

Visualization of Roll Call Data for Supporting Analyses of Political Profiles

Rodrigo Nunes Moni da Silva
Graduate Program in Computer Science
Federal Univ. of Rio Grande do Sul
Porto Alegre, RS, Brazil

Andre Suslik Spritzer
Institute of Informatics
Federal Univ. of Rio Grande do Sul
Porto Alegre, RS, Brazil

Carla M.D.S Freitas
Institute of Informatics
Federal Univ. of Rio Grande do Sul
Porto Alegre, RS, Brazil

Abstract—In this paper, we propose a web-based application where the user can instantiate multiple, coordinated panels for exploring data concerning the votes of representatives in Brazil’s lower legislative house (the Chamber of Deputies). Open data about roll calls made available by the Chamber allowed us to build a set of interactive visualizations to let users explore deputies’ votes and build an understanding of their political profiles. Based on the set of roll call voting results from 1991 to 2016, our application displays the political behaviour of parties in a timeline from which users can select periods and instantiate panels showing the political spectrum of deputies using different methods of dimensionality reduction. Deputies can be separated in clusters based on their position in the political spectrum, and other panels can be instantiated showing details about each cluster. Users can select parts of the timeline and simultaneously analyze the behavior of parties and one or more deputies. Roll calls are represented as a combination of heatmaps and histograms. We illustrate the use of the different visualization techniques in a case study on party cohesiveness over time.

I. INTRODUCTION

Several open data portals provide government information for both the general public and application developers (see, for example, [1]–[4]). These portals let citizens obtain information on specific deputies, senators, and propositions through simple form-based queries. Developers and more advanced users can download datasets such as the detailed federal budget and the voting history of representatives in the National Congress.

Concerning the understanding of legislators’ behavior, however, there is a considerable gap between the works developed by political scientists and those aimed at informing the general public. Political scientists over the past 25 years have used the analysis of roll call voting data for developing empirical models representing the spatial theory of voting [5]. Such models are also known as *ideal point estimators* because they infer the location of a legislator in an abstract space from their roll call votes. Usually represented as a scatterplot, this space (or *political spectrum*) can be as broad as the whole set of roll calls from a legislature or as specific as a set of roll calls related to some particular issue. They have been used in the study of the US Congress as well as in the analysis of legislative behavior more generally [6]–[9]. There has been some effort in communicating findings resulting from these methods to the general public by providing visualizations of the political spectrum of legislators along with explanations or narratives [10]–[13]. However, these visualizations are either

static, merely informative, or allow for little exploration because they rely on showing the political spectrum of legislators and not the roll call data set.

This scenario was improved by CivisAnalysis [14], a web-based tool for the visualization of roll call data of the Brazilian Chamber of Deputies. The main visualizations provided are two kinds of political spectrum: one showing the deputies and the other displaying the roll calls. When legislators are closely located in the political spectrum of deputies, it means that their voting patterns are similar; roll calls that are close in the political spectrum of roll calls mean that they have received similar votes from the same deputies. CivisAnalysis uses a data set of voting records of six legislatures (24 years, 513 deputies), comprising 2,458 roll calls (853,952 votes), as well as the information of six presidential elections, including election results and alliances made for the elections. It lets users observe the distribution of deputies on a political spectrum and compare the position of a specific legislator in relation to the others. Still, it has limitations on how much users can learn about the political profiles of deputies and parties. It is not possible, for example, to compare a deputy’s voting behavior to that of other deputies or the deputy’s party’s policies across different periods of time.

In this paper, we propose a redesigned CivisAnalysis that addresses the original’s limitations by enhancing it with new visualization techniques that let users gain insight into the changing profiles of parties and deputies through the exploration of the entire roll call voting dataset from 1991 to 2016, which includes seven legislatures and nine presidential terms. Visualizations are accessed through a user interface based on multiple coordinated panels that are organized hierarchically.

The main contributions of our work are:

- A robust system that allows users to create visualizations of political data without the original time period limitation. Users can create several coordinated panels, each displaying a (possibly different) visualization.
- A set of interactive techniques that lets users compare deputy and party voting patterns across time.

In the next section, we provide a brief review of the related work. Our proposal is introduced in section III and described in section IV. Section V demonstrates its use in a study of parties cohesiveness along the years. Finally, in section VI we summarize our work and comment on future work.

II. RELATED WORK

In their analyses of roll call data, many political scientists have used spatial models to depict the position of legislators in a political space [5] and to measure and compare the ideology and heterogeneity of political parties [15]. The NOMINATE family of algorithms [16] are widely used in roll call analysis and compare favorably to more modern algorithms [17]. In contrast to the complexity of these methods, fast dimensionality reduction techniques such as Principal Component Analysis (PCA) [18], even if less accurate for roll call analysis, have also been used for obtaining political spectra of deputies and parties [19].

Using NOMINATE scores, Connect 2 Congress (C2C) [10] creates a two-dimensional political spectrum of the U.S. Congress for arbitrary time frames within a two-year period (2007-2008). Users can inspect roll call data as tabular data and filter and highlight representatives by name, state, party, religion, and gender. The time frame can be dynamically modified, resulting in an animation where representatives are continuously repositioned on the spectrum according to their behavior. As it only covers two years, C2C does not have inter-legislature analysis. Social Action [20] takes an alternative approach for visualizing the behavior of U.S. senators: it builds a social network to represent the correlation of votes between senators and displays it using a force-directed layout algorithm, which users can explore by interactively applying filters and statistical tools to uncover patterns of voting groups at single points in time.

