Software Support for Discourse-Based Textual Information Analysis: A Systematic Literature Review and Software Guidelines in Practice

Patricia Martin-Rodilla [1,*] and Miguel Sánchez [2]

[1]   Information Retrieval Lab (IRLab), Facultade de Informática, University of A Coruña,
     C.P. 15008 A Coruña, Spain
[2]   Chocosoft S.L., C.P. 15703 Santiago de Compostela, Spain; miguel@chocosoft.net
*   Correspondence: patricia.martin.rodilla@udc.es

Abstract: The intrinsic characteristics of humanities research require technological support and software assistance that also necessarily goes through the analysis of textual narratives. When these narratives become increasingly complex, pragmatics analysis (i.e., at discourse or argumentation levels) assisted by software is a great ally in the digital humanities. In recent years, solutions have been developed from the information visualization domain to support discourse analysis or argumentation analysis of textual sources via software, with applications in political speeches, debates, online forums, but also in written narratives, literature or historical sources. This paper presents a wide and interdisciplinary systematic literature review (SLR), both in software-related areas and humanities areas, on the information visualization and the software solutions adopted to support pragmatics textual analysis. As a result of this review, this paper detects weaknesses in existing works on the field, especially related to solutions' availability, pragmatic framework dependence and lack of information sharing and reuse software mechanisms. The paper also provides some software guidelines for improving the detected weaknesses, exemplifying some guidelines in practice through their implementation in a new web tool, *Viscourse*. *Viscourse* is conceived as a complementary tool to assist textual analysis and to facilitate the reuse of informational pieces from discourse and argumentation text analysis tasks.

Keywords: discourse analysis; argumentation analysis; information visualization; software assistance; *Viscourse*; systematic literature review; information reuse

1. Introduction

Information management in humanities disciplines necessarily involves natural language textual sources analysis at any level. Recently, digital humanities area is including initiatives and works on how to assist textual analysis tasks via software (see chapters 19–21 from [1] for some examples).

Mainly, this software assistance focuses on two different assistance directions: (1) natural language processing automatization solutions and (2) the development of visualization techniques to visualize results from automation tasks —from group 1— or to help on manual or semiautomatics textual analysis, including annotation tasks, adapted to humanistic disciplines.

Regarding the second kind of software assistance, there are several proposals at visualization level [2,3] for visualizing grammatical or lexical structures of the texts, dealing with morphology, lexicology or syntax levels of linguistic analysis.

However, as long as we focus on more conceptual or relational aspects of the textual sources (at a semantics or pragmatics levels), the previous group 2 of software assistance decreases. Although

textual analysis at the discursive or argumentation level is currently being applied in a wide range of studies and disciplines (political speeches, debates, online forums, historical sources, among others), there are no systematic studies on the software assistance offered at discursive or argumentative levels.

This paper presents a wide and interdisciplinary systematic literature review (SLR), both in software-related areas and humanities areas, on the information visualization and the software solutions adopted to support pragmatics textual analysis at discursive or argumentation levels. Note that, although we are conscious that linguistically and even philosophically the discursive and argumentation approach to text analysis have multiple analysis possibilities and frameworks, this paper treats the discursive-argumentation information at the same level, regardless of the linguistic framework that is taken as a basis, in order to perform a more complete study.

The paper is organized as follows: Section 2 details the materials and methods employed (the systematic literature review methodology, search strategies and inclusion/exclusion criteria employed). Section 3 presents the systematic review results, a first discussion of the weaknesses detected and the answers to the research questions defined for this SLR study. Based on these results, Section 4 presents some software guidelines identified for improvement and an implementation proposal for the guidelines' application in a web tool, *Viscourse*. Finally, Section 5 discusses future directions.

2. Materials and Methods

This section presents the Systematic Literature Review (SLR) performed, detailing the methodology, search strategies and inclusion/exclusion criteria adopted.

2.1. Systematic Literature Review

In order to analyze existing solutions on software visualization for supporting or assisting at discursive or argumentation textual analysis, we have performed a Systematic Literature Review (hereafter SLR). We have followed the SLR guidelines by Kitchenham and Charters in 2007 [4], where a Systematic Literature Review is defined as a methodology for "identifying, evaluating and interpreting all available research relevant to a particular research question, or topic area, or phenomenon of interest" [4]. The rest of this section documents our SLR process.
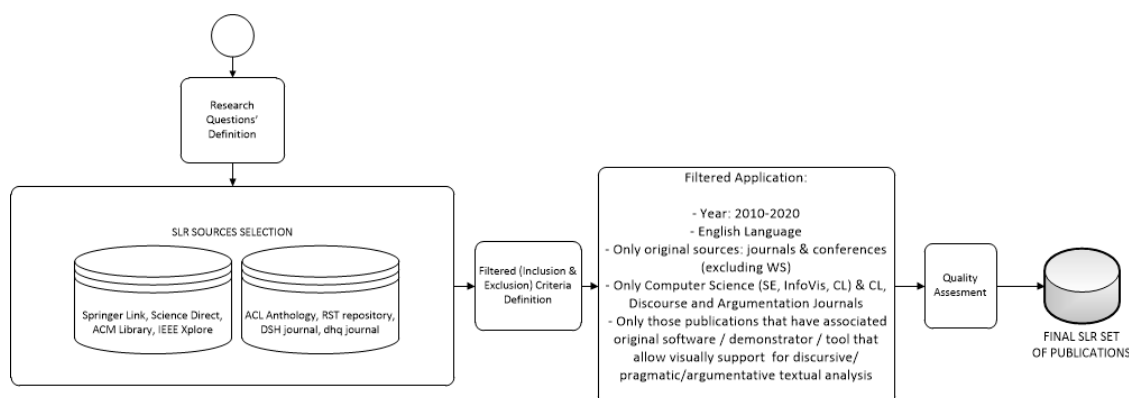
2.1.1. Research Questions

The first step in SLR methodology consists of defining research questions that underpin the SLR process. We have identified 4 relevant research questions to guide the review:

1. RQ1: What evidence is there that discourse and argumentation textual analysis is currently supported via information visualization software?
2. RQ2: What kind of support are these works providing and how is it implemented? We have defined seven categories for analyzing the kind of support provided by the main contribution presented in these works:

    - **InfoVis technique**: The support provided is mainly visual. For example, a new information visualization technique or its application to a new kind of discursive or argumentation textual information.
    - **Linguistic resource**: The main support is provided by offering a new linguistic resource: a new corpus, annotation information, taxonomy, ontological information, etc.
    - Complete **software tool**: The assistance is materialized as an entire new software tool.
    - **Application example**: The support is provided by illustrating a new discursive or argumentation analysis in a new domain of application or corpora.
    - **Discourse metrics measurement**: The support is provided by implementing software mechanisms to calculate discursive or argumentation standard metrics for helping in the textual analysis.

- Fully **automatic analysis support**: The support is provided by implementing automatic software solutions for visualizing automatic discursive or argumentation analysis: fully automatic parsers, new algorithms or machine learning techniques, etc.
- **Survey/Empirical or Qualitative study**: Works focused on qualitative analysis.

3. RQ3: Does the software support offered in these works present any weakness or deficiency reported in the study itself or detected as a result of the review?

4. RQ4: Is it possible to identify some software guidelines for improving the existing information visualization software solutions for supporting or assisting discourse and argumentation textual analysis tasks?

Figure 1 details the SLR process performed. First, we identified the main research repositories for finding relevant work that answer our research questions. Then, we eliminated duplications, defining a filtered inclusion/exclusion strategy. Finally, we analyzed the title and abstract of the resultant publications. In this step, we also applied some quality assessment criteria as a checklist to the works obtained. This quality assessment step ensured that the works reviewed have no important bias and fit the scope and relevance criteria. The resultant set of publications constitutes the relevant work done on visualization software assistance (labeled as group 2 in the Introduction) for supporting or assisting discourse and argumentation textual analysis tasks.

Figure 1. Systematic literature review (SLR) process steps, inspired by the Kitchenham and Charters guidelines (2007) [4].

2.1.2. Sources, Search Strategies and Filtered Criteria

In order to answer the research questions, we have defined a combined search strategy in two different kinds of sources. On the one hand, we have searched in four international and well-known digital libraries of research publications. Specifically, we have chosen *Science Direct* [5], *Springer Link* [6], *ACM library* [7] and *IEEE Xplore* [8] due to the accessibility, degree of reliability and relevance in software engineering. On the other hand, we have also performed the same searches in Digital Humanities relevant repositories that contain works on software engineering and discourse or argumentation textual analysis. Specifically, we have reviewed *ACL Anthology* [9] and *RST repository* (*Rhetorical Structure Theory*) [10–14] as main sources for works on computational linguistics in the area. Moreover, we have included the main two research journals on Digital Humanities on an international level: *Digital Scholarship in the Humanities* (henceforth DSH) [15] and *Digital Humanities Quarterly* (henceforth DHQ) [16].

Relevant terms for extracting computational information visualization solutions on assisting discourse or argumentation textual analysis are included as keywords in the queries, such as "discourse analysis", "argument mining", "information visualization", "software tool", etc. Note that, in order to deal with a manageable number of publications, we also add in this first phase of the SLR filters for the date of publication and publisher. Due to some repositories also acting as hubs of other repositories

publications (e.g., *ACM library*), it is necessary to limit the searches to their own resources in order to avoid duplications. Table 1 summarizes repositories, queries performed, and the number of initial results achieved.

