

Quill: A Framework for Constructing Negotiated Texts – with a Case Study on the US Constitutional Convention of 1787

Nicholas Cole
Faculty of History
University of Oxford
Oxford, United Kingdom
nicholas.cole@history.ox.ac.uk

Alfie Abdul-Rahman
Oxford e-Research Centre
University of Oxford
Oxford, United Kingdom
alfie.abdulrahman@oerc.ox.ac.uk

Grace Mallon
Faculty of History
University of Oxford
Oxford, United Kingdom
grace.mallon@univ.ox.ac.uk

ABSTRACT

This paper describes a new approach to the presentation of records relating to formal negotiations and the texts that they create. It describes the architecture of a model, platform, and web-interface (<https://www.quillproject.net>) that can be used by domain-experts to convert the records typical of formal negotiations in to a model of decision-making (with minimal training). This model has implications for both research and teaching, by allowing for better qualitative and quantitative analysis of negotiations. The platform emphasizes the reconstruction as closely as possible of the context within which proposals and decisions are made. A generic platform, its usability, and benefits are illustrated by a presentation of the records relating to the 1787 Constitutional Convention that wrote the Constitution of the United States.

CCS CONCEPTS

•Information systems →Information retrieval; •Applied computing →Law; •Human-centered computing →Interaction design;

KEYWORDS

Humanities, Data Exploration, Negotiated Texts, User Interfaces

ACM Reference format:

Nicholas Cole, Alfie Abdul-Rahman, and Grace Mallon. 2017. *Quill: A Framework for Constructing Negotiated Texts – with a Case Study on the US Constitutional Convention of 1787*. In *Proceedings of ACM JCDL Conference, Toronto, Canada, June 2017 (JCDL)*, 10 pages. DOI: 10.475/123.4

1 INTRODUCTION

Overview of problem: Many of the foundational texts of the modern world have not been written by individuals but negotiated by groups of people in formal settings. This class of document includes treaties between nations, texts such as the Universal Declaration of Human Rights [31], negotiated at the United Nations between 1946 and 1948, innumerable pieces of legislation negotiated in the world’s legislative assemblies, or Constitutions, such as the one

negotiated by the American States in the Constitutional Convention of 1787 [30], which met between May and September of that year. The records of such negotiations are extremely hard to follow. Typically, the journals published after such negotiations are a record of the proposals made and votes taken, sometimes with a near-verbatim record of what was said, but more often with concise descriptions of debate. The principal purpose of these records is to help those involved in the discussions to keep track of the process of negotiation in which they are immersed; intelligibility for later audiences is only a secondary concern. To fully understand how such negotiated texts are created, it is necessary to understand both the temporal sequence of proposals and votes and the hierarchy of decision-making, as proposed amendments are themselves amended and amended before being finally accepted or rejected.

The difficulty of following such records limits their utility for researchers. It also restricts broader use in education and outreach. Even a relatively short document, such as the Constitution of the United States, can be the product of many thousands of formal proposals and votes taken. A further complication is that there is no particular requirement that votes be taken in close temporal proximity to the proposal to which they relate, meaning that the context in which a vote is taken on a proposal can be considerably different to the context in which the proposal was originally made; indeed, the document may have evolved in the interim to the point that even those who proposed a particular change may vote against it. Providing any kind of detailed commentary on this process, or explaining it in classroom settings, therefore becomes extremely difficult. It can take many pages of text to explain the evolution of particular provisions in prose. For example, a ground-breaking essay attempting to explain the compromises and maneuvers which created the electoral college as the method for choosing the American President had to interweave analysis with an attempt to describe the changing circumstances in which proposals were being made and decided [22].

Particular case study: The Federal Constitution [30] has special significance as the first example of a constitution for a large state that was negotiated in this collaborative way, and a Constitution that at the same time incorporated many novel features, not least of which was the compromise between state and national authority that the Convention was able to agree. Those who recommended adoption of the new Constitution were proud of this aspect of its creation. As Alexander Hamilton put it in the first of the *Federalist* essays [28] (page 301), “It has been frequently remarked that it seems to have been reserved to the people of this country, by their conduct and example, to decide the important

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

JCDL, Toronto, Canada

© 2017 Copyright held by the owner/author(s). 123-4567-24-567/08/06...\$15.00
DOI: 10.475/123.4

question, whether societies of men are really capable or not of establishing good government from reflection and choice, or whether they are forever destined to depend for their political constitutions on accident and force.” Although the idea of a collectively-written constitution may have been novel, the delegates to that Convention were apparently well familiar with a formal, ‘Parliamentary’ style of conducting business. The rules that they adopted for themselves at the start of the process were relatively brief and specify only a few specifics, taking much of the process of debate for granted, a fact which underscores the pre-existing and shared understanding of Parliamentary-style processes to those who took part.

2 OPPORTUNITY FOR MODELING

In these formal processes texts are created through a process that consists of discrete events (such as to introduce new documents, to suggest amendments, to agree or reject proposals) that occur within a sequence that must be understood and also that have a hierarchical relationship to each other. The records typically left by such proposals record both relationships. Participants within any such process would have had access to something which is for the most part lost, which is immediate understanding of the current state of any document they were discussing. However, the very formality of these processes make it possible to reconstruct that context, and doing so transforms the utility of this class of records for a variety of users.

In this paper we present the following contributions:

- We have created a generic platform which can be used to encode the records relating to a wide variety of negotiations, and which can stand alongside and integrate with existing presentations of those records.
- We created a generic model for representing the negotiated of texts proposed by multiple actors.
- We have created a platform where each proposal to change text needs to be entered only once, even if the context in which it is voted on is significantly different to the one in which it had been created.
- We have established the conventions for the use of this model in a consistent fashion.
- We developed a web-based application to enable domain-experts to construct the model of a negotiation with minimal technical training.
- We have evaluated the ease with which this platform can be used by both domain-experts and by non-expert readers.

3 RELATED WORK

Parliamentary records are typically presented online as a photographic representation or transcription of original manuscript or print journals [3, 4, 11–13, 27]. While these provide the benefits of wider access and, if a transcription is available, being more easily searched, the problem for the reader of fully understanding the context of proposals and decisions remains. On the websites of Parliament and other modern legislative assemblies [20, 29], graphics are sometimes provided to illustrate the flow of a draft bill through the stages of being referred between committees and between chambers, but these provide only a very high-level overview of the process. These projects focus on providing access to the

official (and unofficial) records of negotiations. Unlike our project, they do not focus on modeling the negotiations or helping users to understand and analyze the records presented.

