision in the panel. Hovering the cursor over a box or a vote bar will show a textual description of the decision and its votes.

Users also have the option of selecting another person on the committee to compare votes with (T1, T2). If selected, a second person *X* will be added to the vote breakdown: their actions are highlighted in a similar way to *P*, but in red. If both *P* and *X* have participated in a vote or taken the same action, it will be colored purple. Figure 10 shows an example of the voting panel.

5. Usage Scenario

Figure 1 shows an example of workflow using the proposed interface. A History researcher is studying the records for the United States 14th Amendment and The Civil Rights Act of 1866 convention on Quill as part of a collection relating to the complex and extended discussions that resulted in the Amendment of the US Constitution in the wake of the Civil War. The user then wants to investigate how representatives voted in amendment decisions in this convention, and opens the Projection Dashboard interface.

The first thing they notice is that decisions are grouped into two clusters over time: from January to March of 1866, and from April to May of 1866. From using other exploratory tools in Quill, the user knows that the gap between these clusters is due to the way in which Congress managed debate, and in particular the legislative progress of the particular bills and resolutions that are part of this collection. The researcher then decides to investigate if voting behavior has changed from one period to the other (Figure 1-A).

After selecting the desired time windows, the researcher then projects voting data from the House of Representatives committee. When observing the projection, they notice that the two points were relatively distant and became closer together in the second time sample (Figure 1-B). These points refer to Andrew J. Rogers and

Henry Grider. The researcher can then look at the person records for these representatives in the Person Visualize interface.

When visualizing records for Henry Grider, the researcher is shown committee involvement, keywords, relationship neighborhood, and event involvement, displaying highlights of that representative's activity in the convention. By selecting Andrew J. Rogers as a comparison target in the breakdown of the votes, they notice that there is significantly more agreement between the two in the time period from the second sample (Figure 1-C), which is probably the reason why the two dots have moved closer together. By looking at descriptions for these decisions in other Quill interfaces, the researcher notices that there is a difference in the chances of alignment when considering how to treat the southern states after the war and when considering specific changes in constitutional language. Further investigation of the dataset can then determine whether greater consensus was achieved over time on more contentious issues and whether a qualitative judgment was made on how progress was achieved.

6. Evaluation

During the iterative design process of *dataset visualizer* and *person visualizer*, we constantly evaluated the usefulness and usability of the current prototype with domain experts and made continuous improvements in an agile manner. We were in close collaboration with a lead domain expert who has specialist knowledge of constitutional conventions and also understands the entire architecture of Quill. We kept timely communication to obtain immediate feedback from him. This evaluation was carried out throughout the design process in a less formal but immediate way. Upon receiving feedback, features were improved through rapid prototyping until they were ready to be tested by more domain users. We tested milestones with a group of experts who worked with Quill data and interfaces to prevent unrepresentative feedback from a single expert from leading us astray. The evaluation with multi-experts went through the following 3 stages: 1) an inception workshop, 2) an initial survey, and 3) the expert interviews. We discuss the methodology and results of each stage in the remainder of this section.

6.1. Stage 1: Inception Workshop

An inception workshop was conducted with 7 domain experts from the Quill Team on the visual representations of PCA results, which became the key components for subsequent developments of the *dataset visualizer*. These designs were found to be interesting for their ability to extract features from vote and decision data and to be practical from discussions with the lead domain expert.

6.1.1. Methodology

The workshop was conducted with three main intentions: 1) to assess the understanding and acceptance of PCA projection and its related visualizations by domain users, 2) to collect feedback on the current prototype, and 3) to obtain the overall demands on exploring and comparing vote records. The workshop lasted around 3 hours and was carried out through 4 stages: 1) introduction to the PCA background and how to read the related visualizations; 2) demonstration of the prototype in use; 3) free exploration of the

prototype by the domain experts; 4) collection of feedback and requirements through group discussion and a short post-survey.

