ChinaRxiv 中X

AI translation · View original & related papers at
chinarxiv.org/items/chinaxiv-202304.00078

A Survey of Argument Mining Research: Postprint

Authors: Li Yongze, Ou Shiyan

Date: 2023-04-01T16:16:02+00:00

Abstract

[Purpose/Significance] Argument mining aims to identify argument structures within argumentative texts, thereby enabling an understanding of the reasoning and processes underlying conclusion derivation. It holds significant academic and practical value, and has attracted widespread attention in recent years across domains including social media content mining, legal decision support, and decision support, emerging as a novel research direction in the field of text mining. This paper seeks to systematically review and summarize the current state of research and applications in argument mining, identify research frontiers, and provide references for future studies. [Method/Process] In the Association for Computational Linguistics (ACL) database and the Web of Science database, searches were conducted using "argument mining", "argument structure", and "argument component" as keywords. Combined with manual screening, a total of 220 relevant literature items on argument mining were collected. Through close reading, the current research was analyzed and summarized from three perspectives: argument models, argument mining tasks, and argument mining applications. [Results/Conclusion] Research on argument mining remains in its infancy. While numerous studies have focused on simple argumentative texts such as social media content, research on complex argumentative texts like scientific papers is still limited. Future work could investigate complex texts through three avenues: argument annotation schemes, identification of argument components and relations, and argument structure optimization.

Full Text

Preamble

Journal of Library and Information Science, Volume 64, Issue 19, October 2020

A Review of Argument Mining Research

Li Yongze, Ou Shiyan

School of Information Management, Nanjing University, Nanjing 210023

Abstract: [Purpose/Significance] Argument mining aims to identify argument structures in argumentative texts, thereby enabling understanding of the reasons and processes behind conclusions. It holds significant academic and application value and has garnered widespread attention in recent years in areas such as social media content mining, legal assistance in adjudication, and decision support, emerging as a new research direction in text mining. This paper seeks to systematically review and summarize the current state of argument mining research and applications, identify research hotspots, and provide references for future studies. [Method/Process] We searched the ACL and Web of Science databases using the keywords "argument mining," "argument structure," and "argument component," supplemented with manual screening, to collect 220 relevant articles on argument mining. Through intensive reading, we analyzed and summarized current research from three perspectives: argument models, argument mining tasks, and argument mining applications. [Result/Conclusion] Argument mining research is still in its infancy, with numerous studies focusing on simple argumentative texts such as social media content, but relatively few addressing complex texts like scientific papers. Future research could explore complex texts from three aspects: argument annotation schemes, identification of argument components and relations, and optimization of argument structures.

Keywords: argument mining; argument model; argument structure; argument component

Classification Number: G254

DOI: 10.13266/j.issn.0252-3116.2020.19.014

Introduction

Argumentation is an ancient topic that can be traced back to Aristotle, involving philosophy, linguistics, psychology, and other disciplines. The argumentation process refers to a situation where two parties holding different views on an issue aim to persuade or refute each other, analyzing the problem from different perspectives, presenting reasons for or against a certain viewpoint, and ultimately drawing corresponding conclusions [1]. The basic structure of argumentation includes a series of premises and a conclusion, as well as supporting or opposing relationships between premises and between premises and conclusions [2]. Premises and conclusions constitute argument components, while the relationships between components are called argument relations; together they form argument structures. Analyzing argument structures in argumentative texts enables better understanding of the arguer's viewpoints and the reasons behind them. In today's digital and networked environment, there exists a vast amount of argumentative texts, such as product reviews, legal judgments, news comments, persuasive essays, and scientific papers. These texts contain important viewpoints and their reasoning processes but are often implicit in unstructured or semi-structured formats. How to automatically parse argument structures from texts is an urgent problem to be solved, thus giving rise to the emerging research field of argument mining.

Argument mining is an application of natural language processing and machine learning techniques to argumentation, representing a branch of text mining and an extension of opinion mining. As early as 2007, M.F. Moens and N. Kwon conducted argument mining research on legal texts and review texts respectively [3-4], initiating this new research domain. In 2009, R.M. Palau and M.F. Moens formally defined the fundamental tasks of argument mining: using natural language processing techniques to automatically extract arguments from unstructured argumentative texts, analyze the internal structure of arguments and relationships between different arguments, and ultimately provide structured argument knowledge [5]. In recent years, with the proliferation of argumentative texts and advancements in natural language processing technologies, argument mining research has gained widespread attention. Since 2014, numerous international conferences on argument mining have been held, such as the Argument Mining Workshop at the ACL 2014 conference [6] and the SICSA Workshop on Argument Mining [7].

As argument mining research has deepened and developed, researchers have actively explored its applications in different domains, such as extracting argument components and structures from legal texts to assist in adjudication [3, 8], and automatically scoring student essays by analyzing the rationality of their argument structures [9]. Text information mining and organization is one of the core research areas in library and information science, and argument mining represents a technology and method for mining and organizing information from unstructured texts from an argumentative perspective. It is an application of text mining to specific types of corpora—argumentative texts—and thus holds important research and application value in library and information science.

This paper systematically reviews existing literature on argument mining, focusing on current research progress and hotspots in the field, tracking new research findings, and analyzing development trends, with the aim of attracting further attention from domestic library and information science scholars to this research area and providing new perspectives and applications for information mining and organization in the discipline.

2 Argument Models and Argument Corpora

2.1 Argument Models

Argument models provide abstract or conceptual descriptions of the argumentation process, primarily consisting of argument components and relationships between them. J. Bentahar et al. classified argument models into three types based on their content and granularity: monological models, dialogical models, and rhetorical models [11]. Monological models focus on the micro-level internal structure of arguments, describing argument components within a claim and their relationships. Representative models include the Toulmin model [12], the Freeman model, and argumentation schemes [2]. Dialogical models focus on macro-level argument structures, primarily characterizing relationships between different claims rather than internal relationships within a claim, with P.M. Dung's abstract model being a representative example [13]. Rhetorical models focus on the rhetorical structure of argumentation, aiming to arrange argument content and form strategically to persuade the opponent, rather than on macro or micro structures, with Perelman and Olbrechts-Tyteca's New Rhetoric model being representative [14]. Argument mining primarily focuses on micro-level argument structures, so this section mainly introduces monological models that focus on internal claim structures. The key representative models are as follows.