As for the Brazilian Congress, Marino [11], Basômetro [12] and Radar Parlamentar [13] create two-dimensional spectra of deputies. Marino and Basômetro display the political spectrum of Deputies as a scatterplot, where the diagonal divides the government coalition and the opposition. Radar Parlamentar shows the political spectrum in a radial format and only allows for the analysis of distances, as party positions change significantly over short periods of time. Parties appear in clusters, with the radius of the party's corresponding circle being proportional to its number of deputies, and its position being the average position of its deputies. Parties can be expanded to show the original positions of their deputies. Both Marino's work and Radar Parlamentar are based on PCA.

To address the temporal variation of political positions in a single visualization, some techniques provide a long-term overview using political timelines, which are compact visualizations of political trajectories of individuals or parties over time. Friggeri and Fleury's visualization [21] shows the paths of U.S. senators through agreement groups for eight legislatures using a custom clustering algorithm that is applied to roll call data. It is an interactive, web-based application that lets users select a single senator to track his or her cluster alignment.

Finally, CivisAnalysis [14] directly inspired our work due to the possibilities provided by its highly interactive and flexible features that let users explore roll call data in diverse ways. Besides the political spectra of deputies and roll calls,

CivisAnalysis provides: (1) a timeline displaying the relative position of parties (computed from the deputies' political spectrum) along the years, (2) an infographic showing the distribution of seats to political parties in the Chamber of Deputies, (3) a map for filtering deputies by state, (4) a text panel for displaying the proposals selected in the roll call political spectrum, and (5) a search panel that allows searching for one or more deputies or roll calls. It lets users choose between two dimensionality reduction methods for the building of the political spectra: PCA (the default one) and t-Distributed Stochastic Neighbor Embedding (t-SNE) [22].

We based our work on CivisAnalysis, but only used the original code for the Timeline, the Chamber Infographic, and the computation of the deputies' Political Spectrum (both PCA and t-SNE implementations). We named our tool CivisAnalysis 2.0, with the agreement of the previous authors [14].

III. REQUIREMENTS AND OVERVIEW

The design of CivisAnalysis 2.0 was based on analyses that the users might want to perform on roll call data from the Brazilian Chamber of Deputies. Although Borja and Freitas [14] based their work on what they called "Citizens Tasks," we adopted "Visual Analytics Questions" (VAQs) to describe the issues that we aimed to address with our techniques. As we did not provide the roll call political spectrum in our approach (tests performed by Borja and Freitas showed that the general public was not always able to understand that visualization), we kept only three of the original eight tasks.

We have targeted the following visual analytics questions:

- VAQ1 Verify how each deputy voted in individual cases.
- VAQ2 Verify how parties vote: are they cohesive?
- VAQ3 Verify which deputies voted alike and if they continued to do so over time.
- VAQ4 Verify which parties voted alike and if they continued to do so over time.
- VAQ5 Check which deputies are politically divergent from their parties.
- VAQ6 Verify the behavior of parties and deputies across different periods of time.
- VAQ7 Check how deputies switched parties over time.
- VAQ8 Verify the activity of the Chamber of Deputies (i.e., the number of propositions voted in a given period of time).
- VAQ9 Verify how a deputy behaves after switching parties (i.e., is the deputy more aligned with the new party or the old one?).

In order to support these questions, the main requirement is that visualizations must be created on demand, and the exploratory process must allow users to compare data shown in different, simultaneous views. In the original CivisAnalysis, it is only possible to display data from one period at a time, and all visualizations were shown to the user simultaneously, in a single dashboard. As such, we posed the following requirements concerning the design of CivisAnalysis 2.0:

- R1 Provide **new and different visualization techniques**.

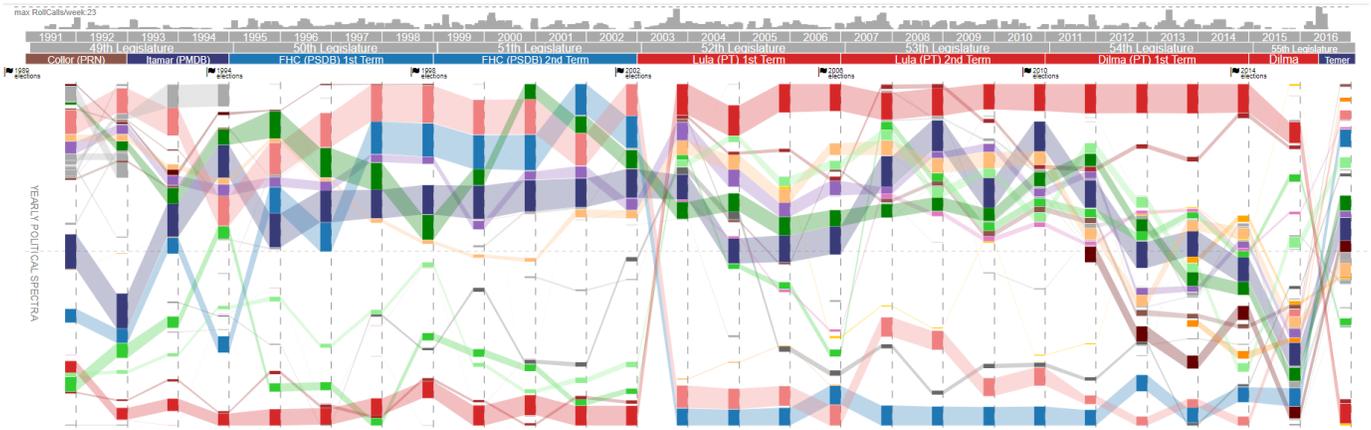


Fig. 1. The political timeline: the main visualization of CivisAnalysis 2.0, showing an overview of the voting behavior of parties across the spectrum. Each party is represented in a different color. Proximity means similarity of voting behavior. Parties closer to the top tend to be pro-government whereas those towards the bottom are in opposition. It covers seven legislatures and nine presidential terms, providing an overview of Brazil’s recent political history.