The most recent versions of Text Encoding Initiative (TEI) have introduced the concept of “genetic editions” [2] that can be used to describe changes to manuscripts, using this technology to describe the evolution of text in a formal setting would be cumbersome and error-prone. The potential of this genetic edition approach is shown by projects such as *Digital Variants* [33]. *Digital Variants* presents the variations between authors’ manuscripts, drafts and published editions, highlighting variations in texts. In each case, however, they are dealing with a relatively small number of texts to compare; by contrast a process of negotiation needs to be understood as (even in our case study) thousands of variant texts. None of the various XML transformation languages that we evaluated seemed appropriate for the task as we had identified it. In particular, they are too sensitive to changes in a document that make it impossible to describe a transformation once and apply it in evolving contexts in a way that would not produce undesired results, meaning that a proposal might have to be encoded multiple times for different contexts.

The distributed version control systems [7, 16, 26] used to track and reconcile changes to computer source code during software development solve many similar problems. Most of these systems, however, are tightly tied to the specific work-flows of developers and the line-oriented texts with which they are dealing.

However, the diff-match-patch algorithm [6] developed and released by Google as part of its Google Wave project (an implementation of Myers’s algorithm [18], coupled with a mechanism for applying “fuzzy” or “inexact” patches to a base text) provided a more promising starting point for a platform concerned with the better presentation of the process of negotiation. Unlike many algorithms for describing the changes to documents as patches and applying those patches, the Google diff-match-patch tools were designed to be used in an environment where multiple authors might be working on a document at once, and the order in which patches were received by the participants might vary. This is analogous to a situation in which a proposal may or may not be incorporated in to a document, depending upon whether a decision has been taken on it, and in which the changes that a particular proposal would make to a document might have to be made to a different base text to the one in which it was suggested, depending upon the decisions that have been made in the meantime. Working through a variety of test-cases proved that this implementation could be configured and used to track the process of negotiating a document.

Unlike the work of Kirschner et al. [8], Reed et al. [21], Shillingsburg [24], and Winograd and Floresour [34] our platform seeks to represent Parliamentary processes as they have historically occurred; it does not seek to use computers to impose a different and better structure on the process of debate and argument. It is focused on understanding the creation of documents through historical (or contemporary) Parliamentary procedure, rather than on visualizing, refining, and influencing the structure of arguments. Our platform focuses on reflecting the timeline and hierarchy imposed by the formal rules of debate in Parliamentary settings, which may not strictly reflect the logical structure of arguments directly.

4 APPLICATION BACKGROUND

The records related to the Constitutional Convention of 1787 include an official journal (kept by the Convention’s Secretary, William Jackson), and a variety of private diaries, the most famous of which was kept by James Madison. The official journal was entrusted to George Washington along with various related papers, and published by Congress in 1819 [1]. Madison’s journal was sold to Congress after his death, and published in various editions. In 1911, Max Farrand published *The Records of the Federal Convention* [5], a compilation of the various extant records, arranged to allow the parts of the various accounts relating to each day to be read alongside each other. Thomas Jefferson’s *Manual of Parliamentary Practice* [14, 32], published in 1801, provides a more detailed explanation of Parliamentary (i.e., formal negotiating) procedure as it was understood at the end of the eighteenth-century. This was compared by the authors to other, similar manuals [17, 23], and a model of formal negotiation produced as a result. Although the intricacies of the rules vary significantly between legislative bodies, the authors envisaged a platform which modeled negotiations, not one which enforced particular restrictions. They therefore constructed a model which could be used to model the creation and negotiation of text, and the passing either of draft documents or amendments between committees, as well as one which could track committee memberships across time. They constructed a series of test-cases based on the analysis of these Parliamentary manuals, to ensure that the platform that they developed would be able to model any likely action by a legislative assembly, recognizing that in practice legislative assemblies frequently suspend their own rules or behave in surprising (and less than entirely logical) ways.

However, although there were certain gaps in the records kept of the Convention (principally, the work of smaller committees), the records that did exist seemed to contain enough detail of formal proposals and votes taken that the process by which the US Constitution had been negotiated could be reconstructed in large part. Certainly, for better documented processes of the same type, the official journals provide enough information to reconstruct every step of debate. Jackson’s 1787 journal did make an effort to record both the wording of formal motions and the outcome of votes. What is often less clearly recorded or absent is the record of which way the various delegations voted on any question. In addition, the precise sequence of events within a given day was sometimes recorded differently by the official journal and the various private journals [5]. What was not known with certainty even to specialists before we began this project was whether the origin of every single clause of the final text could be accounted for by the extant records. We believed that a platform that could satisfactorily model these particular records would have broad application, since in many other processes the kind of uncertainty presented by these sources does not exist.

5 REQUIREMENTS ANALYSIS

Our primary requirement was the creation of a platform that could present the state of documents during any moment of negotiations. This would involve storing a representation of the sequence of events within a negotiation (the linear timeline of each committee), in such a way that the agreed state of documents and related

information could be calculated and presented to users for any selected moment in time. We anticipated that the research-assistants employed to enter the data would need to have excellent historical skills (because of the issues with the source material outlined above), but with little or no programming experience. We were also concerned to make the data capturing interface to be as intuitive as possible for non-technical uses. We wished to have an interface for data entry that would be intuitive for users without much technical training, and which would encourage the model to be used in consistent ways over the course of a long project.

We did not wish to duplicate the efforts of other projects. In particular, the images of manuscripts and historic printed sources, transcriptions of those documents, biographies of those in the 1787 Convention, have all been presented online by projects at public institutions such as The National Archives [19] and The Library of Congress [10], non-profit organizations such as ConSource [25] and The Liberty Fund [9], and by projects based in Universities, such as the Electronic Enlightenment Project [15]. All of these projects offer bespoke tools based on their specific expertise and the nature of the material they are presenting. Rather than compete with their efforts, we decided to make it possible to associate links to these resources with specific objects within our database, and also provide methods for other projects to link to related information within our own platform.