Table 1. SLR repositories, search queries and the number of resultant publications.

Repository	Search Query	Number of Results
Springer Link	("discourse analysis" OR "argument mining") AND ("information visualization" OR "visualization" OR "visual analytics") AND ("software" OR "tool"); Filter 2010-2020	648
Science Direct	("discourse analysis" OR "argument mining") AND ("information visualization" OR "visualization" OR "visual analytics") AND ("software" OR "tool"); Filter 2010-2020	355
ACM Library	[[All: "discourse analysis"] OR [All: "argument mining"]] AND [[All: "information visualization"] OR [All: "visualization"] OR [All: "visual analytics"]] AND [[All: "software"] OR [All: "tool"] OR [All: ""]] AND [Publication Date: (01/01/2010 TO 12/31/2020)]; Filter ACM publisher	96
IEEE Xplore	('discourse AND analysis' OR 'argument mining') AND ('information AND visualization' OR 'visualization' OR 'visual analytics') AND ('software' OR 'tool'); Filter 2010-2020	24
ACL Anthology	("discourse analysis" OR "argument mining") AND ("information visualization" OR "visualization" OR "visual analytics") AND ("software" OR "tool")	298
RST repository	"software"	2
DSH Journal	(discourse analysis OR argument mining AND information visualization OR visual analytics AND software OR tool). Published: After January 2010	54
DHQ Journal	Query: ("discourse analysis" OR "argument mining") AND ("information visualization" OR "visualization" OR "visual analytics") AND ("software" OR "tool")	3
	TOTAL	1480

As Table 1 shows, the preliminary set of publications consists of 1480 works. Following the SLR guidelines, a set of inclusion/exclusion criteria are defined, focusing on original and recent works on information visualization implemented via software solutions for supporting discourse or argumentation analysis tasks. Thus, refinement criteria were applied as:

- The year of publication between 2010 and 2020.
- Original publications written in English language.
- Only original publications: papers in journals and full papers in conferences (also edited as chapters), excluding workshops.
- Only publications with scope on Computer Science (Software Engineering, Information Visualization and Computational Linguistics included) and Linguistics/Discourse/Argumentation-related areas.
- Only those publications that have associated original software/existing software use/demonstrator/ tools that provide visual support for discursive/argumentation textual analysis.

Applying these criteria, we obtained an intermediate reduced pool of 58 different publications. Finally, a step of quality assessment is applied to this intermediate set (see next section).

2.1.3. Quality Assessment

Due to the heterogeneity and diverse source of the publications included in the presented SLR, we systematically applied the following checklist as a quality assessment mechanism, answering yes (Y) or no (N) to the following questions:

- Q1: Are the study goals clearly stated and **related to textual analysis assistance**, and are the software proposals clearly detailed?
- Q2: Are the studies proposing an original software or an original application of existing software for assisting textual analysis **through visual software resources**?
- Q3: Is the proposal validated with **real text analysis cases**?
- Q4: Is the proposal dealing with textual discursive/argumentation analysis information?

- Q5: Is the proposal offering some software mechanisms for **promoting the reuse and sharing of the information generated** during their use?

Each publication of the intermediate set is evaluated applying the checklist (yes = 1 point, no = 0 point). Thus, each publication obtains a quality assessment score point (**Qn**) from Qn = 0 (minimum score) to Qn = 5 (maximum score).

We have decided that only publications that meet the following two quality requirements are finally included in the final set:

- Publications with Qn greater than 3, that is, publications with at least three affirmative answers to the quality questions.

- Publications with an affirmative answer to question 2. This implies that their contribution in assistance is through information visualization mechanisms, which is the main area of this study.

Table A1 shows the final quality assessment evaluation. Only publications marked with a fine gray color in the table are included in the SLR final set repository. After the quality assessment phase, SLR final set repository is composed of 17 publications. The next section presents and deeply discusses the SLR findings and the guidelines extracted.

3. Results

3.1. Systematic Literature Review Results

Multiple readings can be extracted from the SLR performed. In the first place, an important process of reducing the number of publications has been carried out, with a search results of 1480 publications, an intermediate SLR set of 58 publications and a final SLR repository of 17 publications. This refining process has been influenced by two aspects.

First, the selected keywords or combinations of them (multi-words) used in the search queries are commonly used in numerous disciplines, presenting polysemy in their use within the different research communities. Thus, the initial number of queries results is quite high for an SLR process, finding that many of the initially retrieved publications did not meet any of the quality assessment criteria proposed by our SLR process. Therefore, subsequent quality assessment process allowed us to focus on works that specifically proposed software solutions to assist in textual analysis on a discursive and argumentation level.

Secondly, many of the publications reviewed, although they were relevant applying keywords, did not cover the full scope of the work. These works are shown in a white background color in Table A1 and they correspond to:

- Fully automatic solutions and approaches, such as automatics parsers, automatic detection or prediction methods or tools, all of which are referred to in Table A1. Because our goal is focusing on software assistance tools, we have not included the works whose main contribution is based on a full automation of tasks, both at the level of detection of discursive or argumentation structures and at the level of automatic generation of visualizations that do not allow interaction of the end user.
- Approaches based on some pieces of information considered discursive but that do not respond to an analysis of the complete structure of the discourse, such as application of topic modeling, statistical studies (basic analysis of frequencies of terms or similar descriptive statistics), works in metaphors, stems, taxonomies or ontologies, all of which are referred to in Table A1. Many of them also adopt an automation approach.

We present below the synthesis of evidence from our SLR. We begin with a general analysis of the results from the final SLR set. Next, we present the answers to the research questions previously defined.

Availability

It is important to highlight the small number of free-use tools available throughout the study area. Only 12 publications from the SLR intermediate repository (58 publications) present access at least to

one demo for free or, in better cases, to software repositories or fully implemented free software tools (see last column in Table A1). Another small group of the rest could be accessed through institutional credentials, while most of them present a URL's with broken links in the publication, or they were never available online.

Visualization techniques employed

Regarding the visualization techniques used, most of them follow a principle of visualization based on the original text, many of which are based on discourse trees' [17] approaches or similar structures for argumentation. Some alternative proposals, such as conceptual recurrence plots [18] do not allow discursive or argumentation analysis based on the original text. The visualization is carried out after the textual analyses are completed and it is focused only on a specific metric (intervention turns, similarity between utterances, etc.).

Discursive and argumentation framework supported

Another interesting aspect extracting from the SLR is that most of the tools with a high quality score (score Qn = 5 or score Qn = 4) are particularly developed for supporting a specific framework of discursive or argumentation analysis, mainly RST—Rhetorical Structure Theory—[10] (such as [17,19]) or IAT—Inference Anchoring Theory—[20] (such as [21–23]), but also some ad hoc frameworks [24,25]. Although, conceptually, these frameworks present similar ontologies (segments and relationships between them) and similar visual possibilities, most of the tools are conceived to assist in textual analysis using a single discursive or argumentation framework. This causes a dependency between the software tool and the pragmatic framework chosen, since the user must employ that framework to carry out their textual analyses. The dependence on a discursive or argumentation framework is repeating in all proposals. Thus, it is not possible to extend the theoretical framework for creating similar textual analyses with customization in discursive or argumentation schemas.

Sharing and reuse software mechanisms

Finally, we have analyzed the sharing and reuse mechanisms of the final 17 publications. Although most of these software proposals are focusing on the collaborative edition and analysis, only 8 software tools present sharing and reuse functionality, while 9 software tools reviewed the lack of any sharing and reuse planification of the information generated during the textual analyses. The most used mechanism is based on information exportation in standard file formats, mainly XML –eXtensible Markup Language – or derived formats (such as JSON, JavaScript Object Notation) as an interchange mechanism. Only the *rstWeb* tool [19] presents extra mechanisms allowing for the better reuse of the resultant analysis information.

In summary, this SLR shows that the discursive and argumentation aspects continue to present few software alternatives for textual analysis assistance, in comparison with the wide range of proposals in the lexical or grammatical levels of textual analysis, or in fully automatic approaches at any level. The main weaknesses detected in the existing software proposals are the dependency of one specific discursive or argumentation framework (problems to generalize the analysis), availability (problems accessing, using and keeping the software updated) and the lack of sharing and reuse mechanisms that allow for re-analyses, collaborative editing, and comparative reasoning of the textual analysis performed. The following section elaborates on these results, answering the research questions initially proposed.

3.2. Answering the Research Questions

RQ1: What evidence is there that discourse and argumentation textual analysis is currently supported via information visualization software?

Appendix A shows the SLR final repository, with all publications reviewed, their scores and availability. The SLR allows us to answer the RQ1 and RQ2 of this study, regarding evidence about visualization techniques that are applied to support discourse or argumentation textual analysis via software. While the lexical or grammatical levels of textual analysis are gaining in methods and software tools for their assistance, the discursive and argumentation aspects continue to present few software alternatives. There are only a few recent tools [19,23–25] to assist in this textual analysis through an available information visualization software resource.