- R2 Allow **multiple instances** of the same visualization technique (e.g., for different time periods or data selections).
- R3 Allow for the creation of an **unlimited number of simultaneous views**.
- R4 **Resizable** views that can be **freely repositioned** in the workspace for a better exploration flow.
- R5 Explicit display of the relationships between views (shown as lines), so that users can **keep track of the steps taken in their exploration**.
- R6 Let users **show and hide** views to allow for a more customized workspace.

As with the original CivisAnalysis, our application initially displays the political timeline (Fig.1), which provides an overview of the parties’ voting behavior in the spectrum over 26 years of Chamber of Deputies history (helping address VAQ6). The periods are divided into years, presidential terms and legislatures. The timeline works as a hub from which users can launch other visualizations, which is done by selecting a period and clicking on it to open a menu of techniques. We based the user interface on multiple coordinated views that let users create visualizations of different periods of time. The following section describes CivisAnalysis 2.0 in detail.

IV. CIVISANALYSIS 2.0

The timeline serves as the starting point of the exploration process since users can launch other visualizations from it (Fig. 2). Visualizations are displayed in panels (supporting requirement R1) and, depending on the technique, new views can be created from these visualizations. Users can create as many instances of a technique as they desire (R2).

We take a hierarchical approach to manage all connected views: we use a tree structure that stores the necessary data regarding the visualizations and their corresponding windows, with each node representing one panel with a unique identifier (*id*). The interface reflects the tree structure in which nodes are

panels and a line connecting two panels depicts a parent-child relationship between them (R5).

Users can customize the layout of the workspace by creating an unlimited number of views (R3), which are displayed in panels that can be *resized* and *moved* by drag-and-drop (R4). Except for the timeline, all panels can be minimized, maximized, and removed (R6)—actions which are performed through buttons in each panel’s title bar (see Fig. 2). When a panel is minimized, it is replaced with a small icon. Minimized panels are treated just like full-sized ones: they are also *draggable* and their hierarchical relations to other panels are also shown as lines (the only difference being the use of dotted lines instead of solid ones). Double-clicking on a minimized panel restores it to its full size. Depending on the visualization, panels may include a settings menu that lets users interact with the data shown in the view.

A. Political Spectrum of Deputies

A political spectrum of deputies is generated from a time period selected by users. The application loads the deputies and the roll call data relative to this date range and applies a dimensionality reduction (DR) method to obtain the scatterplot that represents the similarity of voting behavior between deputies. Users can choose from three DR methods: PCA [18] by Singular Value Decomposition (SVD), Multidimensional Scaling (MDS) [23], and t-SNE [22].

PCA is a widely known and referenced method for dimensionality reduction. Moreover, it provides similar patterns to those identified by NOMINATE, a method used by social scientists. However, its result is not always easy to interpret and visualize, so we provided MDS and t-SNE as alternatives.

MDS is a classical method for visually representing the distances or dissimilarities between objects. The algorithm takes as input a symmetric matrix containing the pairwise distances between data points (e.g., deputy dissimilarities). In contrast to PCA, which retains the variance of the data, MDS preserves the distance between data points. Its main

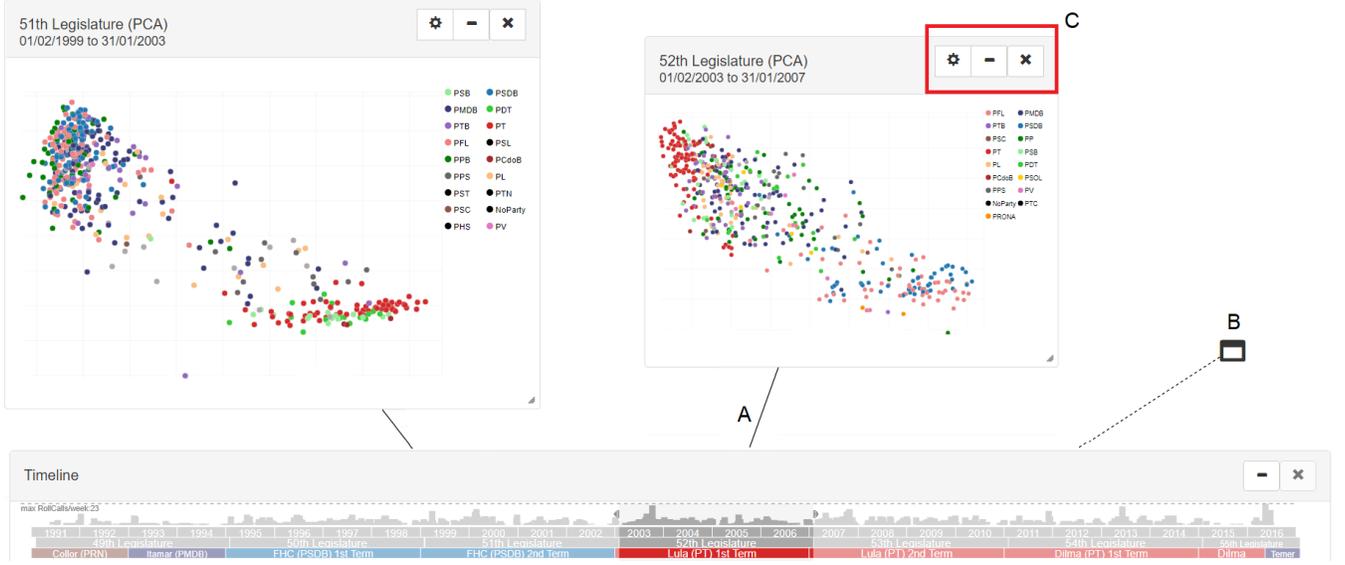


Fig. 2. Three panels are linked (A) to the timeline (cropped at the bottom): two scatterplots depicting different legislatures and one minimized panel, shown as an icon (B). (C) indicates a panel’s interaction buttons: settings, minimize, and close.