2.1.1 Premise-Claim Model The premise-claim model is the most basic and widely applied argument model. It abstracts an argument as consisting of one or a group of premises and a claim. The claim is a viewpoint proposed by the arguer, while the premises provide reasons supporting or opposing that viewpoint [1] (see Figure 1 [FIGURE:1]). This model has a simple structure and cannot distinguish between different roles of premises, making it relatively weak in describing complex argumentative content. In current argument mining research, social media content often adopts the "premise-claim" model due to its relatively simple argument structure.

2.1.2 Toulmin Model In 1958, British philosopher Stephen E. Toulmin proposed the Toulmin model based on his analysis of argumentation in legal texts [12]. The model includes six components: Claim, Datum (premise), Warrant, Backing, Modal Qualifier, and Rebuttal. Among these, claim, datum, and warrant are the core elements of argument structure (see Figure 2

). The claim is an assertion representing the arguer's main viewpoint; the datum provides reasons for the claim; the warrant is the inference chain linking the datum to the claim, indicating how the claim is derived from the datum; backing explains the reliability of the warrant; modal qualifiers (such as "probably," "certainly") indicate the degree of certainty of the claim; and rebuttal refers to counterarguments against the argument. The Toulmin model describes rich types of argument components and complex structures, making it suitable for annotating argument structures in complex monological long texts such as legal judgments. However, in simple texts like social media content, components such as warrant and rebuttal are often not explicitly expressed, making the Toulmin model difficult to apply.

2.1.3 Standard Approach The Standard Approach refers to the organization of premises and can describe relationships between premises through convergence, serial, linked, and divergent patterns [15-16] (see Figure 3

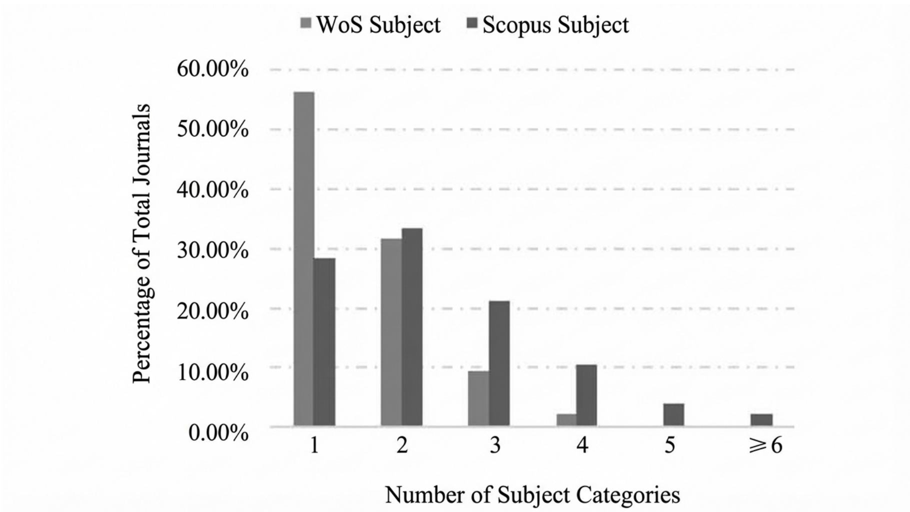

Figure 1: Figure 2

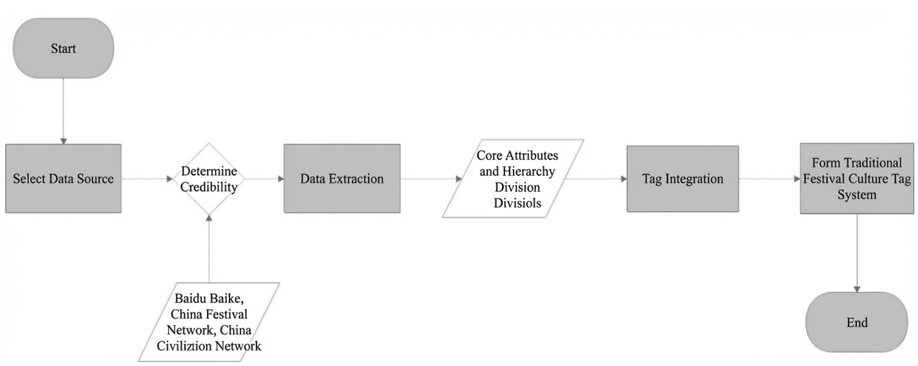

Figure 2: Figure 3

). Convergence means different premises support the same conclusion from different aspects or perspectives (Figure 3(a)). Serial refers to deriving an intermediate conclusion from initial premises, which then serves as a premise to further derive the final conclusion (Figure 3(b)). Linked means two or more premises jointly infer a conclusion (Figure 3(c)). Divergent means one premise can infer two different conclusions (Figure 3(d)). The Standard Approach is more flexible than the Toulmin model and can associate and reorganize multiple sub-arguments through these four patterns to form more complex argumentation structures.