RQ2: What kind of support are these works providing and how is it implemented?

An interesting macro-analysis resultant of the SLR process corresponds to the kind of software support provided by the publications attending our categorization defined in RQ2. The Table A1 "Main contribution" column shows in bold the kind of software support contribution according to our seven categories defined for each publication. Note that, in some cases, one publication could present several contributions for different RQ2 categories, although the common scenario is that one publication focuses on one kind of main software support contribution. Thus, considering only the main category associated with each publication (the first category reported in Appendix A), the distribution of software support in the works reviewed is as follows:

With a total of 58 works reviewed, the majority category (23 publications) presents complete support in the form of a software tool, although the main objective and functionality of each tool may differ. Subsequently, the works present support in the form of InfoVis techniques (12 publications), application examples (seven publications), automatic solutions (seven publications), linguistic resources (four publications), qualitative studies (three publications) and finally the measurement of the basic discourse metrics (two publications). The large number of complete software tools reviewed offers us an idea of he current interest in software support in discursive and argumentation textual analysis. However, as we have already mentioned in the results of the SLR process, many of these tools present the weaknesses detected, and most of them are not even available for evaluation or use.

RQ3: Does the software support offered in these works present any weakness or deficiency reported in the study itself or detected as a result of the review?

Regarding RQ3, we found some weaknesses in the current software tools, especially related to availability, framework independence and a lack of information sharing and reuse software mechanisms. Based on these weaknesses, four software guidelines are defined and exemplified in implementation through *Viscourse*. (See Section 4).

RQ4: Is it possible to identify some software guidelines for improving the existing information visualization software solutions for supporting or assisting discourse and argumentation textual analysis tasks?

Answering RQ4, *Viscourse* tries to act as a complementary software tool, allowing for the generalization of the discursive or argumentation analyzes thanks to their flexibility and independence from the segmentation method (free definition of segments or multi-phrase groups) or of the theoretical framework (RST or IAT, among others), also with a user-customizable visualization criteria (screen position, color palette, color criteria, etc.). Typical users of *Viscourse* can be mainly researchers in humanistic disciplines, but also teachers or students of discourse or argumentation areas.

In addition, *Viscourse* focuses on the sharing and reuse of the analytical information generated, thanks to a black-box mechanism that allows for the encapsulation of import/export formats to researchers, without losing sharing and reused capacity. The next section details the guidelines extracted and the solutions proposed as part of Viscourse implementation for the guidelines' implementation.

4. Extracted Guidelines in Practice

Considering the SLR results, we have defined a non-exclusive set of software guidelines for improving the existing information visualization software assistance provided to discursive and argumentation textual analysis:

1. Textual granularity: Visual mechanisms should be added to change the level of granularity of the textual analysis. This means that the user must be able to change the visual focus of the analysis, being able to focus on the specific text paragraphs, phrases or other textual segments, or to raise the level of abstraction, calculating general metrics for the full text analyzed.

2. Linguistic framework flexibility: Software mechanisms should be developed to allow an independence between the visual mechanisms and the specific discursive or argumentation framework used for the textual analysis, allowing for the extension of the software tools to future discursive or argumentation frameworks. Some guidelines here include separate conceptual modelling strategies for the visual solution and each specific framework applied for each analysis performed.

3. Sharing and reuse mechanisms: Software mechanisms should be developed for allowing the sharing and reuse of the resultant informational pieces for the textual analyses in a transparent way for end users. These mechanisms are particularly useful both in future analysis by the same users and in a collaborative or comparative analysis by other researchers. Some guidelines here include black-box and transparent export/import mechanisms for the informational pieces produced during the textual analyses.

4. Availability alternatives: The software assistance provided should offer some availability and maintenance solutions. This does not necessarily imply free or open models of all the software tools, but rather a prior planning of availability mechanisms so that the learning effort made by users is rewarded.

How these guidelines could be implemented in a real software solution for assisting discursive and argumentation textual analysis? Taking into account previous works on discourse and argumentation studies via software [26–28], *Viscourse* tool [29] is a web platform for discursive/argumentation textual analysis. We detail in the next sub-sections our approach for implementing the guidelines presented above in *Viscourse*.

4.1. Textual Granularity

As a tool for supporting discursive/argumentation textual analysis, *Viscourse* is initially conceived for an analysis attached to the text. This implies that the user details the literal text to be analyzed, segmenting it and performing the analysis at the discursive or argumentative level. Although this is the most common use, *Viscourse* includes some mechanisms related to guideline 1 for ensuring the software support at different degrees of textual granularity. Specifically, two mechanisms have been implemented related to this guideline:

First, the user can vary the level of granularity of the text by grouping the textual segments into larger units, called groups. This allows one to associate discursive or argumentation characteristics to large sets of text, even entire paragraphs.

Secondly, it is possible to hide the literal text of the analysis performed, showing the option "Simplified mode". At that time, all segments and groups acquire the same size and alignment, hiding the literal text to analyze and keeping only the name of the segment or group. This simplified view allows the user to abstract from the literal text, focusing on the discursive or argumentation relationships detected and facilitating the comparison between several analyses on the same text or the calculation of simple metrics (number of discursive relationships of a specific type, etc.). In the future, we plan to offer this automatic metrics calculation and comparative functionality implemented as a module in *Viscourse*. Figures 2 and 3 show the same textual analysis ("Bouquets in the basket" text form RST corpus) with different levels of granularity.
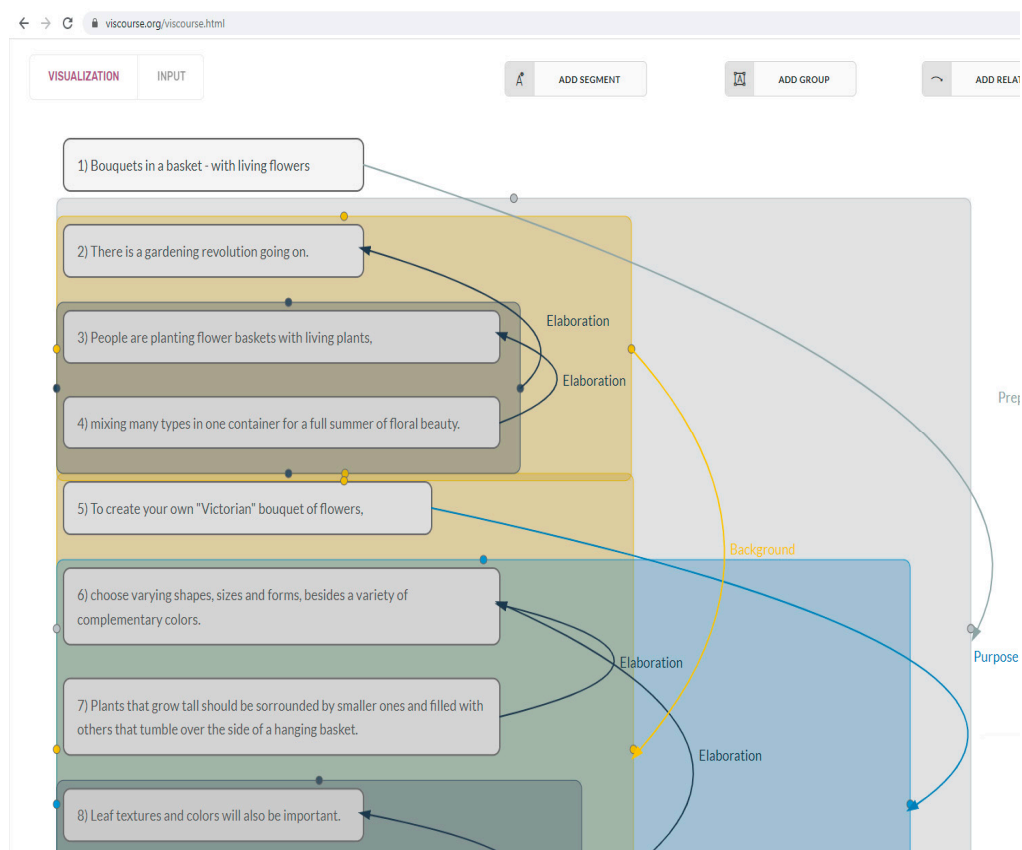
Figure 2. Viscourse textual analysis for the "Bouquets in a basket" Rhetorical Structure Theory (RST) corpus text.
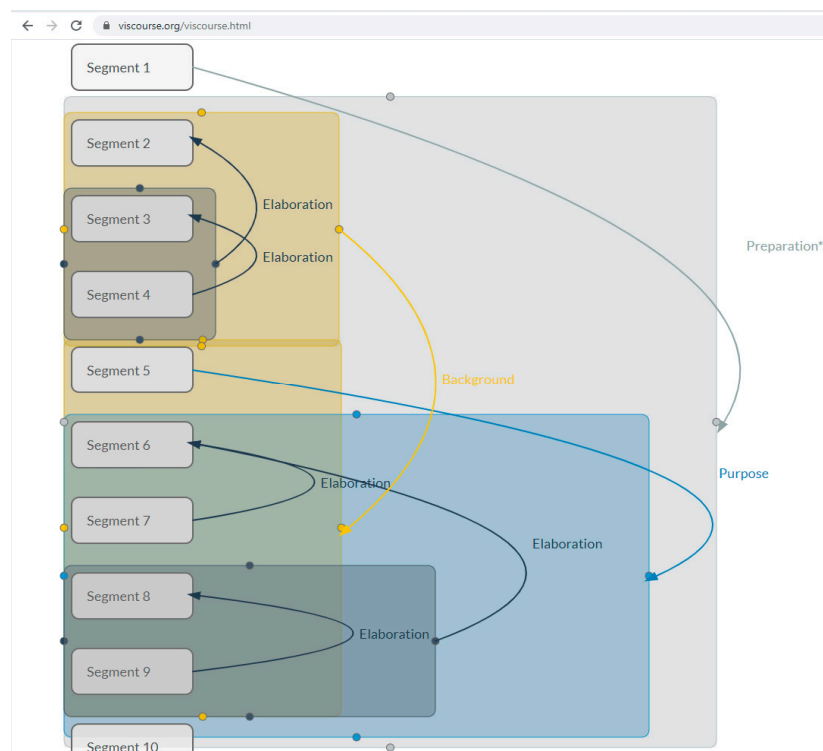
Figure 3. Viscourse textual analysis for the bouquets in the basket RST corpus text. Simplified mode.