6.1.2. Results

Through this workshop, we came to the conclusion that most domain experts had no prior knowledge about PCA but did not resent being exposed to PCA-related visualizations. From the demonstration of the PCA-related visualizations in use, participants recognized their potential to provide insights on voting behaviors and voter relationships. Although confused at the beginning, they were able to understand what the prototype's visual representations were trying to convey after the tutorial and demonstration. The comparison of voting behaviors and decisions between different time spans offered by the prototype benefited domain experts in their exploration of the voting data. However, more explanation was required to improve the ease of understanding the tool. In addition, they suggested features that they would like to see on the current prototype, which can be concluded as new raised domain requirements as follows: 1) More comprehensive filtering capabilities; 2) Highlighting important events in time; 3) Adding detailed vote information; 4) Improving the relevance to other Quill interfaces, such as links to the related document page.

6.2. Stage 2: Initial Survey

The initial survey was conducted after several design iterations and had initial *dataset visualizer* and *person visualizer* deployed. The two interfaces were presented to 10 domain experts from the Quill Oxford Team for testing, and a questionnaire was designed for gathering feedback.

6.2.1. Methodology

The survey was designed with intention to: 1) collect reactions to different components in *person visualizer* and 2) obtain feedback and demands from domain experts on exploring person-typed data. Given that users of Quill interfaces are likely to be domain experts with a background in legislation history, we chose the Expert Sampling [EMA¹⁶, EB17] method to recruit participants and distributed the survey via Qualtrics [Qua05]. The questionnaire had 3 sections: 1) Participants' Backgrounds included demographic questions covering age group, occupation field, education level, and domain expert identification questions; 2) User Experience and Design Preferences included questions regarding the effectiveness and usefulness (Likert scale questions) and general feelings (open-ended questions) of four views on *person visualizer*, which were designed for visualizing keywords, voters' relationships, events timeline, and voting records, respectively; and 3) General Feedback and Demands included open-ended questions regarding how helpful the current visual representations and tools are in addressing domain problems and where future enhancements can be made.

6.2.2. Results

Responses to background questions showed that 90% of the respondents work in legislation history-related fields and all participants had experience with visualization tools.

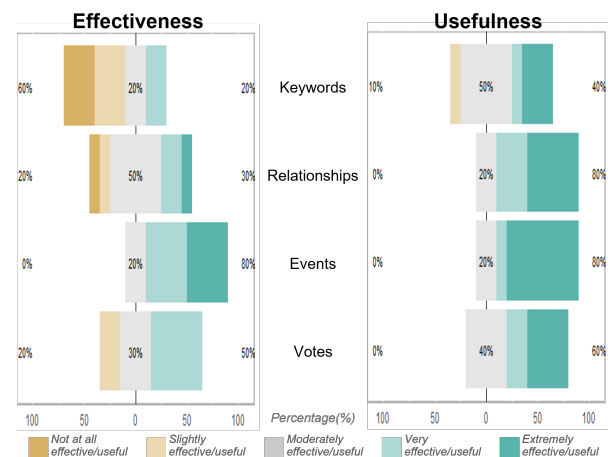


Figure 11: Statistics of the Likert scale question responses regarding effectiveness (left) and usefulness (right).

We received relatively positive feedback from the User Experience and Design Preferences section. The results of the Likert scale questions are shown in Figure 11. Overall, respondents awarded the four *person visualizer* views good rates for their usefulness. Only 10% of the respondents considered the Keywords view as slightly useful, and no participant voted for the below moderate levels of usefulness for the other three views. Feedback on effectiveness relative to the usefulness of each view was diminished to varying degrees. Results indicated that our design is in line with the needs of the domain experts, but the current prototype required future modifications and enhancements to be more effective.