2.1.4 Freeman Model The Freeman model [15] was developed by American scholar J.B. Freeman in 1991 through integrating and modifying the Toulmin model and the Standard Approach [16]. The Freeman model consists of five elements: Premise, Claim, Modifier, Rebuttal, and Counter-rebuttal. In addition to the four components shared with the Toulmin model—premise, claim, modifier, and rebuttal—it adds a counter-rebuttal component for further argumentation on rebuttal content (see Figure 4

). Compared with the Standard Approach, the Freeman model inherits and improves upon the rebuttal and modal qualifier components from the Toulmin model, making it more flexible and providing deeper characterization of argument structures. Compared with the Toulmin model, the Freeman model eliminates the fine distinctions among datum, warrant, and backing, allowing premises to support claims through different patterns such as linked and convergence, thereby overcoming the Toulmin model's limitation in representing complex arguments composed of multiple sub-arguments. Additionally, the inclusion of the counter-rebuttal component enables further argumentation on rebuttals based on counterexamples [17].

2.1.5 Argumentation Schemes Argumentation schemes refer to the inference structures used in natural language conversation or argumentation—common presumptive and plausible reasoning patterns that represent a third type of reasoning distinct from deduction and induction [18]. They describe the inference rules from premises to conclusions, reflecting human reasoning processes such as causal relationships and analogies. In argument mining tasks, the classification of argumentation schemes by Canadian logician D. Walton is typically adopted [19], which includes 96 types of schemes. However, in practice, only a few commonly used schemes are selected to describe argument structures in texts, such as expert opinion schemes and analogy-based schemes [20] (see Table 1).

The inference rules in the Toulmin and Freeman models are implicit in argument components such as the warrant. Different argumentation schemes express different inference rules, and critical questions provide evaluation of argument strength. Therefore, argumentation schemes have received widespread attention in argument mining tasks. However, due to the large number of scheme

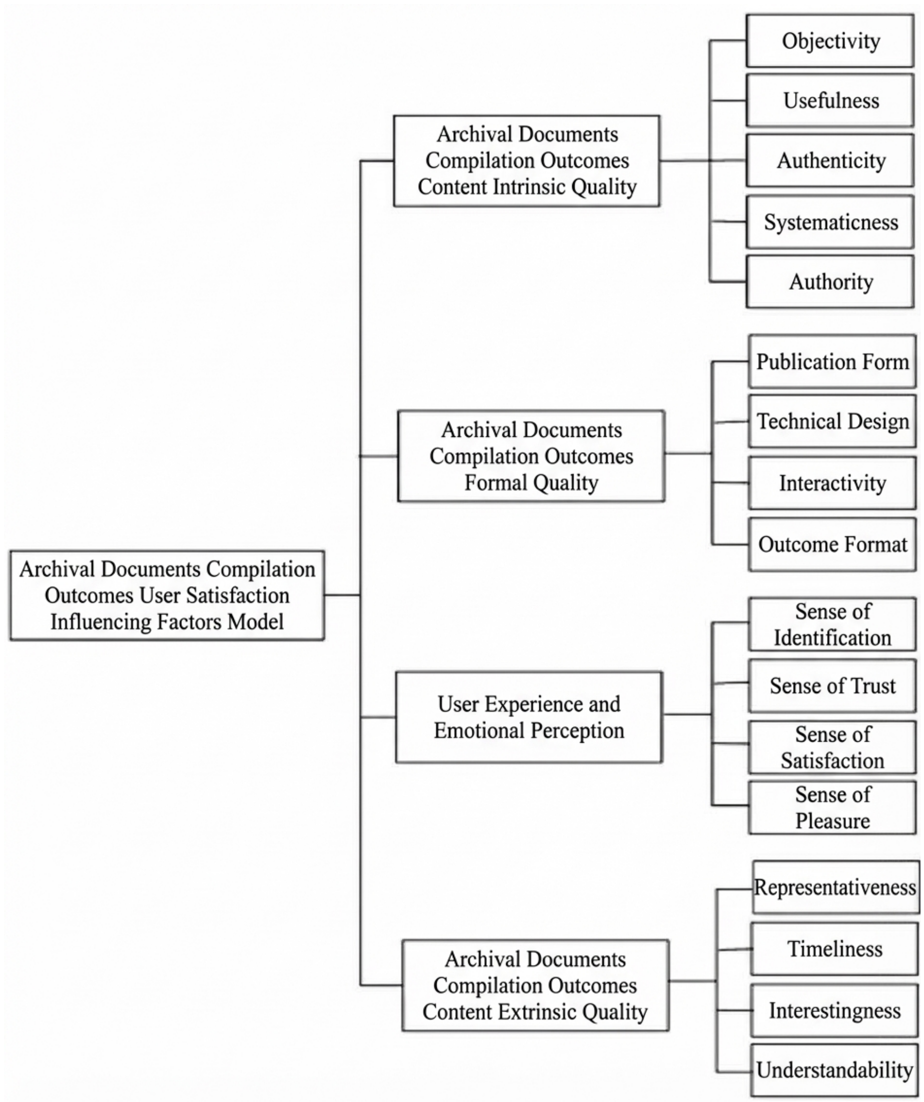

Figure 3: Figure 4

types, they sometimes adversely affect argument mining tasks [21].

2.2 Argumentative Text Corpora and Annotation Schemes

Argument mining requires large amounts of manually annotated training data. Different argumentative texts have different argument structures, necessitating the construction of appropriate annotation schemes. Argument annotation schemes differ slightly from argument models—they represent specific adjustments and comprehensive applications of argument models.

We analyze the construction of argument annotation schemes from two perspectives: text type and argument structure. Different types of texts vary significantly in normativity, which greatly affects the difficulty of argument annotation. Based on normativity and structural characteristics, argumentative texts can be divided into monological long texts and interactive short texts. The former includes student essays [25], scientific papers [22], etc.; the latter includes policy debates, social media discussions on hot issues, and product reviews [26]. In terms of the number of claims, monological texts typically contain multiple claims [23], while interactive texts often focus on only one issue per round of dialogue. Regarding argument structure, monological texts have relatively complex argument content and structures, often containing multiple claims and complex argument components. Their argument structures are closely related to discourse structure, and the argumentation process is reflected across different discourse segments with formal argumentation procedures. Interactive texts have relatively simple argument structures, with argumentation processes mainly manifested in user interaction behaviors, such as direct and indirect replies on social media.