4.2. Linguistic Framework Flexibility

Our SLR showed that most of the current tools are particularly developed for supporting a specific framework of discursive or argumentation analysis, such as *rstWeb* in RST [10] or *OVA* in IAT [20]. Despite the theoretical and linguistic approach differences between them, most of these frameworks base the discursive or argumentation analysis on the same needs in terms of software visualization structures: textual segmentation or grouping, and the definition of discursive or argumentation relationships between segments (or similar textual elements with greater granularity). Thus, guideline 2 defends that a software tool that supports this type of textual analysis should maintain a certain independence between the discursive/argumentation framework used during a certain textual analysis and the visual elements employed in the tool. This conceptual decoupling at the software tool level allows the future extension of support to perform analysis with future discursive or argumentation frameworks that may arise.

For implementing guideline 2, *Viscourse* initial version decouples visual mechanisms and discursive/argumentation information framework used. First, *Viscourse* allows for textual analysis based on the customized segmentation of the text by the user and its grouping into wider levels that also allow for analysis on a multi-phrase or paragraph level. *Viscourse* currently allows for the analysis of relationships between segments or multi-phrase groups, with highly customized features for the user, so any discursive and/or argumentation relationships framework can be used.

In addition, on a visualization level, the text follows existing approaches as discourse trees but adding color visualization for segments or multi-phrase groups. The color of the groups and the discursive or argumentation relationships during the analysis are also selectable by the user, which allows him to apply color criteria (it is common, for example, to use red for discursive and/or argumentation relationships with contrast or disagreement semantics). It is also possible to highlight only a specific type of relation or segment in the tool for visual clarification.

Figure 2 shows an example of the well-known bouquets in a basket example from the RST corpus [10] performed using *Viscourse*. *Viscourse* natively implements the RST analysis [11,14] due to the applicability of the platform to ongoing projects. Besides, thanks to the customization possibilities in terms of segments, groups and relationships definition, it is possible to perform multiple analyses using different discursive or argumentation frameworks as a basis.

4.3. Sharing and Reuse Mechanisms

As our SLR showed, a few cases of the revised tools have automatic mechanisms for exporting the information generated during the analysis [19,21,23–25], which allows some initial steps on sharing and reuse of this information. Most of these works use mechanisms that require knowledge of certain file exchange formats, in most cases requiring that the researcher must also edit the generated files. *Viscourse* also includes the classic import/export mechanisms through editable standard file formats from the web platform (Figure 4 shows the JSON editor included in *Viscourse*, with three options for file view: code, tree or a node view). This mechanism transforms *Viscourse* in a complete JSON editor for the discourse-based information produced during a textual analysis. Note that the tool allows the user to save multiple analyses in their user account, maintaining a visualization carrousel for each user. Each analysis performed is generating an underlined JSON file with information about the original textual sources, the segments, groups, relationships, etc., created by the user and color and position information that allow the software to replicate and export each analysis.
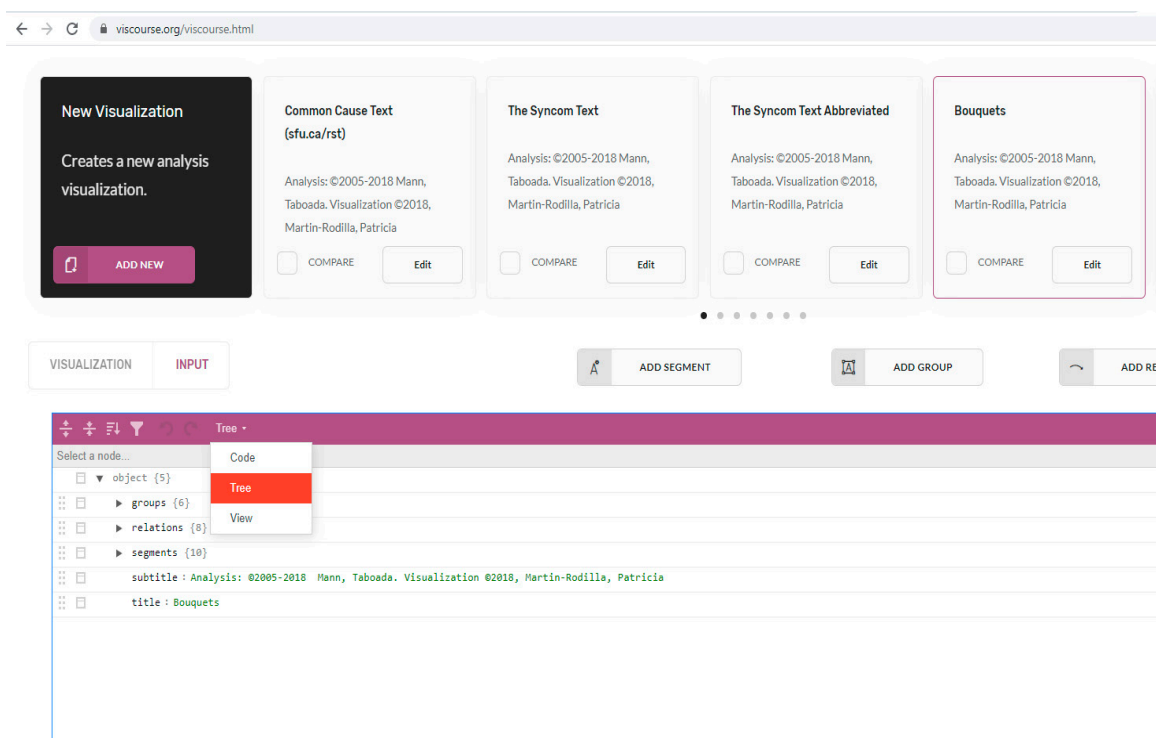
Figure 4. *Viscourse* classic import/export mechanisms through editable JavaScript Object Notation (JSON) files from the *Viscourse* web platform, with three options for JSON file view: code, tree or a node view.

However, in our experience, many of the analyses performed are reused by the researcher himself, or shared with colleagues for comparison or communication purposes, and many of these researchers do not need to edit the files generated by the applications for this purpose. For this reason, *Viscourse* also includes, following guideline 3, an import/export mechanism as a "black box" for the researcher, through its own *Viscourse* code mechanism. The implementation consists of encapsulating all the information produced during a textual analysis (information about the original text, the discursive or argumentation information produced during the textual analysis and the visual decisions about colors, positions, etc., taken by the user during the analysis) in a black box piece associated with a unique code automatically generated.

The black-box import/export mechanism in *Viscourse* follows this workflow:

- Once the textual analysis is finished, the user selects "share selected visualization" in the Options menu. An export message is shown. The user can generate an internal code for *Viscourse* that matches the JSON file created for the textual analysis with exactly their visualization and textual analysis parameters.
- The code generated is shown to the user, with automatic copy options. Sharing the *Viscourse* code with any *Viscourse* user, it is possible to import the textual analysis in any web *Viscourse* session.
- For importing, the user selects the "Import from Code" option near their visualization carrousel in the main screen. Pasting the *Viscourse* code is enough to reproduce the textual analysis performed and all their visualization options in other *Viscourse* sessions or user accounts.