disadvantage is that the graph generated by the algorithm may be distorted and therefore will not always be a faithful depiction of the distances or similarities between objects.

t-SNE (t-Distributed Stochastic Neighbor Embedding) is a modern method used for the visualization of high-dimensional data. The main advantage in comparison to PCA is the reduction of the tendency to crowd points together in the center of the graph, providing well-separated clusters and thus helping users to identify them. However the algorithm is non-deterministic due to its use of random samples and local optimization, so users may have to run it multiple times until they get a good representation.

1) *Input Matrices*: The DR methods are based on matrices built from recorded votes. For PCA and t-SNE, the input is the matrix $R(M \times N)$ where $M = 513$ deputies and $N =$ number of roll calls, so that entry $R_{m,n}$ represents the vote of the m^{th} deputy in the n^{th} roll call. Vote values are defined as follows: 1 for *Yes*, -1 for *No*, and 0 for when the deputy was absent, obstructed, or there is no data.

MDS requires a dissimilarity matrix $D(N \times N)$, where $N = 513$ deputies and the $D_{m,n}$ value represents the dissimilarity between the m^{th} and the n^{th} deputies. This value is computed from the recorded votes by calculating the percentage of roll calls in which each pair of deputies voted differently.

2) *Generating the spectrum*: To obtain PCA-based spectrum, the SVD method is applied to matrix R (recorded votes), producing the following matrices:

$$\underbrace{R}_{513 \times N} = \underbrace{U}_{513 \times 513} \times \underbrace{\Sigma}_{513 \times N} \times \underbrace{V^T}_{N \times N} \quad (1)$$

The matrix U is a $513 \times N$ real unitary matrix and V an $N \times N$ real unitary matrix. To calculate the bidimensional *deputies spectrum* we multiply the two largest singular values found

in Σ by the left-singular vectors of U . t-SNE uses the matrix R , along with the following parameters values, as proposed by [14]: 10 for the *learning rate*, 30 for *perplexity*, and 10 seconds for the *iteration time*. Finally, MDS takes as input the dissimilarity matrix of deputies D that was obtained from the recorded votes. We used the classical version of the algorithm.

B. Visualization Techniques

Two other visualizations can be launched from the timeline view: the *Spectrum of Deputies* and the *Chamber of Deputies Infographic*. The period to be visualized is specified by selecting a year, legislature, or presidential term, or by brushing the timeline. Right-clicking the selected period triggers a menu from the type of visualization can be chosen. Clicking on the desired type will launch the visualization, which will be displayed in a new panel on the workspace.

Spectrum of Deputies: Regardless of the dimensionality reduction method used, the spatial proximity represents the vote similarity between deputies (thus addressing VAQ4). By default, deputies are represented as circles colored according to their respective parties. However, if the Roll Calls Histogram is instantiated, and a single roll call is selected, deputies’ colors correspond to their votes, according to a Vote-to-Color map ([“Yes”, “No”, “Obstruction”, “Abstention”, Chamber President, absence] are respectively mapped to [blue, red, green, purple, yellow, gray]). Users select deputies by clicking on them or by using the brushing tool. Hovering over deputies with the cursor shows more information (e.g., the deputy’s name and state). The Spectrum of Deputies addresses VAQ3, VAQ4, and VAQ5.

Fig. 3 shows how CivisAnalysis 2.0 deals with VAQ3 and VAQ4. Additionally, the users can cluster the deputies using K-Means by selecting the parameter k on the view’s settings, in