We knew that those entering data to our platform based on the interpretation of primary source data would need to exercise a certain amount of judgment in interpreting the sources, and that mistakes in data entry were possible. The source materials would raise issues that needed discussion among the editorial team, and the decisions taken would be documented. Since this project would involve building a model based on the source material, rather than a more mechanical process of transcription, the verification of data-entry would involve human judgment. We would need the system to be able to show those running the project which sections of the data entered in to the platform had been checked and by whom.

Since the purpose of the platform was in part to allow detailed commentary to be more easily written on the process of negotiation, we envisaged a system of “commentary collections” that would be owned by single or multiple authors. These collections would consist of an introductory essay and then explanatory text linked to specific events within the platform. A system for reviewing material was also introduced. These collections were to have two functions. Firstly, they could be presented to users who were viewing a section of the timeline to which they were relevant. Secondly, they would provide an alternative way to navigate the timelines that the platform would present. Users following the debate over particular topics would be able to use the commentaries related to those issues as a guide.

We decided that the main user interface for both users and editors would be built to work within a web browser, while the database and processing would be stored on a central server. Offering a web-based application would be attractive to a broad range of users without the need to install special software, either for viewing or editing, and would improve our ability to collaborate with teams working at other institutions. We would incorporate a flexible permissions system within the platform so that different categories

of user could be given specific permission to view, edit, or verify specific information within the platform. However, once material had been checked and approved for publication, we wanted as much of the platform as possible to be usable without registration. We also want the web-based interface to be highly flexible from a methodological perspective.

Due to the imperfect nature of the records from which we were working, we knew that our platform would also have to be able to capture the variation between manuscripts and, to a certain extent, uncertainty as to what had actually happened at particular moments. This was especially true of the records of particular votes, where it was frequently impossible to be certain who had voted in particular ways, even if the outcome of the vote was known.

6 DATA MODEL

After an analysis of the common features of Parliamentary-style processes and our requirements, we created a data model that captures the discrete and important elements of a process of negotiation. Each process, or ‘Convention’, comprises of two main components drawn from the historical data (the ‘Delegation’ and ‘Committee’ objects) and two that enable us to connect our model to other data (‘Resource Collections’ and ‘Commentary Collections’). Each Delegation is a collection of the ‘Person’ objects that comprise it; while each Committee stores details of its ‘Sessions’ as an ordered list, and within each of those the details of the discrete ‘Events’ that take place within that session, again as an ordered list. The Event types were further broken down in to three categories. Firstly, those that concern the creation, proposed amendment, or decisions taken about particular documents. Secondly those that concern the role of individuals (principally, membership of particular committees). Lastly, those that concern proposals that do not directly (but might indirectly) affect the creation or amendment of documents, such as motions to adjourn, or to rule particular proposals out of order. ‘Voting Records’, storing the details of particular votes, are linked to relevant Event objects.

Consistent use of the model presented by the platform was ensured in three ways. Firstly, a policy document was kept by the editors, and updated as specific issues were encountered. Secondly, the different types of events tracked by the platform was kept to a minimum to allow for an accurate representation of the Parliamentary process. In general, we found that users with a small amount of experience with the platform would use the model in consistent ways because the platform presented them with obvious choices in most situations and required them to make relatively few decisions about how the model would be used. Two to three days of training have proved sufficient with a variety of advanced undergraduates and graduate students. Thirdly, we designed the user interface to force users to capture the sources from which they were working in standard ways, usually by automatically validating the input and requiring active choices within the dialogue boxes presented to users, rather than offering default selections. We had to balance this with ease of use and the likelihood of error, and made adjustments where appropriate during the four months of the data-entry phase of development in 2016 on the basis of feedback from those doing the work of data-capture and verification.

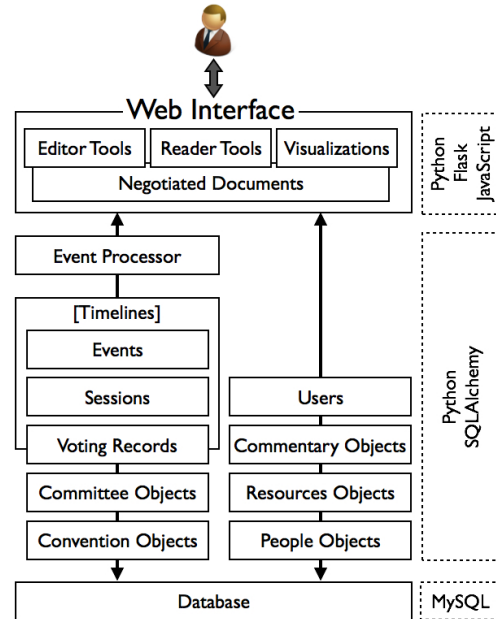


Figure 1: *Quill* platform architecture.

7 ARCHITECTURE

The architecture of the *Quill* platform is presented in Figure 1. The database stores the relationship between committees, sessions, and the ordering of events within each one. These objects are used to build the “timeline” of events within the platform. The database also stores non-timeline related objects, such as information about the actors involved in the Convention, Commentary, and other Resource objects that can be attached to particular events, sessions, committees, people and delegations, and information about the users of the platform, including their various permission levels. Events themselves were subdivided in to three main groups: those related to the creation and editing of documents (including the creation of documents, proposals to amend them, decisions taken on those proposals, and decisions to refer documents to new committees), events related to people (when an individual joined or left a committee, for example) and events related to the “procedure” of a given committee, such as a motion to adjourn, or more complicated motions that have the effect of invalidating or altering previous decisions.

Within this system, documents are represented as a proposal to create them and a series of proposals to amend them, together with the decisions taken on those proposals. An event processing layer is able to reconstruct the state of the documents and committee memberships for any given moment in the timeline reconstructed from this database. It is this event-processing layer that contains the algorithms for merging together documents on the basis of proposals made and votes taken. This algorithm must account for the fact that the state of a document relevant to a particular moment of debate must also take in to account the proposal under discussion. That is to say, that a proposal to change an amendment that has not yet been accepted needs to take in to account its parent amendments and any sibling amendments that have been agreed,