Current argument mining research primarily focuses on interactive short texts (e.g., social media content), and corpora are relatively abundant. However, short text corpora often suffer from insufficient normativity, such as non-standard expressions and incomplete argument structures, which adversely affect argument mining [27].

Argument structures can be divided into macro and micro levels. Different researchers often choose to annotate either macro or micro structures based on corpus characteristics and practical needs. Macro structure refers to relationships between different claims [3], often annotated based on dialogical models [28-29]. Micro structure refers to relationships between different argument components within a single claim, often annotated based on monological models [25]. Some researchers also comprehensively annotate both macro and micro structures of argumentative texts [30]. Current argument mining research mainly focuses on micro-level argument structure annotation.

When applying monological models to annotate actual corpora, the argument components and relationships defined in the models often do not perfectly match the actual situation of the corpora. Therefore, researchers often adjust argument models to construct specific annotation schemes [22, 31-32]. Table 2 lists some corpora and their annotation schemes used in representative studies.

Through summarizing research on argument corpus annotation, we can observe that: (1) Different argument mining tasks require different annotation schemes; argument models are not the sole determining factor and are often adjusted according to actual conditions, with argument components or relationships being refined accordingly; (2) Annotation schemes can be constructed based on multiple argument models simultaneously, such as annotating both macro and micro structures; (3) Current argument corpora mainly focus on two argument components—premise and claim—with less attention to other components; (4) In an argument chain, the roles of argument components are mutually transformable, where the conclusion of one argument can serve as the premise of another [22], but most existing annotation schemes fail to reflect this dynamic transformation; (5) Currently, most annotated argument corpora are in English, with no annotated Chinese corpora identified.

3 Argument Mining Task Flow

Argument mining tasks are typically divided into two major stages: argument structure identification and argument structure optimization. The former can be further subdivided into three subtasks: argument unit identification, argument component classification, and argument relation extraction [25, 33]; the latter is subdivided into implicit argument component identification and global argument structure optimization (see Figure 5

).

3.1 Argument Unit Identification

Argument units refer to individual argument components in an argument structure. Argument unit identification is the foundation of argument mining, mainly including two steps: argumentative sentence identification and argument unit boundary identification—that is, first identifying sentences with argumentative properties from the text, then determining the exact starting position of argument units within sentences [33] (see Figure 6

).

The definition of argument units is related to the granularity of argument components, typically divided into sentence-level and clause-level. There are four possibilities: (1) a complete sentence constitutes one argument component (sentence-level); (2) a sentence contains both argumentative and non-argumentative components (clause-level); (3) a sentence contains multiple different argument components (clause-level); and (4) multiple sentences jointly constitute one argument component (cross-sentence level) [23]. For the first case, the argument unit boundary is the entire sentence; for the second and third cases, the exact boundary of each argument unit needs to be identified; for the fourth case, relationships between different sentences need to be determined.

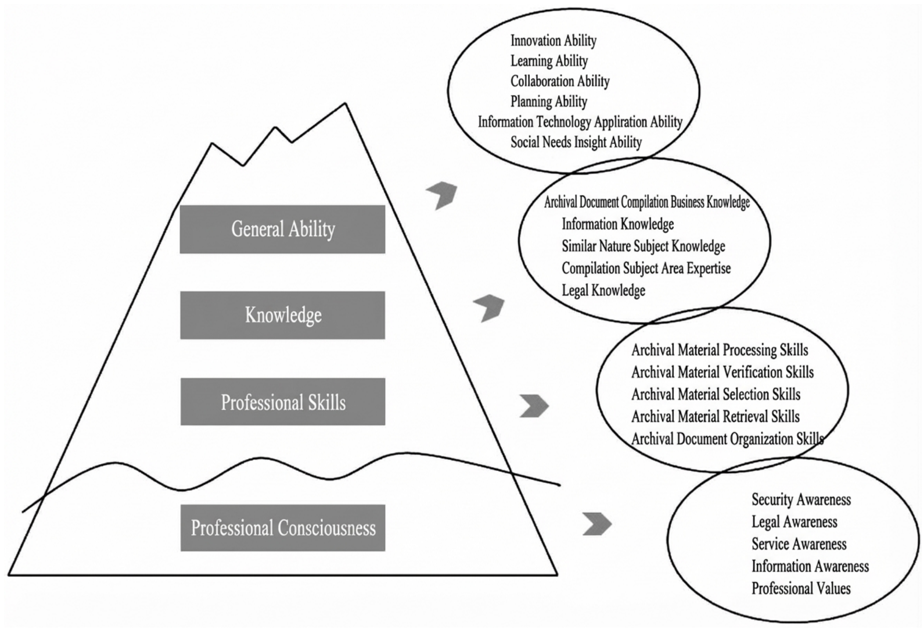

Figure 4: Figure 5

For argumentative sentence identification, the task is typically converted into a sentence classification problem, using text classification algorithms in machine learning to distinguish argumentative from non-argumentative sentences. Common classification algorithms include Bayesian [3, 34], decision trees [35], logistic regression [34, 36], random forests [34, 36], and support vector machines [34], all achieving good results. Early studies used relatively shallow text features such as verbs, sentence length, punctuation count [3], tense and voice [35], entity count, and adjective count [34]. As research progressed, semantic and sentiment features were gradually introduced into classification models, improving classification effectiveness [36]. Since 2018, with the emergence of contextualized word embeddings, scholars have also used dynamic models such as BERT and ELMo to identify argument components. Benefiting from these models' retention of contextual information, particularly when argument topic information is added to the models, good identification results can be achieved [37]. Additionally, some scholars have attempted unsupervised methods to identify argumentative sentences. For example, A. Ferrara et al. determined sentences with high similarity to the argument topic as argumentative sentences, but experimental results showed this method had high recall but low precision, particularly performing poorly in corpora with many non-argumentative sentences [38].