Figure 5 visually summarizes this workflow, showing the different actions in the screen performed by the user in an exporting/importing operation.
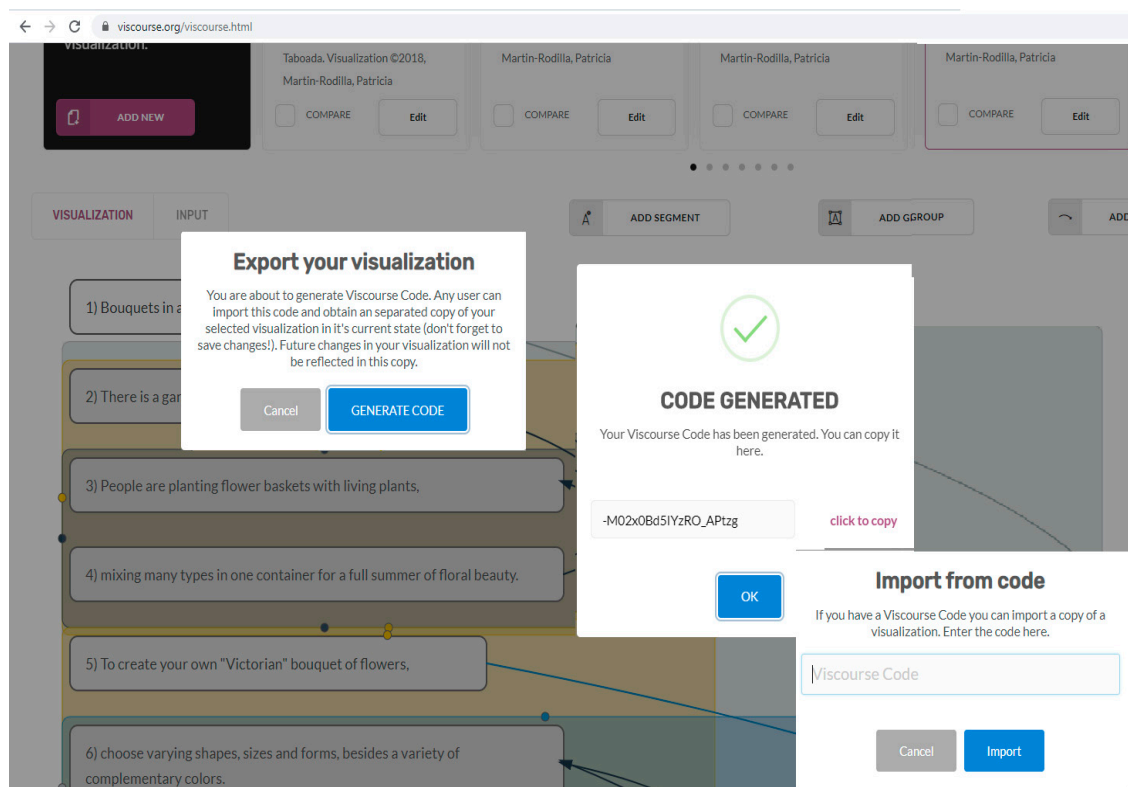
Figure 5. *Viscourse* code, a black-box mechanism for sharing and reusing textual analysis information between users.

4.4. Availability Alternatives

As previously detailed, one of the weaknesses that our SLR analysis points out is the absence of tools currently available to the end user. Although it is not the objective of the paper to go into the many reasons why this can happen, we think that, as guideline 4 details, the prior planning of the software maintenance model and availability of the different tools to support discursive and argumentation textual analysis is necessary. Currently, we are conducting this prior analysis to provide *Viscourse* with different availability and use alternatives: free use, research licenses, free trials, subscription payment, availability to users in code repositories, etc. There are several alternatives to ensure that the work done is effectively available to its end users.

5. Future Steps

As it has been shown throughout this paper, the development of software assistance for textual analysis through information visualization techniques presents great challenges. It is also an area of vital importance for humanistic and social disciplines, where researchers present special needs in terms of textual analysis assistance as compared to other disciplines. Thus, the systematic literature review carried out here is a valuable contribution to the field, including heterogeneity references in terms of source repositories, disciplines and approaches and presenting a broad panorama of the area.

While we believe that the steps taken so far are satisfactory, future steps are planned. Our immediate work includes a formal validation of the *Viscourse* tool and their guidelines on implementation. It is necessary to evaluate the *Viscourse* software tool on a wide set of case studies, by its target users. An evaluation working with researchers in various branches of the digital humanities is planned, in order to identify weaknesses and strengths of the tool, obtaining assessment valorization by end users.

Regarding the connection of the *Viscourse* tool with the current works revised here, future work includes exploring the possible connection between *Viscourse* visualization, sharing and reuse

mechanisms and current natural language processing (NLP) algorithms and approaches to the automatic extraction of linguistic information on a pragmatic level. The flexibility of *Viscourse* in terms of the discursive or argumentation framework used and their black-box sharing and reuse capabilities allow us to explore the use of *Viscourse* as a complementary piece of software for visualizing and interacting with outputs from different NLP discursive parsers.

Besides, the following steps in *Viscourse* tool improvements will focus on the implementation of a comparative visualization solution, which allows for comparing several analyses of the same text from different frameworks, or performed by different users, in the same interface. Adding comparative features to *Viscourse* will further enhance *Viscourse's* reuse and sharing capabilities.

Appendix A

Table A1. Quality assessment SLR phase results, evaluating 59 publications (58 publications eliminating replications) from 7 different source repositories. Gray marked rows constitute publications included in the final SLR repository (17 publications). "Main contribution" column shows in bold the kind of software support contribution according to our seven categories (defined in RQ2) for each publication.

Source-SLRCode	Title	Main Contribution	Q1	Q2	Q3	Q4	Q5	Qn	Availability
Springer Link-SL1 [30]	A survey on information visualization: recent advances and challenges	**InfoVis techniques** for textual analysis: Discourse trees	Y	Y	Y	Y	N	4	N
Springer Link-SL2 [31]	Towards Automatic Argument Extraction and Visualization in a Deliberative Model of Online Consultations for Local Governments	Annotated **corpora** on politics. Argumentation analysis.	N	Y	Y	Y	N	3	N
Springer Link-SL3 [21]	Argumentation in the 2016 US presidential elections: annotated corpora of television debates and social	OVA **tool** application example for supporting argumentation analysis.	Y	Y	Y	Y	Y	5	Y
Springer Link-SL4 [23]	The Argument Web: an Online Ecosystem of Tools, Systems and Services for Argumentation	Ova **tool** + Argument Analytics tool for supporting argumentation analysis.	Y	Y	Y	Y	Y	5	Y
Springer Link-SL5 [32]	SAPTE: A multimedia information system to support the discourse analysis and information retrieval of television programs	SAPTE **tool** for TV domain. Some discourse analysis **metrics**.	Y	N	Y	Y	N	3	N
Springer Link-SL6 [33]	Text to multi-level MindMaps	**InfoVis techniques** for textual analysis: **Manual** Mind maps	Y	N	Y	N	N	2	N
Springer Link-SL7 [34]	Knowledge Building Discourse Explorer: a social network analysis application for knowledge building discourse	KBDeX **tool**. Some discourse analysis **metrics**.	Y	Y	N	Y	N	3	Y
Springer Link-SL8 [35]	PolyCAFe—automatic support for the polyphonic analysis of CSCL chats	PolyCAFe **tool** for automatic analysis of conversations: learning analytics domain.	Y	Y	Y	N	Y	4	N
Springer Link-SL9 [36]	PolyCAFe - Polyphonic Conversation Analysis and Feedback	PolyCAFe **tool** for automatic analysis of conversations.	Y	Y	Y	N	Y	4	N

Table A1. *Cont.*

Source-SLRCode	Title	Main Contribution	Q1	Q2	Q3	Q4	Q5	Qn	Availability
Springer Link-SL10 [37]	Facilitating the Analysis of Discourse Phenomena in an Interoperable NLP Platform	U-Compare **tool** for NLP workflows construction: automatic discourse extensions.	Y	N	Y	Y	N	**3**	N
Springer Link-SL11 [38]	Computer assisted text analysis in the social sciences (Alceste tool)	Alceste **tool**. Some discourse analysis metrics.	Y	Y	N	N	N	**2**	N
Springer Link-SL12 [39]	Mass Collaboration on the Web: Textual Content Analysis by Means of Natural Language Processing	NLP **application example** on mass collaboration domain.	Y	N	N	N	N	**1**	N
Science Direct-SD1 [40]	Towards computational discourse analysis: A methodology for mining Twitter backchanneling conversations	Methodology and **application example** on automatic concept map creation from Twitter data.	Y	Y	Y	N	N	**3**	N
Science Direct-SD2 [41]	MARGOT: A web server for argumentation mining	MARGOT **tool** for automatic argumentation textual analysis.	Y	N	Y	Y	N	**3**	Y
Science Direct-SD3 [42]	Using visual text analytics to examine broadcast interviewing	**InfoVis techniques** for textual analysis: Conceptual Recurrence Plots	Y	Y	Y	Y	N	**4**	N
ACM Lib.-ACM1 [43]	Visual analytics of academic writing	XIP **tool** for automatic textual analysis and application example on scientific discourse	Y	N	N	N	N	**1**	N
ACM Lib.-ACM2 [44]	Temporal analytics with discourse analysis: tracing ideas and impact on communal discourse	Some discourse analysis **metrics**.	Y	N	Y	N	N	**2**	N
ACM Lib.-ACM3 [45]	Humor, support and criticism: a taxonomy for discourse analysis about political crisis on Twitter	**Taxonomy** on politics.	Y	N	Y	N	N	**2**	N
ACM Lib.-ACM4 [25]	Discourse-centric learning analytics	Cohere **tool** for automatic analysis of discourse.	Y	Y	Y	Y	Y	**5**	Y
ACM Lib.-ACM5 [46]	Experiments in automated support for argument reconstruction	**Automatic** topic modelling and argumentation experiments.	N	N	Y	Y	N	**2**	N
ACM Lib.-ACM6 [47]	Highly interactive and natural user interfaces: enabling visual analysis in historical lexicography	**InfoVis techniques** for textual analysis.	Y	Y	Y	N	N	**3**	N
ACM Lib.-ACM7 [22]	Using Argumentative Structure to Interpret Debates in Online Deliberative Democracy and ERulemaking	**Application example** for supporting argumentation analysis on politics	Y	N	Y	Y	N	**3**	N
ACM Lib.-ACM8 [48]	ThemeStreams: visualizing the stream of themes discussed in politics	**InfoVis techniques** for textual analysis: ThemeStreams	Y	Y	Y	N	N	**3**	N
ACM Lib.-ACM9 [49]	Web-Retrieval Supported Argument Space Exploration	**Automatic** information retrieval methods for argumentation analysis.	Y	Y	Y	N	N	**3**	N
ACM Lib.-ACM10 [50]	Visualizing Natural Language Descriptions: A Survey	**Survey** on graphical systems for natural language support.	Y	N	N	N	N	**1**	N
ACM Lib.-ACM11 [51]	Marius, the giraffe: a comparative informatics case study of linguistic features of the social media discourse	**Application example** on social media. Some discourse analysis **metrics**.	Y	Y	N	N	N	**2**	N
ACM Lib.-ACM12 [52]	Single or Multiple Conversational Agents?: An Interactional Coherence Comparison	**Application example** on chatbots. Some discourse analysis **metrics**.	Y	N	Y	N	N	**2**	N
ACM Lib.-ACM13 [53]	Analyzing Wikipedia Deletion Debates with a Group Decision-Making Forecast Model	**Automatic** machine learning technique. **Application example** on debates.	N	N	Y	N	N	**1**	N
IEEE Xplore-IX1 [54]	Current Work Practice and Users' Perspectives on Visualization and Interactivity in Business Intelligence	Qualitative **empirical study** on InfoVis + Business Intelligence uses.	Y	N	N	N	N	**1**	N