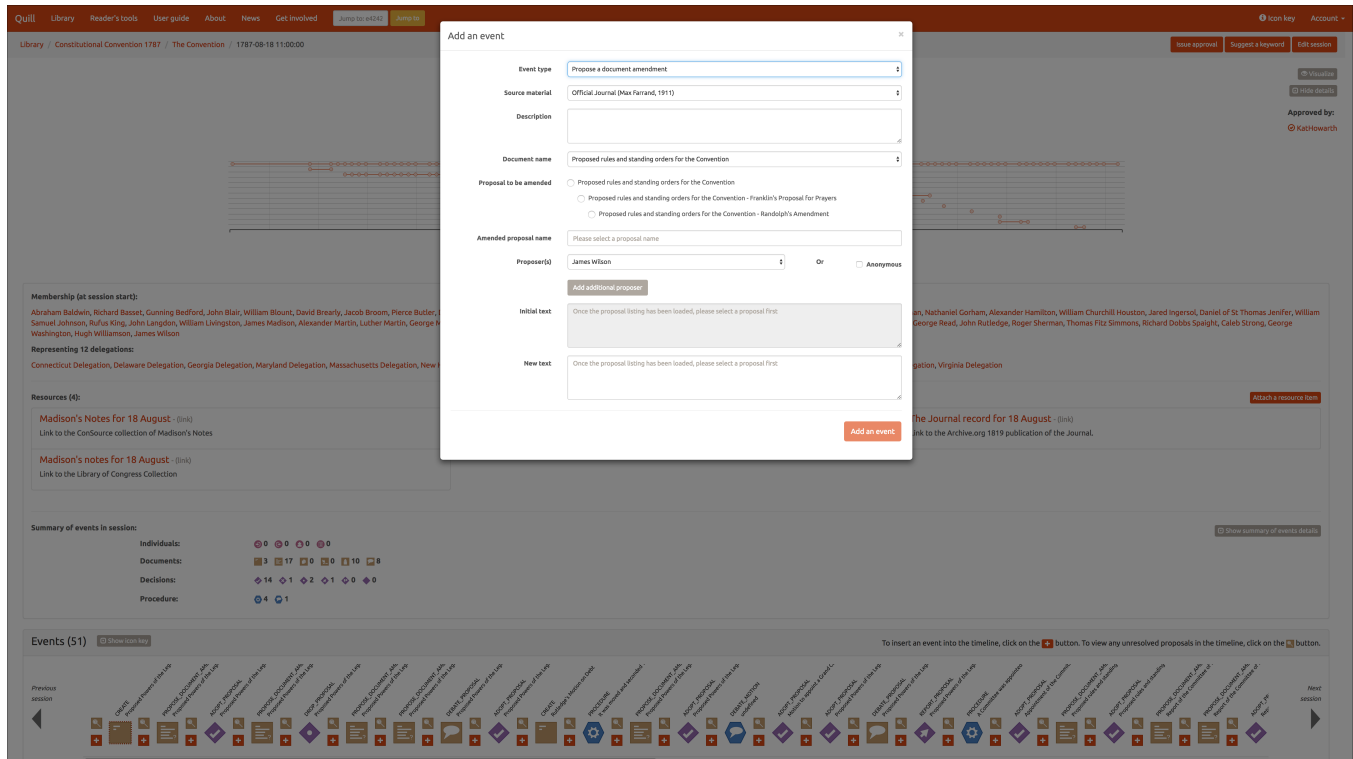


Figure 2: The Full Record view of the *Quill* platform showing the session-view interface for data-entry and editing.

but not any siblings that have not been agreed, nor other pending proposals that have not yet been resolved.

For any given moment of the timeline, therefore, the processing layer is able to calculate:

- A list of documents currently under discussion.
- A list of proposals that have not yet been resolved (the 'pending' proposals).
- The state of any documents currently agreed or under discussion.
- The state that those documents would be in if any of the pending proposals were adopted (if it is possible to generate this).
- The current membership of the committee.
- The information necessary to display various visualizations related to the document and proposal under discussion.

This can be combined with other information, such as links to further resources, that are associated with that moment in the timeline.

7.1 Algorithms Used

The text processing layer of our platform implements an algorithm that takes a series of events and calculates the set of proposals that should be included to create the various versions of a text relevant to a particular moment in time. The formal text of the proposals themselves are stored as a series of transformations encoded as diff objects. Our algorithm produces texts by combining applying these proposals in the order most likely to result in the intended

texts, and makes the necessary adjustments to the sensitivity of the match and patch algorithm to allow the document to be built. The current implementation uses the diff-match-patch algorithms [6] written by Google and work on plain text. Future versions of the platform may adopt an XML-based solution if algorithms with the required characteristics emerge.

8 VISUAL INTERFACE

We created a web application called *Quill* (<https://www.quillproject.net>) that would be used both for data-capture and by readers. We were aware that readers would not all require the same level of detail. Whereas those working on data-capture would necessary be most concerned with the detailed sequence of events within a committee session, many users of the platform would be better served with a more general overview of events.

For the bulk of data-entry, which concerns the creation of the model timeline, a view of the committee session in question is presented (see Figure 2). The timeline of the committee is represented by a horizontal series of icons. Around this is displayed information useful to those translating the source material into the model. The current membership of the committee is shown, together with summary lists of proposals that are pending for debate. Between each event of the timeline, users are offered a button to allow them to see the exact state of any documents or proposals at that point, and a button that allows them to add a new event to the timeline at that location. The latter causes a pop-up form to appear, in which the user is invited to select the type of event he/she wishes