We compared and summarized these three types of argumentative sentence identification methods, concluding that: text classification algorithms are easy to implement and can achieve good results even with small training datasets, but

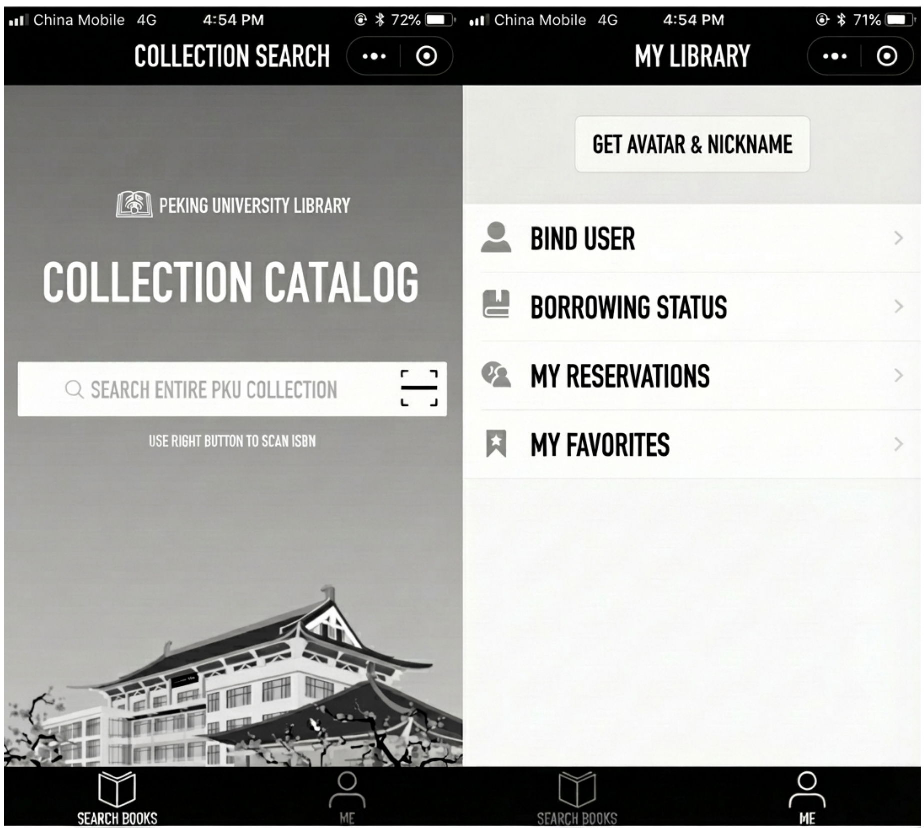

Figure 5: Figure 6

require manual construction of classification features with weak model transferability; contextualized word vector models can retain contextual information and achieve better identification results, but have certain requirements for training data scale and quality; unsupervised algorithms use text similarity to identify argumentative sentences without building text features and requiring only small amounts of annotated data, making them simple to implement but with significant limitations, particularly poor performance when non-argumentative sentences are numerous.

For argument unit boundary identification, the simplest method is to use text classification techniques to determine whether a word is the start or end of an argument unit [39]. Although this method can achieve good results with small training samples, it requires training multiple classifiers, resulting in high overall training costs and is not commonly used. The distribution of argument components in texts has obvious sequential characteristics—for example, premises are often given after claims. Therefore, argument unit boundary identification is more often converted into a sequence labeling problem, with Conditional Random Fields (CRF) being the most commonly used method [21, 34, 40]. In recent years, with significant advances in deep learning, scholars have begun using neural networks (such as RNN or LSTM) combined with CRF to identify argument unit boundaries [41-42], achieving better results than using CRF alone. This is because argumentative texts have strong semantic correlations in context, and recurrent neural networks can retain longer-distance contextual information, while CRF primarily retains intra-sentence context. Some scholars have also combined BERT word vector models with CRF models, achieving excellent identification results that even surpass human performance [43]. This is because the BERT model utilizes position encoding and attention mechanisms to achieve deeper mining and utilization of contextual information, but simultaneously has high requirements for training data and hardware/software.

3.2 Argument Component Classification

Argument component classification refers to categorizing argument units into different component types based on their roles in argumentation, such as premise, claim, warrant, etc. When argument components include only premises and claims, this can be viewed as a simple binary classification problem. When there are three or more components, the problem transforms into multiple binary classification problems or a multi-label classification problem; existing research mainly adopts multiple binary classification approaches. Some studies combine argument unit identification with argument component classification, identifying argument units while simultaneously classifying them as premise, claim, or non-argument component.

From a technical perspective, argument component classification mostly uses supervised text classification algorithms such as Support Vector Machines [5, 44], Bayesian [45], random forests [45, 46], and decision trees [46], among which Support Vector Machines often achieve the best results [46]. Features used mainly include lexical features (such as cue words [49] and topic words [50]) and syntactic features (such as verb count and named entity count in sentences [34]). With the development of deep learning, scholars have applied neural network models (such as CNN [47] and LSTM [51]) to argument component classification, or combined them with manually constructed features, achieving better results [47]. Additionally, some scholars have used semi-supervised methods based on rules and sentence similarity to identify argument components [48], but this method is premised on the assumption that the same argument components have similar syntactic structures, thus having certain limitations.

We summarized and compared these three argument component identification methods: traditional classification algorithms require manual feature construction; deep learning algorithms can automatically learn features, but incorporating manually constructed features can improve classification effectiveness; unsupervised methods require no feature construction, only manual rule construction, but have limited applicability. In future research, applying BERT word vector models and attention mechanisms to argument component classification will be a key focus.