According to the distribution of the ratings, the keyword view received a lower rating compared to the others, especially in its effectiveness. Combining responses to open-ended questions, we concluded that the main aspect reducing effectiveness was imprecise and unrepresentative keywords; in contrast, respondents gave the interactive visual representation of the wordcloud and lollipop chart very positive reviews. *“The principle of a wordcloud for frequently used words in relation to a particular congressman is appealing. It could be a helpful overview for the topics ... However, the words are so general that doesn't have the desired effect.”* The stop-word list was further expanded with terms considered to be important by YAKE but not helpful to domain experts to make the YAKE-derived keywords more representative. It is worth noting that YAKE-derived keywords were used as a temporary measure to fill the gaps in keywords supported by the Quill platform.

The neighborhood map and voters' relationships view received praise for its ability to reveal similarities and differences in voting behavior. Distance estimations provided domain experts with a new way of looking at data, which also resulted in difficulties: requests for more explanatory features were mentioned frequently, and there was confusion among certain users on how names in the list and their position on the map did not always match their expectations.

Visualization views of events and votes received positive feedback. Participants' suggestions centered on improving the visualization's aesthetic appeal and clarity, and adding explanatory features. These two views followed Quill's original visual designs for

User	Duration	Role
P1	3.5 years	Documentary Editor <i>primary data inputting and editing</i>
P2	2 years	Documentary Editor <i>editorial inconsistencies checking</i>
P3	1.5 years	Documentary Editor <i>document and timeline checking</i>
P4	1.5 years	Documentary Editor <i>document transcriptions inputting</i>
P5	5 years	Quill Research Assistant <i>teaching and educational impact assessing</i>

Table 1: Summary of the interview participants' background.

event lists and voting records by expanding and improving on them. The respondents were already comfortable with operating and interpreting them. The familiarity with the design directly resulted in better scores on the effectiveness of the event and vote views compared to the newly added views for keywords and relationships.

We summarized points for improvement based on the feedback from the General Feedback and Demands section: 1) improve the *person visualizer*'s relevancy and connection with other Quill interfaces, such as linking the vote data to the relevant documents; 2) enhance the general interpretability by adding more instructions and explanations; and 3) offer more contextual and textual information related to the voters, i.e. add close-reading functionalities on top of the current prototype. The respondents were generally complimentary of the current prototype and listed several areas in which it may improve their workflow, notably in terms of overview and data exploration, decision-making, and comparison.

6.3. Stage 3: Expert Interview

The two interfaces were adjusted and improved based on the feedback from the previous workshop and survey. The usefulness and usability were once again evaluated by conducting expert interviews in conjunction with the think-aloud protocol. We observe the interaction and specific use of the tool by domain experts and conduct a qualitative analysis based on interview transcripts.

6.3.1. Methodology

We interviewed 5 domain experts in person or remotely depending on their location. Each interview lasted for 60 to 90 minutes.

The first four participants are from the Quill Oxford Team. They are users of the Quill system and are in charge of inputting, editing, and proofreading different types of Quill data. Participant P5 is from Utah Valley University. She focuses on exploring Quill's impact on law and education. She provides training for students on how to use the platform and model in Quill and works with K-12 teachers to evaluate its educational impact. In addition, she uses the visualizations in Quill for research purposes. The participants' background is described in Table 1.

The procedure and the duration of activities are listed in Table 2. Interviews were conducted with introductory questions and answers to understand participants' backgrounds and walk them through the toolset. Participants were then asked to screen-share and practice activities we prepared, thinking aloud during the entire process. During think-aloud sessions, they were first encouraged to

Order of Procedure	Activities	Duration
Preliminary Preparation	1) Introductory questioning 2) Tool walkthrough	15 min
Think-aloud Evaluation	1) Test in the wild 2) Test via predefined tasks	40 min
Reflective Discussions	1) Reflection on the tool 2) Future direction	20 min

Table 2: Evaluation Procedure and Duration.

interact with the tool freely to get familiar with it in practice. We also learned about how they used the tool without directional tasks from this Test in the Wild session. The participants were then asked to perform two activities with the tools, consisting of multiple designated tasks. Tasks were derived from the pre-collected user demands and domain requirements. Both activities aimed to test the usefulness of the proposed interfaces. The first activity focused on the *dataset visualizer*, while the second aimed at the *person visualizer*. The ending phase was a reflective discussion on the usability of the tool and how helpful it was in meeting domain requirements. The future development of the tool was the subject of discussion.