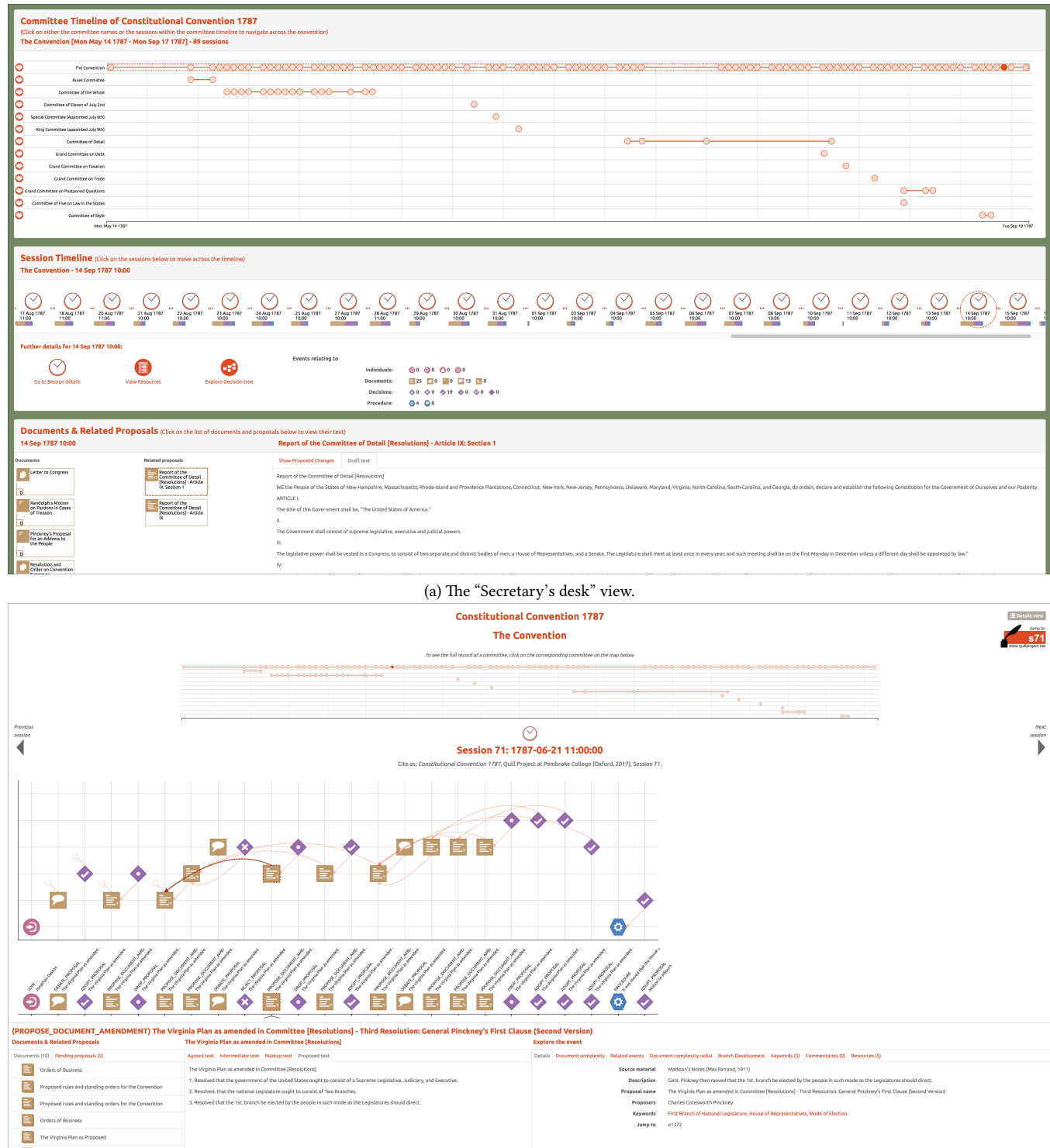


Figure 3: A comparison of two levels of detail which users can view, showing similar metaphors used in both.

to add (see Figure 2) — for example, a person joining a committee, or a proposal to amend a document. The fields of the rest of the

form are adjusted based on this choice. This session display also includes tools for editing existing events and deleting them from the timeline, and other functions needed by those entering data or verifying data entry.

Most of the other information required by the platform is entered on the “Full Record” view. This is where the names and members of particular delegations are stored, where the names of different committees are created. A page devoted to each committee shows a listing of all of its sessions and allows those to be added to. This page also shows users with the appropriate permissions an overview of which committee sessions have been verified, and by which users.

New users of our platform are guided to the “Secretary’s Desk” view (see Figure 3(a)). This combines a list of all of the Convention’s Subcommittees, a sense of when they met (represented simply as a timeline showing their first and last session), and a timeline of the individual sessions for any selected committee. A small chart under each session gives a quick sense of the number of individual events contained within it, while mouse-over information provides a more detailed view. For any selected committee session, we display a list of documents under discussion, indicating any whose text is altered by that day of debate, and any unresolved proposals that relate to a selected document. If users select a document, its current text is displayed in the center of the screen, and users can choose to highlight the text that was altered by the selected session’s debate. If they select an unresolved proposal, they can likewise see the affect that adopting that particular proposal would have on the state of the document. If users want more detail they can easily click through to the more detailed session view.

A similar set of metaphors is maintained in the more detailed visualization of each session (see Figure 3(b)). Along the top of the screen is a horizontal representation of the timeline of that committee session. Down the left hand side of the screen are lists of documents and proposals currently under discussion. Users who click on any of these documents are presented with their current agreed state, and clicking on any of the pending proposals shows the state of the documents that they would create. The center of the screen contains the text relevant to the proposal selected in the timeline. For proposals to amend documents, this is:

- The “agreed text” of the document. This is the text of the document if the document were simply accepted as final in its current state, with no further debate.
- The “proposed text” envisaged by a particular proposal.
- The “intermediate text” that this proposal amends. That is to say, the state of the document including any parent proposals of this text including any relevant sibling proposals.
- A display (the “markup text”) that highlights the difference between the “intermediate” and “proposed” texts.

On the right hand side of the screen is an area where users can choose to display either the details of the selected event or a variety of other tools. A ‘Document Complexity Tree’ shows all of the proposals and decisions that have formed the document relevant to the selected proposal in to its current state (see Figure 4). The proposal under discussion appears highlighted at the top. By navigating both the linear timeline at the top of the screen and using the complexity tree presented on the right hand side of it, users can quickly understand the relationship between different proposals

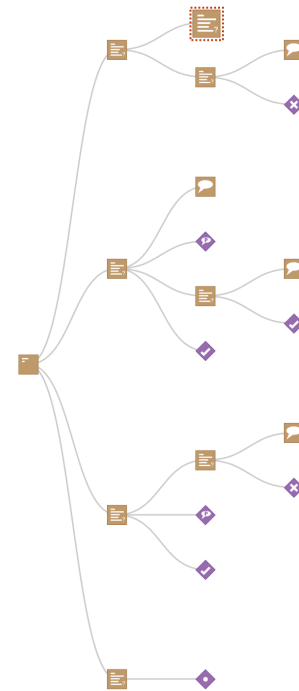


Figure 4: An example Document complexity display from early in the Constitutional Convention.

and the way in which they shape the creation of documents. The display of the tree of decisions that make up a document shows the extremely careful and often word-by-word nature of these negotiations, and provide an alternative method of navigation, allowing users to navigate decisions by hierarchy rather than by timeline. As other negotiations are modeled, it will become possible to compare the structure of decision-making between processes. Further tabs provide access to commentary collections relevant to this event or links to resources held in other collections relevant to this event.