Table A1. *Cont.*

Source-SLRCode	Title	Main Contribution	Q1	Q2	Q3	Q4	Q5	Qn	Availability
IEEE Xplore-IX2 [55]	Visualization of Sensory Perception Descriptions	Wine Fingerprints + Topics2Themes InfoVis **tools**. Sentiment Analysis and Topic Modelling applications.	Y	Y	Y	Y	N	**4**	Y
IEEE Xplore-IX3 [18]	Conceptual Recurrence Plots: Revealing Patterns in Human Discourse	**InfoVis techniques** for textual analysis: Conceptual Recurrence Plots	Y	Y	Y	Y	N	**4**	N
IEEE Xplore-IX4 [56]	Visual unrolling of network evolution and the analysis of dynamic discourse	**InfoVis technique:** DCRA visualization prototype	Y	Y	Y	N	N	**3**	N
IEEE Xplore-IX5 [57]	A survey on computer assisted qualitative data analysis software	**Survey** on data analysis software.	Y	N	N	N	N	**1**	N
IEEE Xplore-IX6 [58]	A Tool for Discourse Analysis and Visualization	**Tool** for supporting discourse analysis.	Y	Y	Y	Y	N	**4**	N
IEEE Xplore-IX7 [59]	Robust Adaptive Discourse Parsing for E-Learning Fora	Agora **tool** for **automatic** contrast parsing on Internet forums	Y	N	Y	Y	N	**3**	N
IEEE Xplore-IX8 [60]	Assessing Collaborative Process in CSCL with an Intelligent Content Analysis Toolkit	Some discourse analysis **metrics**.	Y	N	Y	N	N	**2**	N
IEEE Xplore-IX9 [61]	Epicurus: A platform for the visualisation of forensic documents based on a linguistic approach	Epicurus **tool** for supporting discourse analysis.	Y	Y	Y	Y	N	**4**	N
IEEE Xplore-IX10 [62]	Text cohesion visualizer	Text cohesion **tool** for InfoVis techniques. Some discourse analysis **metrics**.	Y	Y	Y	Y	N	**4**	N
IEEE Xplore-IX11 [63]	A Pilot Study of CZTalk: A Graphical Tool for Collaborative Knowledge Work	**InfoVis techniques:** graph visualizations for discourse.	Y	N	Y	N	N	**2**	N
IEEE Xplore-IX12 [64]	The competency building process of human computer interaction in game-based teaching: Adding the flexibility of an asynchronous format	**Application example** on Massively Multiplayer Online Games (MMOG) domain. Some discourse analysis **metrics**.	N	N	Y	N	N	**1**	N
ACL Anth.-ACL1 [65]	ArguminSci: A Tool for Analyzing Argumentationand Rhetorical Aspects in Scientific Writing	ArguminSci **tool: Automatic** argumentation and discourse parsing.	Y	Y	Y	Y	N	**4**	Y
ACL Anth.-ACL2 [66]	Two Practical Rhetorical Structure Theory Parsers	**Automatic** discourse parsers.	Y	N	Y	Y	N	**3**	Y
ACL Antl.-ACL3 [67]	Capturing Chat: Annotation and Tools for Multiparty Casual Conversation	**Infovis technique** + STAVE **tool** for conversational analysis.	Y	N	Y	N	N	**2**	N
ACL Antl.-ACL4 [68]	Interactive Exploration of Asynchronous Conversations: Applying a User-centered Approach to Design a Visual Text Analytic System	**Infovis techniques** for conversational analysis.	Y	N	N	N	N	**1**	N
ACL Anth.-ACL5 [19]	rstWeb–A Browser-based Annotation Interface for Rhetorical Structure Theory and Discourse Relations	rstWeb **tool** for supporting discourse analysis.	Y	Y	Y	Y	Y	**5**	Y
ACL Anth.-ACL6 [69]	Tree Annotator: Versatile Visual Annotation of Hierarchical Text Relations	Tree Annotator **tool**: graphical tool for annotating tree-like structures	Y	Y	Y	Y	Y	**4**	Y
ACL Anth.-ACL7 [70]	The Impact of Modeling Overall Argumentation with Tree Kernels	**Automatic** representation methodology for argumentation	Y	N	Y	Y	N	**3**	N
ACL Anth.-ACL8 [24]	iLCM - A Virtual Research Infrastructure for Large-Scale Qualitative Data	iLCM **tool** for discourse analysis.	Y	Y	Y	Y	Y	**5**	Y
RST-RST1 [71]	The GUM Corpus: Creating Multilayer Resources in the Classroom	Annotated **corpora** on education. RST discourse analysis.	Y	Y	Y	N	N	**3**	Y
RST -Duplicated	*rstWeb – A Browser-based Annotation Interface for Rhetorical Structure Theory and Discourse Relations*	-	-	-	-	-	-	**-**	-

Table A1. *Cont.*

Source-SLRCode	Title	Main Contribution	Q1	Q2	Q3	Q4	Q5	Qn	Availability
DSH Journal-DSH1 [72]	PaperMiner—a real-time spatiotemporal visualization for newspaper articles	**InfoVis techniques**. **Application example** on newspapers. Some discourse **metrics**.	Y	Y	Y	N	N	3	N
DSH Journal-DSH2 [73]	Mining ethnicity: Discourse-driven topic modelling of immigrant discourses in the USA	**Automatic** topic modelling. **Application example** on historical texts.	Y	N	Y	N	N	2	N
DSH Journal-DSH3 [74]	Exploratory Thematic Analysis for Digitized Archival Collections	TOME **tool**: **Automatic** topic modelling. **Application example** on historical texts.	Y	N	Y	N	N	2	N
DSH Journal-DSH4 [75]	Non-representational approaches to modeling interpretation in a graphical environment	**InfoVis techniques** for textual analysis.	N	Y	N	N	N	1	N
DSH Journal-DSH5 [76]	Supporting exploratory text analysis in literature study	**Application example** on literature. Some discourse analysis **metrics**.	Y	N	Y	N	N	2	N
DSH Journal-DSH6 [77]	Non-traditional prosodic features for automated phrase break prediction	**Automatic** phrase break prediction review.	N	N	Y	N	N	1	N
DSH Journal-DSH7 [78]	Analysis of variation significance in artificial traditions using Stemmaweb	Stemmaweb **tool** for stemmatology.	Y	N	Y	N	N	2	N
DSH Journal -DSH8 [79]	Networks of networks: a citation network analysis of the adoption, use, and adaptation of formal network techniques in archaeology	**Automatic** network analysis techniques. **Application examples** on archaeology.	Y	N	Y	N	N	2	N
DSH Journal-DSH9 [80]	Ontology-based analysis of the large collection of historical Hebrew manuscripts	**Manual** Ontology analysis. **Application example** on ancient texts.	N	Y	Y	N	N	2	Y
dhq Journal-dhq1 [81]	A Pedagogy for Computer-Assisted Literary Analysis: Introducing GALGO (Golden Age Literature Glossary Online)	**Taxonomy**-Glossary resource.	Y	N	Y	N	N	2	N

References

1. Schreibman, S.; Siemens, R.; Unsworth, J. *A New Companion to Digital Humanities*; John Wiley & Sons: Chichester, UK, 2016.