3.3 Argument Relation Extraction

The focus of argument structure identification is argument relation identification. Argument relations can be divided into two categories: basic support, oppose, and neutral relations; and rhetorical relations as supplementary argument relations based on support and opposition, such as explanation and elaboration [23, 32]. Although the second category provides richer semantic information between argument components, it also increases the difficulty of argument structure identification. Most argument relation identification focuses on the first category. Argument relation identification is divided into two stages: relation detection and relation classification—that is, first determining whether an argumentative relation exists between a pair of argument components (whether they can be associated), then further determining the type of association (the specific argument relation).

Determining whether an argumentative relation exists between a pair of argument components can be approached from two angles. On the one hand, the problem can be converted into a binary classification problem, concatenating two argument components within a certain range into a text segment, then using text classification to determine whether an association relation exists in that text [21]. On the other hand, it can be judged from the perspective of topic similarity, calculating topic similarity between different argument components based on topic models and determining that texts with higher similarity have argumentative relations [39, 52]. This method requires no manual feature construction but has relatively low recall.

Argument relation classification methods can be roughly divided into two categories: textual entailment recognition-based and text classification-based. Textual entailment recognition technology is mainly used to infer semantic relations between two texts. For argument components, if premise statement T can infer conclusion statement H, it means premise T entails (supports) conclusion H [53]. Therefore, some scholars have used it to determine simple support or opposition relations between premises and assumptions [54], but it cannot identify more complex argument relations. Argument relation identification can also be viewed as a text classification task, using classification algorithms to classify text segments formed by concatenating two argument components into support or opposition categories [21]. Common classification features include linguistic features [50] and discourse structure features [57-58]. Additionally, contextual semantic information is also an important basis for argument relation classification. Since Recurrent Neural Network (RNN) models and their variants (such as LSTM) can effectively capture contextual information, they have been widely applied in argument relation classification [55-56]. However, when data volume is small, deep learning algorithms do not perform as well as traditional machine learning algorithms [55].

Through analysis and summary of these three methods, we find that: textual entailment recognition methods can determine inference relations between different components to a certain extent, but they are not based on strict logical reasoning, making their reliability difficult to guarantee and complex relations hard to identify; using traditional text classification algorithms for argument relation classification requires more complex feature construction than argument unit identification and argument component classification, resulting in poorer classification effectiveness; since recurrent neural network models more easily capture sequential information between argument components, they can achieve better results in argument relation classification.

3.4 Argument Structure Optimization

Argument structure optimization includes two aspects: first, identifying implicit argument components in argument structures and supplementing them into existing structures; second, analyzing whether there are contradictory or conflicting relations in argument structures to achieve global optimization.

3.4.1 Identification of Implicit Argument Components Due to writing techniques and purposes, some content may be omitted in argumentative texts, resulting in missing argument components [32]. While this does not affect the overall semantic expression of the text, it may impact the clarity and persuasiveness of the argumentation [59]. One of the tasks of argument mining is to reconstruct and supplement these implicit components, making argument structures more complete and intuitive, which is of great significance for users to understand argument logic and represents an important future research direction in argument mining.

The reconstruction and supplementation of implicit argument components largely depend on argumentation logic. Only by fully understanding the argument structure of a text can missing components be identified. N. Green believes that constructing complete argument structures requires consideration of argumentation schemes, domain knowledge, assumed common understanding between authors and potential readers, and the context of argumentation [59]. Currently, given the complexity of argument structures themselves, the discovery and supplementation of implicit argument components still mainly rely on manual work. Some researchers manually identify missing argument components based on argumentation schemes or argument structures after automatically identifying "premise-claim" pairs [60-62]. Others partially use machine learning methods for semi-automatic identification of implicit components, such as using text classification techniques to automatically determine the correct "warrant" between a "premise-claim" pair [63].

Overall, existing machine learning methods cannot fully automatically identify implicit argument components. Additionally, manual identification of implicit components requires substantial human and material resources, resulting in a lack of corpora with accurate annotations and rich content, which poses significant obstacles to automatic identification of implicit argument components.

3.4.2 Global Optimization of Argument Structures C. Stab et al. found that argument structures in texts are often ambiguous—i.e., the same text can be described using different argument structures [25], a situation that becomes more pronounced when structural indicators are lacking [23]. The ambiguity of argument structures poses significant challenges to manual annotation and automatic mining of argumentative texts, requiring corresponding standards and metrics to evaluate argument structures and identify optimal structures. There is relatively little research directly comparing and analyzing argument structures; most studies indirectly evaluate argument structures using methods for assessing argument quality from logic.

Since argument components can form graph structures through argument relations, A. Lauscher et al. proposed graph structure-based evaluation metrics for argument structures, including graph diameter, number of subgraphs, and structure depth [22]. Based on generated argument graphs, some researchers use minimum spanning trees or linear programming to globally optimize graph structures to ensure no conflicts exist between argument relations [50, 64]. Experimental results show that the optimization effects of these two methods are not significantly different [65]. Currently, most research focuses on argument structure identification, with relatively few studies on argument structure optimization. The characteristics, attributes, and evaluation metrics of argument structures also lack corresponding theoretical foundations.

4 Applications of Argument Mining

Currently, argument mining applications primarily focus on argumentative texts such as persuasive essays, scientific papers, social media content, and legal judgments, covering fields including education, scientific research, social media, and law.

4.1 Argument Mining in Legal Texts

The legal domain is one of the earliest application areas of argument mining technology. Legal texts such as court judgments are rich in reasoning and argumentation processes for case decisions. By parsing argument structures, analyzing evidence types, and constructing rules for handling conflicting evidence, argument mining can assist in adjudication [66]. In 2007, M.F. Moens first identified argumentative content in legal texts [3] but did not distinguish specific argument components. In 2009, R. Mochales and M.F. Moens extracted simple syllogisms from legal judgments and formed tree structures to clearly display final conclusions, premises for conclusions, and premises for premises [5], greatly facilitating reading and understanding for relevant personnel. Subsequently, some researchers analyzed fine-grained argument structures in legal texts, refining argument component types, but only provided annotated corpora without achieving automatic extraction of argument components and relations [66-67].