6.3.2. Results

We consolidate feedback from the expert interviews as follows:

Reflection on Usability. Referring to the definition and five usability indicators proposed by Nielsen [Nie94], the feedback of the participants and their performance indicated good levels for *satisfaction*, *efficiency*, and *low-error*. *Memorability* was not the main focus of the evaluation. Regarding *learnability*, the participants were able to quickly understand and use most of the features, while two features were frequently queried for lack of clarity: 1) two participants (P1 & P2) got confused by the dark and light gray colors used in the *Timeline* (Figure 1 DV1). Due to the presence of anonymous labels in the Quill data, we decided to leave the existing design (to differentiate whether decisions without voting records have anonymous labels with different shades of gray) to maintain consistency with the original data. To increase readability, we added more detailed legends and textual descriptions on the explanation page; 2) participants found the *PCA Chart* (Figure 1 DV3) somewhat difficult to understand at first. After understanding that the information conveyed from the chart is the extent to which the two main components (x and y) explain the data, the participants' confidence in the PCA results increased from their feedback. To further enhance the first impression of the PCA Chart, the default data being visualized was changed to 'x+y', which was recognized by participants as the most intuitive and understandable of all the choices.

Reflection on Usefulness. The usefulness of a tool is assessed by how well it contributes to the achievement of domain requirements. From the interview transcripts, all participants found the toolset very effective and useful in providing insights into voting interactions and people's voting behavior and roles (**R1 & R2**). The *Timeline* was praised for its contribution to data exploration: "[...] you may need a way to explore [the dataset] to find out moments of change, that is where the Timeline is most helpful." The *scatter plot* (Figure 1 DV2) was the most acclaimed in gaining insight into voting behavior changes across time: "I think it is the most effective when comparing the changing patterns and voting behavior, especially if you're comparing two people that were previously identical

and now changed.” The **Relationships view** combined with the **Vote Breakdown** (Figure 1 PV3 & PV5) was mentioned to be useful in reflecting the relationship between voters based on their voting behaviors. They can also be used as visual evidence for historical hypotheses on the relationship between members of a committee. The **Event List** (Figure 1 PV4) was shown as an efficient way to show a person’s participation in events. By inheriting the original Quill icon designs, participants were able to quickly identify the different types of events in which the person was involved, which meet the requirement of adhering to the existing Quill platform (**R4**). The newly added filtering function increased the effectiveness of obtaining information from the Event List and the Vote breakdown: “I love that you can remove the sessions that don’t have their participation, especially because it shortens the loading time, and made things more efficient.” Users are satisfied with the linkage between data and visualization in both interfaces (**R3**): “The connections from voting and decision data to person and documents are really useful in terms of adding information, proofreading, and checking.”

Summary. Through expert interview sessions, the complete functionality of *dataset visualizer* and *person visualizer* was verified. The panels and views in both interfaces performed their tasks properly, and the combination of the two interfaces worked stably and gave the corresponding results. In conclusion, the usability and effectiveness of both interfaces were validated. In terms of usability, the interface is generally easy to understand and use. The connection between the two interfaces is smooth, and the data presented in the visualization panels is well connected to the other data interpretation interfaces of the Quill framework. Some users mentioned difficulty in interpreting parts of the PCA Chart, and we believe it to be mainly due to the lack of understanding of how the principal components are derived. However, this did not affect users’ positive opinion of the usefulness of the PCA Projection for extracting features from the data. Based on the evaluation results, we added more textual descriptions to the original *dataset visualizer* interface to improve readability. Regarding usefulness, all the domain requirements were satisfied. The PCA Projection, the Relationship Neighborhood, and the Vote Breakdown were deemed useful in offering fresh perspectives toward voting data and acting as evidence to support hypotheses. Timeline and the Involvement Donuts Chart were considered to be good overview visualizations. Experts in the field who input document data to Quill and study Quill’s impact on education thought the Keyword Cloud was particularly relevant to their line of work, and they expected it to be extended to other Quill interfaces that deal with documents.