2. Kucher, K.; Kerren, A. (Eds.) Text visualization techniques: Taxonomy, visual survey, and community insights. In Proceedings of the 2015 IEEE Pacific Visualization Symposium (PacificVis), Hangzhou, China, 14–17 April 2015.

3. Alharbi, M.; Laramee, R.S. Sos textvis: An extended survey of surveys on text visualization. *Computers* **2019**, *8*, 17. [CrossRef]

4. Kitchenham, B.; Charters, S. *Guidelines for Performing Systematic Literature Reviews in Software Engineering*; EBSE Report No. 2007-01; Durham University: Durham, UK, 2007.

5. Elsevier. ScienceDirect® Elsevier, B.V. 2020. Available online: https://www.sciencedirect.com/ (accessed on 7 May 2020).

6. Springer. Springer Link. Springer Nature Switzerland AG. 2020. Available online: https://link.springer.com/ (accessed on 7 May 2020).

7. ACM. ACM Digital Library. Association for Computing Machinery. 2020. Available online: https://dl.acm.org/ (accessed on 7 May 2020).

8. IEEE. IEEE Xplore. 2020. Available online: https://ieeexplore.ieee.org/Xplore/home.jsp (accessed on 7 May 2020).

9. ACL. ACL Anthology. Association for Computational Linguistics (ACL). 2020. Available online: https://www.aclweb.org/anthology/ (accessed on 7 May 2020).

10. Mann, W.C.; Taboada, M. RST—Rhetorical Structure Theory 2005–2018. Available online: https://www.sfu.ca/rst/06tools/index.html (accessed on 7 May 2020).

11. Mann, W.C.; Thompson, S.A. Rhetorical structure theory: Toward a functional theory of text organization. *J. Study Discourse* **1988**, *8*, 243–281. [CrossRef]

12. Mann, W.C.; Thompson, S.A. *Rhetorical Structure Theory: A Theory of Text Organization*; Information Sciences Institute: Los Angeles, CA, USA, 1987.

13. Taboada, M.; Mann, W.C. Rhetorical structure theory: Looking back and moving ahead. *Discourse Stud.* **2006**, *8*, 423–459. [CrossRef]

14. Taboada, M.; Mann, W.C. Applications of rhetorical structure theory. *Discourse Stud.* **2006**, *8*, 567–588. [CrossRef]

15. DSH. Digital Scholarship in the Humanities. Oxford University Press. 2020. Available online: https://academic.oup.com/dsh (accessed on 7 May 2020).

16. DHQ. Digital Humanities Quarterly. Association for Computers and the Humanities (ACH) and the Alliance of Digital Humanities Organizations (ADHO). 2020. Available online: http://digitalhumanities.org/dhq/about/about.html (accessed on 7 May 2020).

17. Zhao, J.; Chevalier, F.; Collins, C.; Balakrishnan, R. Facilitating discourse analysis with interactive visualization. *IEEE Trans. Vis. Comput. Graph.* **2012**, *18*, 2639–2648. [CrossRef] [PubMed]

18. Angus, D.; Smith, A.; Wiles, J. Conceptual recurrence plots: Revealing patterns in human discourse. *IEEE Trans. Vis. Comput. Graph.* **2011**, *18*, 988–997. [CrossRef]

19. Zeldes, A. (Ed.) rstWeb—A browser-based annotation interface for Rhetorical Structure Theory and discourse relations. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations, San Diego, CA, USA, 12–17 June 2016.

20. Budzynska, K.; Reed, C. *Whence Inference? Technical Report*; University of Dundee: Dundee, UK, 2011.

21. Visser, J.; Konat, B.; Duthie, R.; Koszowy, M.; Budzynska, K.; Reed, C. Argumentation in the 2016 US presidential elections: Annotated corpora of television debates and social media reaction. *Lang. Resour. Eval.* **2019**, *54*, 123–154. [CrossRef]

22. Lawrence, J.; Park, J.; Budzynska, K.; Cardie, C.; Konat, B.; Reed, C. Using argumentative structure to interpret debates in online deliberative democracy and erulemaking. *ACM Trans. Internet Technol.* **2017**, *17*, 1–22. [CrossRef]

23. Reed, C.; Budzynska, K.; Duthie, R.; Janier, M.; Konat, B.; Lawrence, J.; Pease, A.; Snaith, M. The argument web: An online ecosystem of tools, systems and services for argumentation. *Philos. Technol.* **2017**, *30*, 137–160. [CrossRef]

24. Niekler, A.; Bleier, A.; Kahmann, C.; Posch, L.; Wiedemann, G.; Erdogan, K.; Heyer, G.; Strohmaier, M. iLCM-A Virtual Research Infrastructure for Large-Scale Qualitative Data. *arXiv*, 2018; arXiv:80511404.

25. De Liddo, A.; Shum, S.B.; Quinto, I.; Bachler, M.; Cannavacciuolo, L. (Eds.) Discourse-centric learning analytics. In Proceedings of the 1st International Conference on Learning Analytics and Knowledge, Banff, AL, Canada, 27 February–1 March 2011.

26. Martín-Rodilla, P.; Gonzalez-Perez, C. (Eds.) An ISO/IEC 24744-derived modelling language for discourse analysis. In Proceedings of the IEEE 8th International Conference on Research Challenges in Information Science (RCIS), Marrakech, Morocco, 28–30 May 2014.

27. Gamallo, P.; Martín-Rodilla, P.; Calderón, B. (Eds.) Identifying Causal Relations in Legal Documents with Dependency Syntactic Analysis. In Proceedings of the 8th Symposium on Languages, Applications and Technologies (SLATE 2019), Coimbra, Portugal, 27–28 June 2019.

28. Martin-Rodilla, P. *Digging into Software Knowledge Generation in Cultural Heritage*; Springer: Cham, Switzerland, 2018; ISBN 978-3-319-69187-9.

29. Martin-Rodilla, P.; Sanchez, M. Viscourse. 2019. Available online: https://viscourse.org/ (accessed on 7 May 2018).

30. Liu, S.; Cui, W.; Wu, Y.; Liu, M. A survey on information visualization: Recent advances and challenges. *Vis. Comput.* **2014**, *30*, 1373–1393. [CrossRef]

31. Bembenik, R.; Andruszkiewicz, P. (Eds.) Towards automatic argument extraction and visualization in a deliberative model of online consultations for local governments. In Proceedings of the East European Conference on Advances in Databases and Information Systems, Prague, Czech Republic, 28–31 August 2016.

32. Pereira, M.H.; de Souza, C.L.; Pádua, F.L.; Silva, G.D.; de Assis, G.T.; Pereira, A.C. SAPTE: A multimedia information system to support the discourse analysis and information retrieval of television programs. *Multimed. Tools Appl.* **2015**, *74*, 10923–10963. [CrossRef]

33. Elhoseiny, M.; Elgammal, A. Text to multi-level MindMaps. *Multimed. Tools Appl.* **2016**, *75*, 4217–4244. [CrossRef]

34. Oshima, J.; Oshima, R.; Matsuzawa, Y. Knowledge Building Discourse Explorer: A social network analysis application for knowledge building discourse. *Educ. Technol. Res. Dev.* **2012**, *60*, 903–921. [CrossRef]

35. Trausan-Matu, S.; Dascalu, M.; Rebedea, T. PolyCAFe—Automatic support for the polyphonic analysis of CSCL chats. *Int. J. Comput.-Supported Collab. Learn.* **2014**, *9*, 127–156. [CrossRef]

36. Dascalu, M. PolyCAFe-Polyphonic Conversation Analysis and Feedback. In *Analyzing Discourse and Text Complexity for Learning and Collaborating*; Springer: Cham, Switzerland, 2014; pp. 107–135.

37. Batista-Navarro, R.T.; Kontonatsios, G.; Mihăilă, C.; Thompson, P.; Rak, R.; Nawaz, R.; Korkontzelos, I.; Ananiadou, S. Facilitating the analysis of discourse phenomena in an interoperable NLP platform. In Proceedings of the International Conference on Intelligent Text Processing and Computational Linguistics CICling 2013, Samos, Greece, 24–30 March 2013.

38. Brier, A.; Hopp, B. Computer assisted text analysis in the social sciences. *Qual. Quant.* **2011**, *45*, 103–128. [CrossRef]

39. Habernal, I.; Daxenberger, J.; Gurevych, I. Mass Collaboration on the Web: Textual Content Analysis by Means of Natural Language Processing. In *Mass Collaboration and Education*; Springer: Cham, Switzerland, 2016; pp. 367–390.

40. Lipizzi, C.; Dessavre, D.G.; Iandoli, L.; Marquez, J.E.R. Towards computational discourse analysis: A methodology for mining twitter backchanneling conversations. *Comput. Hum. Behav.* **2016**, *64*, 782–792. [CrossRef]

41. Lippi, M.; Torroni, P. MARGOT: A web server for argumentation mining. *Expert Syst. Appl.* **2016**, *65*, 292–303. [CrossRef]

42. Angus, D.; Fitzgerald, R.; Atay, C.; Wiles, J. Using visual text analytics to examine broadcast interviewing. *Discourse Context Media* **2016**, *11*, 38–49. [CrossRef]

43. Simsek, D.; Shum, S.B.; De Liddo, A.; Ferguson, R.; Sándor, Á. (Eds.) Visual analytics of academic writing. In Proceedings of the 4th International Conference on Learning Analytics and Knowledge, Indianapolis, IN, USA, 24–28 March 2014.