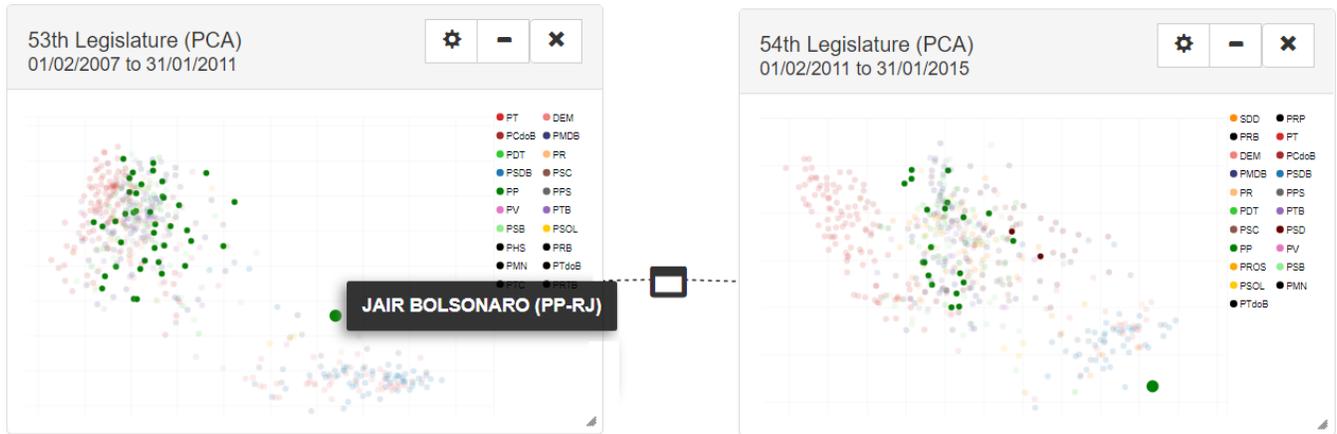


Fig. 3. All deputies of the PP party in the 53rd legislature are selected, showing cohesiveness. One deputy, Jair Bolsonaro, has a divergent behavior. To support VAQ3 and VAQ4, we instantiated the 54th legislature: his evolved to more divergent, and the party as a whole has moved to the center of the spectrum.

which case a convex hull is displayed to identify each cluster. Two additional views (described later in this section) can be launched for each subset created by this clustering: the Bar Chart and the Clustered Force Layout.

Chamber of Deputies Infographic: Often used to depict the distribution of seats in legislatures, this visualization shows deputies and parties in a semi-circle (see Fig. 4). Deputies are positioned according to their positions in the political spectrum and parties according to the average of their deputies, which results in a party-based clustering of the representatives. Additionally, a half-donut chart wraps the seat representations to display the proportion of each party in the legislature. User actions include the selection of deputies and parties and hovering over items for more information (as in the Political Spectrum).

From both the Political Spectrum and the Chamber of Deputies Infographic, another two visualizations can be launched: the *Heatmap Histogram* and the *Cropped Timeline*.

Roll Calls Heatmap Histogram: This visualization, shown in Fig. 4, helps users inspect the roll calls and provides them with an idea of how many motions were voted during a given time period. The data are shown as a horizontal histogram: a stack of rectangular cells (each representing a roll call), with the y-axis representing the months and the x-axis representing the number of roll calls. Roll calls are selected by clicking and hovering displays a tooltip containing a pie chart of the proportion of *Yes* and *No* votes. This visualization addresses VAQ8, as users can identify the most active periods and see what was happening in the country at the time in order to relate them with their historical context. Roll call cells are colored according to a colorblind-safe adaptation of the Agreement-to-Color scale of the first version of CivisAnalysis: the vote agreement scale ([100% “No” votes to 100% “Yes” votes]) is mapped to a color scale ([red, yellow, blue]) in which brightness corresponds to the number of votes.

The Roll Calls Heatmap Histogram is directly connected to its parent visualization. The default cell colors represent

the voting ratio of selected deputies in the parent visualization, which can be either the spectrum or the chamber of deputies infographic. Colors change dynamically according to the selected deputies, and when only one deputy is selected, the roll calls cells are colored according to the Vote-To-Color map. Along with the Chamber of Deputies Infographic, this technique addresses VAQ1 and VAQ2.

Cropped Timeline: The Cropped Timeline (Fig. 5) presents a timeframe snapshot of the main timeline. It is very similar to its parent view: it displays the parties in the spectrum as area elements anchored on the y-axis and divided into one-year intervals. This visualization’s new feature is the possibility of adding deputies to the timeline. The deputies’ behavior (their paths on the timeline) are represented by simple lines instead of area elements. As in other views, users can hover over the lines to obtain more information about the deputies and parties. This visualization allows for the observation of party switching by the displayed deputies, which addressed VAQ9. This is noticeable through the changes in the color of their respective lines, addressing VAQ7. This visualization is useful for analyzing specific deputies’ voting trajectories, which can be compared to that of other deputies or parties, which helps address VAQ6. Fig. 5 illustrates this by comparing the behavior of deputy Jair Bolsonaro, a frequent party switcher, to that of some the parties he belonged to.

Clustered Force Layout: The clustered force-based layout (Fig. 6A) provides a view of the composition of a cluster selected in a clusterized political spectrum. Deputies are represented as circles colored according to their respective parties, and a force-based layout uses only the party information to determine attraction between them. As usual, users can hover over circles to obtain more details about deputies.

Bar Chart: This visualization (Fig. 6B) represents the composition of a cluster selected in a clusterized political spectrum. Each bar corresponds to a party and its size represents the percentage of deputies of that party in the cluster.

Both the bar chart and the clustered force layout support

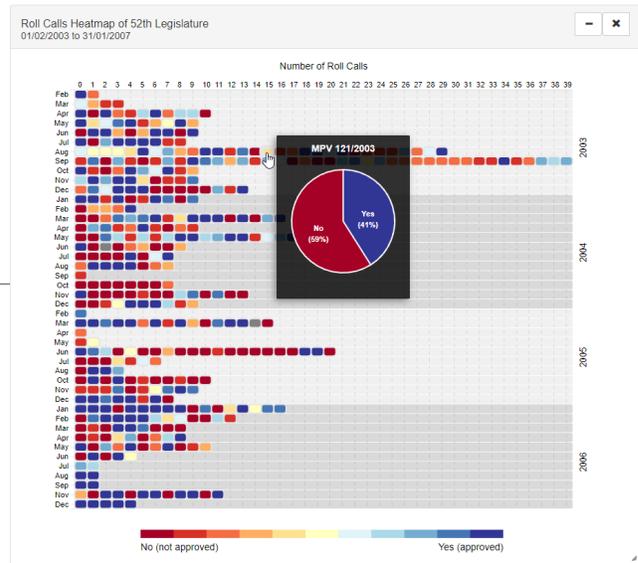
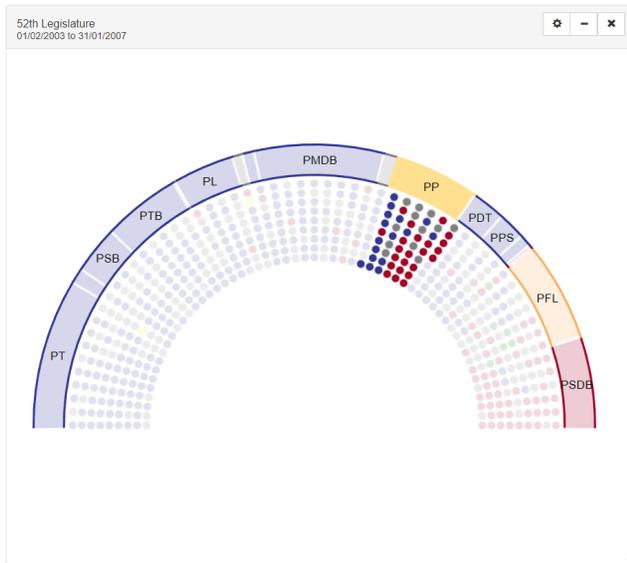