7. Discussion & Takeaway

After analyzing our evaluation results, as well as reflecting on the design and development of the toolset and our interactions with domain experts, we can raise the following conclusions:

Technique Complexity. During our studies, we noticed a difference in feedback between users who were familiar with PCA and users who were not; while the former would be interested in increasing the level of reported statistical information, coefficients, and understanding the composition of each dimension, the latter would prefer the visualization to be kept more minimalist to avoid an increase in the complexity of visualization not so familiar to

them. Our refinement changes were then focused on interpretability aiming to bridge this gap, i.e., providing the information more experienced users wanted but in ways that could favor less experienced users.

Familiarity Tradeoff. While adhering to representations that were already familiar to Quill users whenever possible was a requirement, we observed that users tended to also prefer the visualizations that employed familiar abstractions, such as the Events panel. Even though not as scalable, sequences of events are a familiar way of representing data in Quill and users would often default to them. While there is a tradeoff between efficacy and familiarity, there are often ways to present an efficient visual abstraction with a familiar package.

The limitations in our toolset are, for the most part, tied to existing constraints in the Quill platform or related to previous user experience. Components that are related to the length of the dataset, such as the event and voting panels in the *person visualizer*, may suffer from scalability issues when investigating particularly long or eventful conventions. In the same manner, observing an entire timeline at once in the *dataset visualizer* may not be feasible for a large dataset/low screen resolution combination. These components do offer collapsible elements and can still be used when zoomed in. Further summarization, such as semantically aggregating events and decisions for each type of visualization, is considered future work.

The PCA scatterplot presents some limitations: due to its nature, the PCA projection may not be able to properly convey certain complex relationships in voting data, and the changes observed between two samples are not a precise representation as we would wish it to be due to the differences in the Data spaces, which at times are not fully compatible. We tried to compensate by leveraging elements in the visualization, such as the covariance ratios displayed, which are also used as “signposts” to keep users aware of these limitations as much as possible. The PCA projection is not intended to be an end in itself, but rather a discovery tool to assist qualitative research as part of a broader research platform.

8. Conclusion

In this paper, we present a set of visualization tools designed to help navigate and interact with voting records and historical documents. We discuss how we navigated designing these tools to meet a set of user requirements while also ensuring that they fit seamlessly within a larger platform that already had a legacy set of expectations and visual language associated with it. Our design process was informed by constant communication with a domain expert, as well as multiple evaluation sessions with his team.

Overall the designed tools received positive feedback from users and we obtained valuable insight for future work. As for future work directions, we have uncovered several opportunities for building more encompassing visualization tools including the addition of features to further expand the exploration space, such as options to support a deeper investigation of interpersonal relationships in document descriptions and speeches. The tools designed here can also be a starting point for visualization of collaborative writing and voting in different domains, such as social media platforms.

Acknowledgements

This work was supported in part by the Engineering and Physical Sciences Research Council (EP/V028871/1) and the King's-China Scholarship Council PhD Scholarship programme (K-CSC).