Overall, legal texts contain many types of argument components, and understanding their content requires certain legal background knowledge, which increases the difficulty of argument extraction. Moreover, current research only focuses on basic support and opposition relations; other types of argument relations in legal texts require further study.

4.2 Argument Mining in Persuasive Essays

Persuasive essays often analyze controversial topics to elaborate authors' viewpoints and have obvious argumentation processes, making them important corpora for argument mining. Currently, argument mining for persuasive essays is mainly used for automatic essay scoring, typically evaluating argument strength and credibility. In 2014, Y. Song was the first to use argumentation schemes to describe essay structures, automatically determining the degree of match between essays and argumentation schemes and finding that matching degree positively correlated with essay quality, but did not achieve automatic scoring [9]. In 2015, I. Persing et al. evaluated argument strength by analyzing whether contradictions existed between argument components and automatically scored essays based on this [68]. In 2017, H. Wachsmuth et al. evaluated argument quality based on the organization patterns of premises and claims (e.g., premise + premise + claim) and the number of arguments after automatically extracting essay argument structures [69-70], but this research only considered argument components while ignoring the role of argument relations.

Additionally, argument mining can be used for assisted writing of persuasive essays by parsing argument structures to help authors discover missing or conflicting components. For example, in 2014, G. Stab supported essay writing and evaluation by automatically assessing the reliability of extracted premises and the degree of support for claims [21]. However, current research has not examined assisted writing from perspectives such as claim arrangement and missing component supplementation, which represent future research directions.

4.3 Argument Mining in Social Media

User discussions and analyses of issues on social media exhibit clear argumentative characteristics. Argument mining for social media includes decision support, truthfulness judgment, and key issue identification. Social media is an important platform for users to express viewpoints and exchange opinions. By analyzing users' attitudes toward a particular issue and the reasons supporting or opposing them, we can gain deeper understanding of user viewpoints [29] and provide support for government decision-making [73]. Additionally, some researchers have automatically classified the persuasiveness of user comments in forums to identify highly persuasive comments [48, 71], which has reference value for guiding public opinion.

User-generated content on social media is sometimes difficult to distinguish as true or false. Some researchers assist in judging the truthfulness of conclusions by analyzing the sources of argument premises [36]. User arguments on social media often focus on certain important issues, and identifying core argumentative issues is key to resolving disputes. For example, B. Konat et al. analyzed discussions of social events in forums, parsed their argument structures, constructed argument graphs, and thereby identified the most controversial argument components [72].

Argument mining has many application directions in social media, but existing studies only focus on individual applications without systematic approaches, such as using social platforms for decision support while also needing to judge the truthfulness of source information. Future research could conduct more in-depth studies based on argument structures, such as public opinion monitoring and rumor identification.

4.4 Argument Mining in Scientific Papers

The purpose of scientific papers is to convey authors' scientific viewpoints and conclusions to readers and persuade them to accept these ideas—essentially also an argumentation process. Compared with persuasive essays, scientific papers have more complex argumentation processes, and argumentation methods differ across disciplines. Currently, there is relatively little argument mining research on scientific papers. Some studies mainly analyze argument components in scientific papers from a discourse structure perspective, such as the work of S. Teufel [74], T. Mayer [77], and A.J. Yepes [75], but they do not identify argument relations and thus do not constitute strict argument mining. Only a few researchers have conducted argument mining on parts of scientific papers. For example, in 2015, N. Green automatically identified argument structures in the introduction and discussion sections of biomedical scientific papers and then manually identified implicit argument components based on argumentation schemes to help readers understand arguments about gene-disease relationships in the papers [62, 76].

Argument mining for scientific papers currently faces several issues: (1) Lack of argument corpora—existing corpora are small in scale and cannot meet data requirements for argument mining; (2) No argument mining for full scientific papers—most research focuses on sections with more obvious argumentation processes such as abstracts, introductions, and discussions; (3) Few application studies on argument mining for scientific texts.

Conclusion

This paper systematically reviews the concept, task flow, and main implementation methods and technologies of argument mining. Additionally, it introduces and evaluates the current research and application status of argument mining.

From the perspective of argument structure, existing research mainly focuses on the argument structure of single claims, lacking mining of argument structures among multiple claims, such as the argument structure of full scientific papers or the argument structure between original posts and comments forming multiple documents on social media. In arguments composed of multiple claims, premises and conclusions are mutually transformable—a premise of one claim is often the conclusion of another. How to reflect this role transformation of argument components and construct argument chains is a worthy research question. Moreover, current identification of implicit argument components in argument structures is not yet mature. For example, how to identify unexpressed consensus in argument contexts is important for constructing complete argument structures.

From a technical perspective, argument structure identification mainly relies on supervised learning methods requiring large amounts of high-quality annotated corpora. However, due to the complexity of argument structures themselves, corpus annotation requires substantial human and material resources, and there is currently a lack of large-scale standard corpora. Most existing studies use traditional machine learning techniques to identify argument components and extract argument relations, but these require manual construction of numerous features with weak model generalization capabilities. With continuous development of deep learning technology, its application to argument mining is a foreseeable direction. However, since deep learning has high requirements for corpus quality and scale, applying frontier artificial intelligence technologies such as meta-learning and few-shot learning to argument mining may solve the current dilemma of corpus scarcity and represents an important future research direction.

From an application perspective, current argument mining application research mainly focuses on short texts in social media, with relatively few applications for complex long texts such as scientific papers. However, long texts such as news comments, scientific papers, and legal judgments often contain more valuable argument information. Therefore, argument mining for long texts represents an important future research direction for the field.