Readers might want more of a sense of the structure of decision-making within the Convention as a whole. We can display the flow of documents between committees, or a display representing the overall hierarchy of decision-making within a process of negotiation, presented as a radial tree with the various committees, the documents they consider, amendments on those documents, and any subordinate amendments or decisions, radiating out from the center. A particular challenge is to fit this display on to smaller screens, and we offer users a choice of compact and expanded views.

The influence of particular delegations within the Convention is captured on a summary screen that shows two graphs. The first of these is created from a principle component analysis of the matrix of votes within the Convention, and gives a sense of how likely different delegations were to ally with each other. Users can choose a one-, two- (the default), or three-axis display. A bar graph represents how many proposals were made by members of a particular delegation, and how many of those proposals were accepted or rejected. A separate display allows the success or failure of different delegations during particular votes, presented as a spine chart.

9 USE OF THIS MODEL IN PRACTICE

The overwhelming majority of effort, as far as data-capture was concerned, was to convert the records of the Federal Convention in to a timeline of specific events. Those converting the records in to this model would read through the parallel records of each committee session (where multiple records existed), and decide how to reconcile any conflicts between the records. They recorded such decisions in a private “editors’ commentary” as part of the process. It was initially envisaged that reconciling the records in this way might be impossible — especially if competing forms of words were found recorded for the same proposal in different sources. The platform was therefore designed to allow for competing versions of the Convention timeline to be captured. However, it was found that in practice these features were not needed and that (where records existed at all) it was always possible to reconstruct the timeline of particular sessions if records were carefully reconciled by subject experts. The advantage of designing a platform for ease of use by non-technical users was that those recruited for data-entry could be selected for their subject-matter expertise. Due to the nature of the material and the model, an automated process for ensuring the accuracy of data entered in to our system was not possible. We implemented a system that would allow the data entry for each session to be marked as verified and for those in charge of the project to view who had entered and who had checked each section of the data.

Data entry was made intuitive for non-technical users with a few hours of training and supervised practice. The most frequent type of event in our model of the Constitutional Convention debates is the “Document Amendment”. Those working to enter the data select the point in the timeline where they wish to insert an amendment. They then select the document they wish to amend, and whether they are amending the base document or one of the proposed amendments. Once they have made this choice, the platform presents them with the current state of the text at that moment, which they are invited to edit to reflect the state of the document as it would be if the new amendment were to be accepted. They also enter other information, such as the source from which this event is taken, a free-form description of the event, and any known proposers of the amendment. When they have finished, the platform calculates the difference between what the user was presented and what they returned, and associates that patch with the new event.

The next most frequent type of event is a “Decision Event”. This records a decision on a particular proposal, be it to alter the text of an amendment, to adopt a section of text in to a document, or to accept or reject a document as a whole. There was considerable inconsistency in the records as to the level of detail with which such decisions were recorded. Sometimes the records record with certainty exactly which delegations voted for or against particular motions; sometimes only the totals on each side were known; sometimes only the outcome was known. Again, it was feared that the extant records might provide conflicting accounts, and so the platform was designed to allow competing accounts of the votes on particular questions to be displayed, or simply the uncertainty created by the records itself to be represented.

It was sometimes necessary to infer from the records that a particular decision had been taken. For example, it was the practice

of the Convention to debate and amend sections of text and then to approve or reject the amended section as a whole. Sometimes, this approval is not recorded in any of the extant sources. This may reflect the fact that the Convention was inconsistent in applying its own procedure, but it is equally likely that a unanimous consent to accept a section as amended and move on to the next order of business was simply not recorded by the Secretary as such. Our model, however, required the insertion of “Decision Events” to capture what the editors inferred to be a decision to agree text and move on. Such interpolated events are clearly marked within our platform. The need to include them highlights the fact that this project produces a model of negotiations, not a literal transcription of source material. It may be obvious from the sources that a particular piece of text has been agreed, even if there is nothing in them explicitly stating the fact.

The implicit rejection of text is a little more complicated. It will be the case in the course of a negotiation that particular suggestions that have been scheduled for debate have simply been overtaken by events — the section of the document to which they refer may have been altered in ways that make the suggestion redundant, or a similar suggestion may have been debated and agreed. In some cases, debate on an issue may simply be managed in such a way that a formal conclusion is never taken, perhaps to avoid the embarrassment of those involved. In these situations, there may never been a formal rejection of a proposal, and even to infer one at a specific point in the timeline may be misleading. For this reason, as well as marking proposals as “accepted” and “rejected”, the *Quill* platform’s model includes the ability to mark a proposal as “dropped”. From the point of view of the model, this has an identical effect to marking a proposal as rejected. It is removed from the list of pending proposals, along with any child amendments, and it is not incorporated in to the document. However, including this as a specific type of event allows a more accurate representation of the process of negotiation than the simple binary choice of accepting or rejecting a proposal, and can be made visually distinct for users.

The most surprising aspect of the platform for new users is that most documents debated by the Convention need to be represented at least twice. Most committees do not work from a blank sheet of paper, but work from an initial base text, either suggested by one of their members or passed to them from another committee. Frequently, they work through this document line by line or paragraph by paragraph, and in so doing produce a new report. The Constitutional Convention operated in the following way: a framework set of proposals, or suggested document (such as the famous “Virginia Plan”) would be offered to the Convention. This would be referred to a sub-committee — in the case of the Virginia Plan, the whole Convention sitting as a subordinate committee. This committee would work through the document section by section and clause by clause, and produce a report for the Convention to consider. The Convention would then work through this report, again amending section by section and clause by clause. In this way, everything would have been considered at least twice, once by each committee.

In a world of paper, quill, and ink, this process created a significant record-keeping challenge, which it would have been the task of the secretary to manage. As the Convention or subcommittees worked through the documents referred to them, he would have had to write out the new text on clean sheets of paper. No doubt these

sheets of paper rapidly became untidy and even hard to follow, and perhaps it is for this reason that they were not entrusted to Washington for safe-keeping but were instead deliberately destroyed, even though copies of the various base documents are extant.

When represented in the *Quill* platform, this process looks identical. If a committee is working through one document to create its own report, the initial document is not shown as amended, but rather the clauses of it are modeled as being gradually incorporated in to a new document, which represents the report of the committee.