References

- [ALH*15] ANDRIS C., LEE D., HAMILTON M. J., MARTINO M., GUNNING C. E., SELDEN J. A.: The Rise of Partisanship and Super-Cooperators in the U.S. House of Representatives. *PLOS ONE* 10, 4 (04 2015), 1–14. [2](#)
- [ARLC*13] ABDUL-RAHMAN A., LEIN J., COLES K., MAGUIRE E., MEYER M., WYNNE M., JOHNSON C. R., TREFETHEN A., CHEN M.: Rule-based visual mappings – with a case study on poetry visualization. *Computer Graphics Forum* 32, 3pt4 (2013), 381–390. [2](#)
- [BEAC*18] BRADLEY A. J., EL-ASSADY M., COLES K., ALEXANDER E., CHEN M., COLLINS C., JÄNICKE S., WRISLEY D. J.: Visualization and the Digital Humanities: Moving toward stronger collaborations. *IEEE Computer Graphics and Applications* 38, 6 (2018), 26–38. [2](#)
- [BGCC16] BRIGADIR I., GREENE D., CROSS J. P., CUNNINGHAM P.: Dimensionality reduction and visualisation tools for voting record. In *24th Irish Conference on Artificial Intelligence and Cognitive Science (AICS'16)*, University College Dublin, Ireland, 20–21 September 2016 (2016), CEUR Workshop Proceedings. [2](#)
- [BL16] BATEMAN D. A., LAPINSKI J.: Ideal points and american political development: Beyond dw-nominate. *Studies in American Political Development* 30, 2 (2016), 147–171. [2](#)
- [BSS20] BENITO-SANTOS A., SÁNCHEZ R. T.: A data-driven introduction to authors, readings, and techniques in visualization for the Digital Humanities. *IEEE Computer Graphics and Applications* 40, 3 (2020), 45–57. [2](#)
- [CARM17] COLE N., ABDUL-RAHMAN A., MALLON G.: Quill: A framework for constructing negotiated texts – with a case study on the US Constitutional Convention of 1787. In *2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL)* (2017), IEEE, pp. 1–10. [1](#), [2](#), [3](#)
- [CMP*20] CAMPOS R., MANGARAVITE V., PASQUALI A., JORGE A., NUNES C., JATOWT A.: Yake! keyword extraction from single documents using multiple local features. *Information Sciences* 509 (2020), 257–289. [5](#), [6](#)
- [D3] D3.js: Data-Driven Documents. URL: <https://d3js.org/>. [3](#)
- [EB17] ETIKAN I., BALA K.: Sampling and sampling methods. *Biometrics & Biostatistics International Journal* 5, 6 (2017), 00149. [8](#)
- [EMA*16] ETIKAN I., MUSA S. A., ALKASSIM R. S., ET AL.: Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics* 5, 1 (2016), 1–4. [8](#)
- [EMK*19] ESPADOTO M., MARTINS R. M., KERREN A., HIRATA N. S., TELEA A. C.: Toward a quantitative survey of dimension reduction techniques. *IEEE Transactions on Visualization and Computer Graphics* 27, 3 (2019), 2153–2173. [4](#)
- [FBM16] FULDA J., BREHMER M., MUNZNER T.: Timelinecurator: Interactive authoring of visual timelines from unstructured text. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 300–309. [2](#)
- [FSR*21] FILIPOV V., SCHETINGER V., RAMINGER K., SOURSOS N., ZAPKE S., MIKSCH S.: Gone full circle: A radial approach to visualize event-based networks in digital humanities. *Visual Informatics* 5, 1 (2021), 45–60. [2](#)
- [GGJ*21] GUO Y., GUO S., JIN Z., KAUL S., GOTZ D., CAO N.: Survey on visual analysis of event sequence data. *IEEE Transactions on Visualization and Computer Graphics* 28, 12 (2021), 5091–5112. [2](#)
- [HAC*15] HINRICHS U., ALEX B., CLIFFORD J., WATSON A., QUIGLEY A., KLEIN E., COATES C. M.: Trading consequences: A case study of combining text mining and visualization to facilitate document exploration. *Digital Scholarship in the Humanities* 30 (10 2015). [2](#)