44. Lee, A.V.Y.; Tan, S.C. (Eds.) Temporal analytics with discourse analysis: Tracing ideas and impact on communal discourse. In Proceedings of the 7th International Learning Analytics & Knowledge Conference, Vancouver, BC, Canada, 13–17 March 2017.

45. Teixeira, C.R.G.; Kurtz, G.; Leuck, L.P.; Tietzmann, R.; de Souza, D.R.; Lerina, J.M.F.; Manssour, I.H.; Silveira, M.S. Humor, support and criticism: A taxonomy for discourse analysis about political crisis on Twitter. In Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age, Delft, The Netherlands, 30 May–1 June 2018.

46. Winkels, R.; Douw, J.; Veldhoen, S. Experiments in automated support for argument reconstruction. In Proceedings of the 14th International Conference on Artificial Intelligence and Law, Rome, Italy, 10–14 June 2013.

47. Therón, R.; Seguín, C.; de la Cruz, L.; Vaquero, M. Highly interactive and natural user interfaces: Enabling visual analysis in historical lexicography. In Proceedings of the First International Conference on Digital Access to Textual Cultural Heritage, Madrid, Spain, 19—20 May 2014.

48. De Rooij, O.; Odijk, D.; De Rijke, M. Themestreams: Visualizing the stream of themes discussed in politics. In Proceedings of the 36th international ACM SIGIR Conference on Research and Development in Information Retrieval, Dublin, Ireland, 28 Jul–1 August 2013.

49. Thiel, M.; Ludwig, P.; Mossakowski, T.; Neuhaus, F.; Nürnberger, A. Web-retrieval supported argument space exploration. In Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval, Oslo, Norway, 7–11 March 2017.

50. Hassani, K.; Lee, W.-S. Visualizing natural language descriptions: A survey. *ACM Comput. Surv.* **2016**, *49*, 1–34. [CrossRef]

51. Zimmerman, C.; Chen, Y.; Hardt, D.; Vatrapu, R. Marius, the giraffe: A comparative informatics case study of linguistic features of the social media discourse. In Proceedings of the 5th ACM International Conference on Collaboration across Boundaries: Culture, Distance & Technology, Kyoto, Japan, 20–22 August 2014.

52. Chaves, A.P.; Gerosa, M.A. Single or Multiple Conversational Agents? An Interactional Coherence Comparison. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018.

53. Mayfield, E.; Black, A.W. Analyzing Wikipedia Deletion Debates with a Group Decision-Making Forecast Model. *Proc. ACM Hum. Comput. Interact.* **2019**, *3*, 1–26. [CrossRef]

54. Aigner, W. Current Work Practice and Users' Perspectives on Visualization and Interactivity in Business Intelligence. In Proceedings of the 17th International Conference on Information Visualisation, London, UK, 16–18 July 2013.

55. Kerren, A.; Prangova, M.; Paradis, C. Visualization of sensory perception descriptions. In Proceedings of the 15th International Conference on Information Visualisation, London, UK, 13–15 July 2011.

56. Brandes, U.; Corman, S.R. Visual unrolling of network evolution and the analysis of dynamic discourse. *Inf. Vis.* **2003**, *2*, 40–50. [CrossRef]

57. Reis, L.P.; Costa, A.P.; de Souza, F.N. A survey on computer assisted qualitative data analysis software. In Proceedings of the 11th Iberian Conference on Information Systems and Technologies (CISTI), Las Palmas, Spain, 15–18 June 2016.

58. Chiru, C.-G.; Trausan-Matu, S. A tool for discourse analysis and visualization. In Proceedings of the 3rd International Conference on Emerging Intelligent Data and Web Technologies, Bucharest, Romania, 19–21 September 2012.

59. Lucas, N.; Giguet, E. Robust adaptive discourse parsing for e-learning fora. In Proceedings of the 8th IEEE International Conference on Advanced Learning Technologies, Santander, Spain, 1–5 July 2008.

60. Li, Y.; Wang, J.; Liao, J.; Zhao, D.; Huang, R. Assessing collaborative process in CSCL with an intelligent content analysis toolkit. In Proceedings of the 7th IEEE International Conference on Advanced Learning Technologies (ICALT 2007), Niigata, Japan, 18–20 July 2007.

61. Somaraki, V.; Xu, Z. Epicurus: A platform for the visualisation of forensic documents based on a linguistic approach. In Proceedings of the 22nd International Conference on Automation and Computing (ICAC), Colchester, UK, 7–8 September 2016.

62. Nukoolkit, C.; Chansripiboon, P.; Mongkolnam, P.; Todd, R.W. Text cohesion visualizer. In Proceedings of the 6th International Conference on Computer Science & Education (ICCSE), Singapore, 3–5 August 2011.

63. Lam, H.; Fisher, B.; Dill, J. A Pilot Study of CZTalk: A Graphical Tool for Collaborative Knowledge Work. In Proceedings of the 38th Annual Hawaii International Conference on System Sciences, Big Island, HI, USA, 6 January 2005.

64. Emad, S.; Halvorson, W.; Broillet, A.; Dunwell, N. The competency building process of human computer interaction in game-based teaching: Adding the flexibility of an asynchronous format. In Proceedings of the IEEE International Professonal Communication 2013 Conference, Vancouver, BC, Canada, 15–17 July 2013.

65. Lauscher, A.; Glavaš, G.; Eckert, K. Arguminsci: A tool for analyzing argumentation and rhetorical aspects in scientific writing. In Proceedings of the 5th Workshop on Argument Mining, Brussels, Belgium, 31 October–1 November 2018.

66. Surdeanu, M.; Hicks, T.; Valenzuela-Escárcega, M.A. Two practical rhetorical structure theory parsers. In Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations, Denver, CO, USA, 31 May–5 June 2015.

67. Gilmartin, E.; Campbell, N. Capturing Chat: Annotation and Tools for Multiparty Casual Conversation. In Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC'16), Portorož, Slovenia, 23–28 May 2016.

68. Hoque, E.; Carenini, G.; Joty, S. Interactive exploration of asynchronous conversations: Applying a user-centered approach to design a visual text analytic system. In Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces, Baltimore, MA, USA, 27 June 2014.

69. Helfrich, P.; Rieb, E.; Abrami, G.; Lücking, A.; Mehler, A. TreeAnnotator: Versatile visual annotation of hierarchical text relations. In Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018), Miyazaki, Japan, 7–12 May 2018.

70. Wachsmuth, H.; Da San Martino, G.; Kiesel, D.; Stein, B. The impact of modeling overall argumentation with tree kernels. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 7–11 September 2017.

71. Zeldes, A. The GUM corpus: Creating multilayer resources in the classroom. *Lang. Resour. Eval.* **2017**, *51*, 581–612. [CrossRef]

72. Kutty, S.; Nayak, R.; Turnbull, P.; Chernich, R.; Kennedy, G.; Raymond, K. PaperMiner—A real-time spatiotemporal visualization for newspaper articles. *Digit. Scholarsh. Human.* **2019**, *35*, 83–100. [CrossRef]

73. Viola, L.; Verheul, J. Mining ethnicity: Discourse-driven topic modelling of immigrant discourses in the USA, 1898–1920. *Digit. Scholarsh. Human.* **2019**. [CrossRef]

74. Klein, L.F.; Eisenstein, J.; Sun, I. Exploratory thematic analysis for digitized archival collections. *Digit. Scholarsh. Human.* **2015**, *30*, i130–i141. [CrossRef]

75. Drucker, J. Non-representational approaches to modeling interpretation in a graphical environment. *Digit. Scholarsh. Human.* **2018**, *33*, 248–263. [CrossRef]

76. Muralidharan, A.; Hearst, M.A. Supporting exploratory text analysis in literature study. *Lit. Linguist. Comput.* **2012**, *28*, 283–295. [CrossRef]

77. Brierley, C.; Atwell, E. Non-traditional prosodic features for automated phrase break prediction. *Lit. Linguist. Comput.* **2011**, *26*, 279–284. [CrossRef]

78. Andrews, T.L. Analysis of variation significance in artificial traditions using Stemmaweb. *Digit. Scholarsh. Human.* **2016**, *31*, 523–539. [CrossRef]

79. Brughmans, T. Networks of networks: A citation network analysis of the adoption, use, and adaptation of formal network techniques in archaeology. *Lit. Linguist. Comput.* **2013**, *28*, 538–562. [CrossRef]

80. Zhitomirsky-Geffe, M.; Prebor, G.; Miller, Y. Ontology-based analysis of the large collection of historical Hebrew manuscripts. *Proc. Assoc. Inf. Sci. Technol.* **2018**, *55*, 958–959. [CrossRef]

81. García, N.A.; Caplan, A.; Mering, B. A Pedagogy for Computer-Assisted Literary Analysis: Introducing GALGO (Golden Age Literature Glossary Online). *DHQ Digit. Hum. Q.* **2017**, *11*, 21–31.