Fig. 4. Chamber of Deputies Infographic and Heatmap Histogram displayed in coordinated views. The PP party was selected on the infographic and the heatmap histogram reflects its votes using the Agreement-to-Color scale. The cursor is hovering over a roll call, triggering a tooltip that shows the results as a pie chart. This is reflected back on the infographic, where the deputies' colors represent their individual votes based on the Vote-To-Color map.

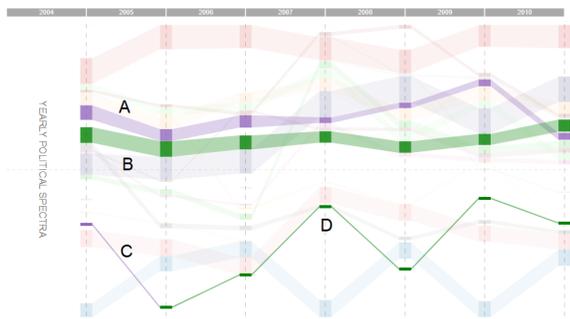


Fig. 5. Cropped Timeline showing the voting patterns of deputy Jair Bolsonaro and the parties he has belonged to from 2005 to 2010. (A) and (B) show, respectively, the patterns of his first and second parties (PTB and PP). (C) shows his voting behavior when a member of the PTB and (D) depicts his votes after switching to the PP.

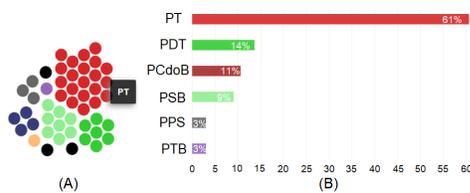


Fig. 6. The proportion of deputies by party shown as a **Clustered Force Layout (A)** and as a **Bar Chart (B)**. (A) uses a force-based layout in which deputies that belong to the same party attract one another, forming clusters that help users see the number of deputies of each party in a given subset (we can see here that most deputies belong to the PT). (B) is a traditional bar chart, with the percentage of deputies that belong to a bar's corresponding party being displayed inside each bar. Hovering over a bar shows a tooltip containing more information about its party.

the analysis of subsets of interest.

C. Additional Features

Visual Filter: Filtering is supported by clicking, hovering, or brushing deputies or roll calls and is coordinated between views (i.e., actions taken in one view will also be applied on the others).

District Filter: One or more districts (i.e., states) can be selected from a combo box, and deputies corresponding to the chosen districts will be highlighted.

Textual Filter: A text input box accessible from the panels' settings menu can be used to search for deputies based on their names. Found deputies will be highlighted.

V. CASE STUDY: PARTY COHESIVENESS OVER TIME

To illustrate how CivisAnalysis 2.0 can be used in practice, we conducted an informal, but broad case study, analyzing party cohesiveness over time. Our analysis covered each presidential term included in our data and was based on each legislatures t-SNE political spectrum visualization and the Chamber of Deputies composition infographic. The analysis was also complemented by historical information that helps contextualize and understand what is being seen.

We report herein our findings for only three presidential terms: Collor (the first one after the military period), FHC's first term, and Lula's second term. The whole case study, covering the nine presidential terms, is provided as a supplementary file that can be downloaded from <http://www.inf.ufrgs.br/~rnmsilva/CivisAnalysisDoc/>

A. Collor (PRN) 1990-1992

Main parties (at least 20 deputies): PFL (76), PL (20), PTB (27), PDS (46), PP (27), PDC (20), PMDB (104), PSDB (41), PDT (44), PT (35)

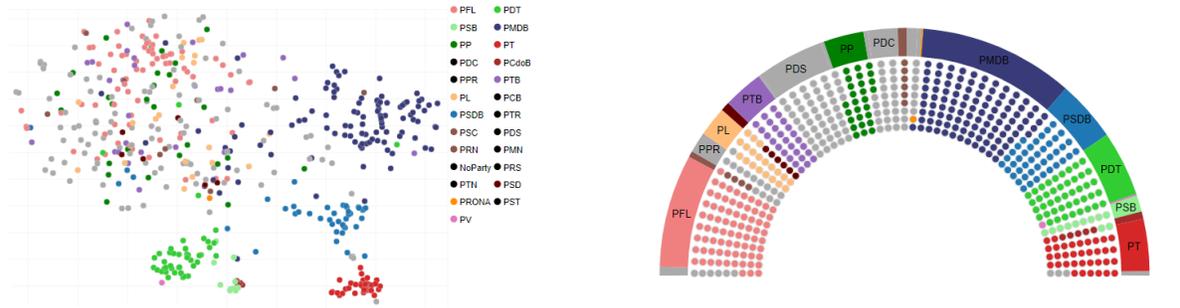


Fig. 7. Political Spectrum built from roll call data during Fernando Collor de Mello (PRN) term (1990-1992)