We were also concerned to make it easy to encourage precise commentary and accurate record-keeping. For users with appropriate permissions, a button to add commentary to any event within a timeline was presented, which presents the user with a pop-up form. Authors creating commentary collections use this button to add their comments. The data-entry team used this same system to flag issues within the timeline which required review.

10 USER EVALUATION

We built the core of the platform and web interface over the course of the academic year 2015–6. The platform was then opened to three interns working at the Utah Valley University’s Constitutional Studies Center. These interns were given several days of training via video-link and then encouraged to experiment with the platform. They were encouraged to try modeling parts of the 1787 Convention records using the platform. They were able to accurately and consistently use the model, and highlighted a number of deficiencies in the user interface for both data-capture and readers. A detailed record of their observations and suggestions was kept, and used to inform modifications to the platform and a set of editorial conventions that would be used to model the records of negotiations consistently. Although other team-management tools were initially used, it eventually became apparent that the *Quill* software itself provided by far the best record of this set of editorial decisions.

From June to October 2016, a recent Oxford University graduate in history (a co-author of the paper) was employed to do the work of the data-entry for the 1787 Convention, using the 1911 publication of the records, and a graduate in law employed at Oxford from September onward to assist with verification. We were ourselves surprised by the complexity even of the short negotiation covered by our case study. We had originally guessed that this would require around five hundred events, but the final model of the Convention required close to four thousand. Interns at the Utah Valley University continued to assist with verification, with some of the ancillary and less historically difficult data-entry tasks. Data capture for the 1787 Convention was completed in approximately three and a half months in total. We estimated that approximately a day of work on data-capture was required for each day of the Convention itself, though this time varied significantly depending upon the state of the records. None of those working on the data-entry or verification required specific technical training beyond a few days’ training on the specifics of the *Quill* platform and a discussion of the assumptions of the model. Throughout the process of data-capture, a record was kept of any aspect of the Convention records that had been difficult to model and the ways in which

those difficulties were resolved. At the end of this process the two graduate students involved in the process of data-capture wrote a detailed data-entry guide, that incorporated the decisions taken by the editorial team during the process as particular circumstances had been encountered. This data-entry guide is intended for use in future work.

We have subsequently held a workshop to evaluate both the readers’ interface and the interface for data-capture. In the workshop we invited users unfamiliar with the platform to explore the records of the 1787 Constitutional Convention, and to attempt to encode one day of negotiation which was part of the process for creating the Universal Declaration of Human Rights (negotiated at the United Nations 1946–8) [31]. The participants in this workshop were unfamiliar with the platform, from a range of disciplines, and ranging in education level from graduate-students through to senior researchers. After an hour of introduction to the platform, all of them were able to grasp the basic use of the platform from a readers’ perspective and understand the basic conventions and metaphors used by the web interface. All of them were able to understand how to create the basic records related to the work of a process of negotiation. Most of them were able to translate the documents given in to an accurate model of the start of the day of negotiation given. In feedback, users commented that the hardest part of this process was understanding the conventions of the minutes in question not the use of the technology they had been given. This confirmed our view that, as far as data-entry is concerned, our platform was appropriate for use by domain-experts working within properly-agreed guidelines.

We have invited several domain-experts to prepare commentary collections for us relating to specific questions raised by the records of the Convention. We have already published the first of these, by Lindsay Chervinsky, who was completing her PhD on the idea of a President’s cabinet.¹ Those who have no other experience of data-entry have found it easy to attach commentary to specific objects within our timeline.

10.1 Other Feedback and Suggestions

We have been asked by various researchers to make adjustments to the platform to allow it to be used for a wider-range of material (especially that relating to foreign-language material), and to be used to capture debates in real-time as well as working from historical materials. In the latter case, we believe this can be achieved mostly through hiding options, and in particular the ability to edit the timeline arbitrarily.

Adopting a suggestion frequently made, we will be extending the platform to improve the machine-readable interfaces. In particular, we will implement an XML output, using TEI conventions [2]. Although this is likely to result in an extremely complicated set of XML documents, but it would be suitable for archiving and importing purposes.

10.2 Conclusions and Future Work

We have developed a system that allowed domain-experts with minimal technical training to model the almost four thousand proposals and votes that created resulted in the text of the United

¹https://www.quillproject.net/commentary_collections/7

States Constitution, even when working from imperfect records. The history of the final text of the Constitution can be accounted for within our model, from beginning to end, and although the platform does have the ability to present places where the text is uncertain because of conflicts in the manuscripts, there were no discrepancies substantive enough to warrant it for the 2016 presentation of the records. Such variations as there are relate to extremely minor points of capitalization and spelling. The nature of the records was such that it required significant expertise to accurately model. The reconstruction offered would not have been possible from any one of the surviving sources, but taken together and used systematically and rigorously, we believe they capture the complete work of the Convention's formal business for all of the committees where James Madison and William Jackson were present. For other, smaller committees, we have been able to show the specific text that was given to them to consider and the report that they returned. These areas of darkness, however, are much smaller than might be assumed, and the process of producing this edition has generally reassured, rather than challenged, our confidence in the extant records.

Our model is not tied to the presentation of these particular records, but rather intended to be deployed for a wide range of records. We are currently identifying targets for future work, and will be expanding the platform to assist with the presentation of multi-language texts. Expanding the range of materials held within our database will enable us to both quantitatively and qualitatively compare and contrast different processes of negotiation.

We are also keen to make the material we are presenting useful in a classroom setting, and especially in the classroom settings of the United States. We are currently in discussions with non-profit organizations that work to generate classroom material to highlight several potential approaches. Firstly, we would need to provide an easy way for content-creators to integrate our material in to their existing lesson plans. Secondly, we would need to provide them with an interface that would let them create resources suitable for classroom use within our platform.

We believe that this platform is relevant to the presentation of records held on the negotiation of treaties, constitutions, and innumerable pieces of legislation created in Parliamentary settings. We believe that this platform has the potential to democratize understanding of these complicated processes and transform the utility of existing digitized collections for a wide range of audiences.

11 ACKNOWLEDGMENTS

The authors would like to thank K. Howarth (University of Oxford), L. Merrell, M. Nolte, and C. Sponseller (Utah Valley University) for assisting with the verification and data-entry tasks. The authors would also like to thank D. Burt, D. De Roure, O. Durand, G. Mantegari, Y. Nasifoglu, and P. Willcox (University of Oxford) for participating in the workshop and their helpful feedback.