- [HLL14] HEIMERL F., LOHMANN S., LANGE S., ERTL T.: Word cloud explorer: Text analytics based on word clouds. In *2014 47th Hawaii International Conference on System Sciences* (2014), IEEE, pp. 1833–1842. [5](#)
- [IA12] ITOH M., AKAISHI M.: Visualization for changes in relationships between historical figures in chronicles. In *2012 16th International Conference on Information Visualisation* (2012), pp. 283–290. [2](#)
- [JFCS17] JÄNICKE S., FRANZINI G., CHEEMA M. F., SCHEUERMANN G.: Visual text analysis in digital humanities. In *Computer Graphics Forum* (2017), vol. 36, Wiley Online Library, pp. 226–250. [2](#)
- [KK15] KUCHER K., KERREN A.: Text visualization techniques: Taxonomy, visual survey, and community insights. In *2015 IEEE Pacific Visualization Symposium (PacificVis)* (2015), pp. 117–121. [2](#)
- [KOTM13] KIMURA F., OSAKI T., TEZUKA T., MAEDA A.: Visualization of relationships among historical persons from Japanese historical documents. *Literary and Linguistic Computing* 28, 2 (02 2013), 271–278. [2](#)
- [MCL17] MUN S., CHOI G., LEE S., LEE K.: Visual analysis for voting relationships in Joseon Dynasty, Korea. In *10th International Conference on ICT, Society and Human Beings* (2017), pp. 111–118. [2](#)
- [MHM18] MCINNES L., HEALY J., MELVILLE J.: Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426* (2018). [6](#)
- [MLCM16] MCCURDY N., LEIN J., COLES K., MEYER M.: Poemage: Visualizing the sonic topology of a poem. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 439–448. [2](#)
- [MMM22] MÉNDEZ G. G., MORENO O., MENDOZA P.: Legislatio: A visualization tool for legislative roll-call vote data. In *15th International Symposium on Visual Information Communication and Interaction* (2022), pp. 1–8. [2](#)
- [Nie94] NIELSEN J.: *Usability engineering*. Morgan Kaufmann, 1994. [9](#)
- [NMdSSDSF18] NUNES MONI DA SILVA R., SPRITZER A., DAL SASSO FREITAS C.: Visualization of roll call data for supporting analyses of political profiles. In *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)* (2018), pp. 150–157. [1](#), [2](#)
- [NMM*14] NOVAK J., MICHEEL I., MELENHORST M., WIENEKE L., DÜRING M., MORÓN J. G., PASINI C., TAGLIASACCHI M., FRATER-NALI P.: Histogram – a visualization tool for collaborative analysis of networks from historical social multimedia collections. In *2014 18th International Conference on Information Visualisation* (2014), pp. 241–250. [2](#)
- [OIK*11] OSAKI T., ITSUBO S., KIMURA F., TEZUKA T., MAEDA A.: Visualization of relationships among historical persons using locational information. In *Web and Wireless Geographical Information Systems* (Berlin, Heidelberg, 2011), Tanaka K., Fröhlich P., Kim K.-S., (Eds.), Springer Berlin Heidelberg, pp. 230–239. [2](#)
- [PR99] POOLE K., ROSENTHAL H.: D-nominate after 10 years: A comparative update. In *Annual Meeting of the American Political Science Association* (1999). [2](#)
- [PRDC20] POSSIERI C., RAVAZZI C., DABBENE F., CALAFIORE G. C.: A new metric for understanding hidden political influences from voting records. *PLOS ONE* 15, 9 (09 2020), 1–17. [2](#)
- [Qua05] QUALTRICS., 2005. Accessed: 2022-11-21. URL: <https://www.qualtrics.com>. [8](#)
- [SM07] SLEZ A., MARTIN J. L.: Political action and party formation in the United States Constitutional Convention. *American Sociological Review* 72, 1 (2007), 42–67. [2](#)
- [VdMH08] VAN DER MAATEN L., HINTON G.: Visualizing data using t-sne. *Journal of Machine Learning Research* 9, 11 (2008). [2](#)