The first election after the military period resulted in a very fragmented parliament and a president, Fernando Collor de Mello, from a minor party, PRN, which had only four seats in the chamber of deputies. The two largest parties were the democratic reinventions of the two parties that existed during the dictatorship: the rightwing PFL (Liberal Front Party), which emerged from the conservative, pro-regime ARENA, and the PMDB (Brazilian Democratic Movement Party), a rebranding of opposition party MDB (Democratic Movement Party), which served as an umbrella for ideologies covering the entire political spectrum. Other sizable parties include PDS (Social Democratic Party), PP (Progressive Party), and PDC (Christian Democratic Party), the social-liberal PL (Liberal Party), the centrist PSDB (Party of the Brazilian Social Democracy) and PTB (Brazilian Labor Party) and the leftwing PDT (Democratic Labor Party) and PT (Workers Party).

In CivisAnalysis 2.0, we can see (Fig. 7) that some deputies form well-defined clusters, revealing the cohesiveness of their parties in terms of how similarly deputies voted. From this, we can see that PDT, PSDB, PT, and PMDB were very cohesive in this first legislature of Brazil's new democracy. PFL was slightly less cohesive, and most of the other parties were a lot less so while not straying too far from the PFL's positions (their deputies are spread over a large area overlapping PFL's, but they remain far from the other major parties).

B. FHC (PSDB) 1995-1999

Main parties (at least 20 deputies): PFL (110), PTB (25), PSDB (91), PPB (79), PMDB (86), PDT (24), PT (53)

The success of Fernando Henrique Cardoso's (PSDB) term as finance minister under former president Itamar Franco brought him to the presidency, with most of the major parties' deputies voting in cohesive patterns (Fig. 8). The exception to the rule was the PMDB, whose deputies divided into smaller clusters. Notable changes from the previous presidential term include the shrinking of the PMDB and the growth of the PFL (which became the largest party in the chamber of deputies), the PSDB (which almost doubled in size), and the PT. Also notable was the fusion of the PP and the PPR into the PPB (Brazilian Progressive Party), which formed a mostly cohesive bloc of deputies.

C. Lula (PT) 2007-2011

Main parties (at least 20 deputies): PT (83), PTB (20), PDT (24), PR (46), PMDB (85), PSB (27), PP (41), DEM (50), PSDB (59)

Despite a vote-buying scandal that threatened to bring down Lula's (PT) government during his first term in office, an economic boom led to his reelection and very high approval ratings. This was reflected in the voting behavior of deputies that can be very clearly seen in CivisAnalysis 2.0 (Fig. 9): the opposition, made up mostly of the PSDB and the DEM (Democrats a rebranded PFL), forms a small and cohesive but completely isolated bloc. All other parties are bundled together, although roughly overlapping and cohesive clusters can still be seen. It is interesting to note that the PT is still especially cohesive, but there is now a breakaway cluster that is very far from the main body of the party. Closer to the PT's main clusters are the PDT and about half of the PSB. After years of being very loose, the PMDB is again very cohesive. Slightly less so are the PP and the PR (Party of the Republic), a new party resulting from the fusion of the PL and the tiny PRONA (Party of the Reconstruction of the National Order). Also notable during this period was the beginning of the proliferation of very small parties.

VI. CONCLUSIONS

We presented a web-based application for the visualization of roll call data aiming at supporting users in gaining insights about the voting behavior of legislators and political parties in Brazil. Inspired by a previous solution, we designed CivisAnalysis 2.0 to address some limitations of that application by providing a set of techniques that can be integrated through a user interface based on hierarchically organized multiple coordinated views. From a political timeline that covers 25 years of Brazil's Chamber of Deputies activity, users can generate different views that depict the voting behavior of parties and deputies over time. Our application can be of use not only to citizens who want to be better informed but also to journalists interested in telling data-driven stories, and our case study is a partial example of this type of application.

As future work, we aim to generate political spectra of deputies and parties based on subsets of roll calls to allow

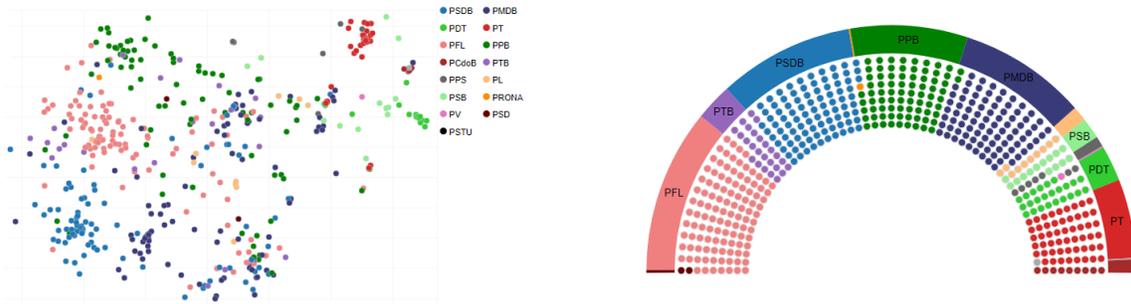


Fig. 8. Political Spectrum built from roll call data during Fernando Henrique Cardoso (PSDB) term (1995-1999)