REFERENCES

- [1] M. S. Bilder. 2012. How Bad Were the Official Records of the Federal Convention?. In *The George Washington Law Review*, Vol. 80. 1620–1682.
- [2] L. Burnard, F. Jannidis, E. Pierazzo, and M. Rehbein. 2010. An Encoding Model for Genetic Editions. (Apr 2010). <http://www.tei-c.org/Activities/Council/Working/tcw19.html> (Accessed on: 27 Jan 2017).
- [3] EDSITEment. 2016. The Constitutional Convention of 1787. (2016). <https://edsitement.neh.gov/curriculum-unit/constitutional-convention-1787> (Accessed on: 29 Jan 2017).
- [4] Z. Elkins and T. Ginsburg. 2005. Comparative Constitutions Project. (2005). <http://comparativeconstitutionsproject.org/> (Accessed on: 29 Jan 2017).
- [5] M. Farrand. 1911. *The Records of the Federal Convention of 1787*. New Haven: Yale University Press.
- [6] N. Fraser. 2006. google-diff-match-patch: Diff, Match and Patch libraries for Plain Text. (2006). <https://code.google.com/p/google-diff-match-patch/> (Accessed on: 27 Jan 2017).
- [7] GNU RCS. 2013. GNU RCS. (2013). <http://www.gnu.org/software/rcs/rcs.html> (Accessed on: 27 Jan 2017).
- [8] P. A. Kirschner, S. J. Buckingham Shum, and C. S. Carr (Eds.). 2003. *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making*. Springer, London, UK.
- [9] Liberty Fund. 2016. Liberty Fund. (2016). <http://www.libertyfund.org/> (Accessed on: 30 Jan 2017).
- [10] Library of Congress. 2016. Part of: James Madison Papers, 1723-1859. (2016). <https://www.loc.gov/search/?fa=partof:james+madison+papers,+1723-1859:++subseries+5e,+madison%27s+original+notes+on+debates+in+the+federal+constitutional+convention,+1787> (Accessed on: 29 Jan 2017).
- [11] Lilian Goldman Law Library. 2008. Notes on the Debates in the Federal Convention. (2008). http://avalon.law.yale.edu/subject_menus/debcont.asp (Accessed on: 29 Jan 2017).
- [12] D. Linder. 2016. The Constitutional Convention 1787. (2016). <http://law2.umkc.edu/faculty/projects/Fttrials/conlaw/convention1787.html> (Accessed on: 29 Jan 2017).
- [13] G. Lloyd. 2016. The Constitutional Convention. (2016). <http://teachingamericanhistory.org/convention/themes/> (Accessed on: 29 Jan 2017).
- [14] T. E. May. 1844. *A Treatise upon the Law, Privileges, Proceedings and Usage of Parliament*. Charles Knight & Co., London, United Kingdom.
- [15] R. McNamee. 2016. Electronic Enlightenment Project. (2016). <http://www.e-enlightenment.com/> (Accessed on: 29 Jan 2017).
- [16] Mercurial. 2016. Mercurial. (2016). <https://www.mercurial-scm.org/> (Accessed on: 27 Jan 2017).
- [17] Ministry of Parliamentary Affairs, Government of India. 2017. Manual of Parliamentary Procedures in the Government of India. (Feb 2017). (Accessed on: 4 Feb 2017).
- [18] E. W. Myers. 1986. An O(ND) difference algorithm and its variations. *Algorithmica* 1, 1 (1986), 251–266.
- [19] National Archives Catalog. 2016. Journal of the Federal Convention. (2016). <https://catalog.archives.gov/id/7347105> (Accessed on: 29 Jan 2017).
- [20] Parliament. 2016. Bills before Parliament 2016-17. (2016). <http://services.parliament.uk/bills/> (Accessed on: 29 Jan 2017).
- [21] C. Reed, S. Wells, M. Snaith, K. Budzynska, and J. Lawrence. 2011. Using an argument ontology to develop pedagogical tool suites. In *Tools for Teaching Logic (Lecture Notes in Computer Science)*, P. Blackburn, H. van Ditmarsch, M. Manzano, and F. Soler-Toscano (Eds.). Springer, 207–214.
- [22] W. H. Riker. 1984. The Heresthetics of Constitution-Making: The Presidency in 1787, with Comments on Determinism and Rational Choice. *The American Political Science Review* 78, 1 (1984), 1–16.
- [23] H. M. Robert. 1923. *Parliamentary law*. Century Co., New York.
- [24] P. L. Shillingsburg. 1996. *Scholarly Editing in the Computer Age: Theory and Practice* (3rd ed.). University of Michigan Press.
- [25] J. Silverbrook and M. Johnson. 2007. ConSource. (2007). <http://www.consource.org/> (Accessed on: 29 Jan 2017).
- [26] Software Freedom Conservancy. 2016. Git. (2016). <https://git-scm.com/> (Accessed on: 27 Jan 2017).
- [27] H. Surden. 2015. Constitution Explorer (Beta). (2015). http://www.harrysurden.com/projects/visual/USCode_D3/constitution/US-Constitution.Tree.html (Accessed on: 29 Jan 2017).
- [28] H. C. Syrett (Ed.). 2011. *The Papers of Alexander Hamilton Digital Edition*. Vol. IV. University of Virginia Press, Rotunda, Charlottesville.
- [29] J. Tauberer. 2004. Bills and Resolutions. (2004). <https://www.govtrack.us/congress/bills/> (Accessed on: 29 Jan 2017).
- [30] The Library of Congress. 2016. United States Constitution. (Oct 2016). <http://www.loc.gov/rr/program/bib/ourdocs/Constitution.html> (Accessed on: 4 Feb 2017).
- [31] United Nations. 1948. Universal Declaration of Human Rights. (10 Dec. 1948). <http://www.un.org/en/universal-declaration-human-rights/> (Accessed on: 27 Jan 2017).
- [32] U.S. Government Publishing Office. 2016. House Rules and Manual. (2016). <https://www.gpo.gov/fdsys/browse/collection.action?collectionCode=HMAN> (Accessed on: 27 Jan 2017).
- [33] Digital Variants. 1996. Digital Variants. (1996). <http://www.digitalvariants.org> (Accessed on: 4 Feb 2017).
- [34] T. Winograd and F. Flores. 1987. *Understanding Computers and Cognition: A New Foundation for Design*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.