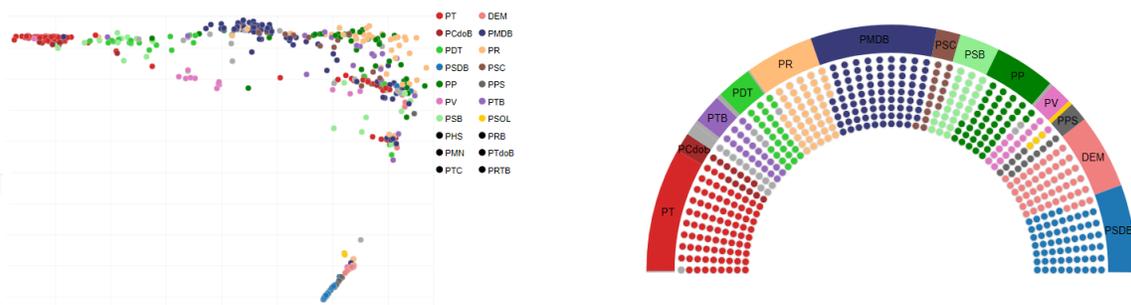


Fig. 9. Political Spectrum built from roll call data during Luís Inácio Lula da Silva (PT) term (2007-2011)

for the analysis of deputy votes on specific topics such as education and health. Other possible features include filtering roll calls by the number of votes each motion type requires to be approved (e.g., simple majority, absolute majority, three fifths of the votes).

ACKNOWLEDGMENTS

We would like to thank Francisco Gerda de Borja, for providing the code that allowed us to implement the timeline and the political spectrum of deputies. We also thank CNPq and CAPES, Brazilian research funding agencies, for financial support.

REFERENCES

- [1] "Tracking the united states congress," <https://www.govtrack.us>.
- [2] "Open data & apis," <https://www.govtrack.us/developers>.
- [3] "Senado federal - transparência," <https://www12.senado.leg.br/transparencia/>.
- [4] "Câmara dos deputados - informações no portal da câmara," <http://www2.camara.leg.br/transparencia/acesso-a-informacao/portal-da-camara-informacoes>.
- [5] R. Carroll and K. Poole, *Roll call analysis and the study of legislatures*. Oxford University Press, 2014, ch. 5, pp. 103–125.
- [6] K. T. Poole and H. Rosenthal, "Patterns of congressional voting," *American journal of political science*, pp. 228–278, 1991.
- [7] C. Schonhardt-Bailey, "Ideology, party and interests in the british parliament of 1841–47," *British journal of political science*, vol. 33, no. 4, pp. 581–605, 2003.
- [8] H. Rosenthal and E. Voeten, "Analyzing roll calls with perfect spatial voting: France 1946–1958," *American Journal of Political Science*, vol. 48, no. 3, pp. 620–632, 2004.
- [9] S. Hix, A. G. Noury, and G. Roland, *Democratic politics in the European Parliament*. Cambridge University Press, 2007.
- [10] P. Kinnaird, M. Romero, and G. Abowd, "Connect 2 congress: visual analytics for civic oversight," in *CHI'10 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2010, pp. 2853–2862.
- [11] R. Marino, "A valsa dos partidos, de collar a dilma," <http://www.todasasconfiguracoes.com/2014/04/24/a-valsa-dos-partidos/>, 2014.
- [12] "Basômetro," <http://estadaodados.com/basometro/>, 2013.
- [13] "Radar parlamentar," <http://radarparlamentar.polignu.org/>.
- [14] F. G. de Borja and C. M. Freitas, "Civisanalysis: Interactive visualization for exploring roll call data and representatives' voting behaviour," in *Graphics, Patterns and Images (SIBGRAPI), 2015 28th SIBGRAPI Conference on*. IEEE, 2015, pp. 257–264.
- [15] R. Carroll and H. Kubo, "Measuring and comparing party ideology and heterogeneity," *Party Politics*, 2017.
- [16] K. T. Poole and H. Rosenthal, "A spatial model for legislative roll call analysis," *American Journal of Political Science*, pp. 357–384, 1985.
- [17] R. Carroll, J. B. Lewis, J. Lo, K. T. Poole, and H. Rosenthal, "Comparing nominate and ideal: Points of difference and monte carlo tests," *Legislative Studies Quarterly*, vol. 34, no. 4, pp. 555–591, 2009.
- [18] I. Jolliffe, "Principal component analysis," in *International encyclopedia of statistical science*. Springer, 2011, pp. 1094–1096.
- [19] A. Jakulin, W. Buntine, T. M. L. Pira, and H. Brasher, "Analyzing the u.s. senate in 2003: Similarities, clusters, and blocs," *Political Analysis*, vol. 17, no. 3, pp. 291–310, 2009.
- [20] A. Perer and B. Shneiderman, "Integrating statistics and visualization: Case studies of gaining clarity during exploratory data analysis," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '08. New York, NY, USA: ACM, 2008, pp. 265–274.
- [21] A. Friggeri and E. Fleury, "Agreement Groups in the United States Senate," in *IOGDC - International Open Government Data Conference*, 2012. [Online]. Available: <http://hal.inria.fr/hal-00705682>
- [22] L. van der Maaten and G. Hinton, "Visualizing data using t-sne," *The Journal of Machine Learning Research*, vol. 9, no. 11, pp. 2579–2605, 2008.
- [23] W. S. Torgerson, "Multidimensional scaling: I. theory and method," *Psychometrika*, vol. 17, no. 4, pp. 401–419, 1952.