References

[1] Patrick Hurley. A Concise Introduction to Logic [M]. Translated by Chen Bo, Song Wengan. 10th Edition. Beijing: World Book Publishing Company, 2010: 5-7. [2] Walton D. Argumentation theory: a very short introduction [G] // Argumentation in Artificial Intelligence. Boston: Springer, 2009: 1-22. [3] Moens M F, Boiy E, Palau R M, et al. Automatic detection of arguments in legal texts [C] // Proceedings of the 11th International Conference on Artificial Intelligence and Law (ICAIL). Stanford: ACM Press, 2007: 225-230. [4] Kwon N, Hovy E, Zhou L, et al. Identifying and classifying subjective claims [C] // Proceedings of the 8th Annual International Conference on Digital Government Research. Maryland: ACM Press, 2007: 76-81. [5] Palau R M, Moens M F. Argumentation mining: The detection, classification and structure of arguments in text [C] // Proceedings of the 12th International Conference on Artificial Intelligence and Law. New York: ACM, 2009: 98-107. [6] First Workshop on Argumentation Mining [EB/OL]. [2020-03-07]. https://www.uncg.edu/cmp/ArgMining2014/. [7] SICSA Workshop on Argument Mining 2014 [EB/OL]. [2020-03-07]. http://www.arg-tech.org/index.php/sicsa-workshop-on-argument-mining-2014/. [8] Mochales R P, Moens M F. Study on sentence relations in the automatic detection of argumentation in legal cases [C] // Proceedings of the 2007 Conference on Legal Knowledge and Information Systems. The Netherlands: IOS Press, 2007: 89-98. [9] Song Y, Heilman M, Beigman B, et al. Applying argumentation schemes for essay scoring [C] // Proceedings of the First Workshop on Argumentation Mining. Baltimore: Association for Computational Linguistics, 2014: 69-78. [10] Xiong Caiquan, Sun Xianbin, Ouyong Yong. A survey of logical models for debate [J]. Pattern Recognition and Artificial Intelligence, 2010, 23(3): 362-368. [11] Bentahar J, Moulin B, Bélanger M. A taxonomy of argumentation models used for knowledge representation [J]. Artificial Intelligence Review, 2010, 33(3): 211-259. [12] Toulmin S E. The Uses of Argument [M]. London: Cambridge University Press, 1958. [13] Dung P M. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and N-person games [J]. Artificial Intelligence, 1995, 77(2): 321-357. [14] Perelman C, Olbrechts-Tyteca L. The New Rhetoric: A Treatise on Argumentation [M]. Translated by Wilkinson J, Weaver P. Notre Dame: University of Notre Dame Press, 1971. [15] Freeman J B. Argument Structure: Representation and Theory [M]. Netherlands: Springer, 2011: 3-6. [16] Thomas S N. Practical Reasoning in Natural Language [M]. New Jersey: Prentice-Hall, 1986. [17] Wang Jianfang. Research on argument structure based on argumentation—A comparison between Freeman's model and Toulmin's model [J]. Logical Studies, 2016, 9(3): 42-56. [18] Zhang Binfeng, Hou Goulei. The characteristics and functions of argumentation schemes [J]. Journal of Hubei University (Philosophy and Social Science Edition), 2018,

45(6): 66-72, 176. [19] Walton D. Argumentation Schemes [M]. London: Cambridge University Press, 2008. [20] Lawrence J, Reed C. Argument mining using argumentation scheme structures [J]. Frontiers in Artificial Intelligence and Applications, 2016, 287: 379-390. [21] Stab C. Argumentative Writing Support by Means of Natural Language Processing [D]. Darmstadt: Technical University of Darmstadt, 2017. [22] Lauscher A, Glavaš G, Ponzetto S P. An argument-annotated corpus of scientific publications [C] // Proceedings of the 5th Workshop on Argument Mining. Brussels: Association for Computational Linguistics, 2018: 40-46. [23] Kirschner C, Eckle-Kohler J, Gurevych I. Linking the thoughts: Analysis of argumentation structures in scientific publications [C] // Proceedings of the 2nd Workshop on Argumentation Mining. Denver, CO: Association for Computational Linguistics, 2015: 1-11. [24] Peldszus A. Automatic Recognition of Argumentation Structure in Short Monological Texts [D]. Potsdam: University of Potsdam, 2017. [25] Stab C, Kirschner C, Eckle K J, et al. Argumentation mining from discourse structure perspective [C] // Proceedings of the Workshop on Frontiers and Connections between Argumentation Theory and Natural Language. Bertinoro, Italy: 2014: 1-10. [26] Passon M, Lippi M, Serra G, et al. Predicting the usefulness of Amazon reviews using off-the-shelf argumentation mining [C] // Proceedings of the 5th Workshop on Argument Mining. Brussels: ACL, 2018: 35-39. [27] Boltužić F, Šnajder J. Fill the gap! Analyzing implicit premises between claims from online debates [C] // Proceedings of the Third Workshop on Argument Mining. Berlin: ACL, 2016: 124-133. [28] Ghosh D, Muresan S, Wacholder N, et al. Analyzing argumentative discourse units in online interactions [C] // Proceedings of the First Workshop on Argumentation Mining. Baltimore: ACL, 2014: 39-48. [29] Boltužić F, Šnajder J. Back up your stance: Recognizing arguments in online discussions [C] // Proceedings of the First Workshop on Argumentation Mining. Baltimore: ACL, 2014: 49-58. [30] Morio G, Fujita K. Annotating online civic discussion threads for argument mining [C] // International Conference on Web Intelligence 2018. Marcachile: IEEE, 2018: 546-553. [31] Hidey C, Musi E, Hwang A, et al. Analyzing the semantic types of claims and premises in an online persuasive forum [C] // Proceedings of the 4th Workshop on Argument Mining. Copenhagen: ACL, 2017: 11-21. [32] Accuosto P, Saggion H. Discourse-driven argument mining in scientific abstracts [C] // Natural Language Processing and Information Systems. Cham: Springer, 2019: 182-194. [33] Lippi M, Torroni P. Argumentation mining: State of the art and emerging trends [J]. ACM Transactions on Internet Technology, 2016, 16(2): 1-25. [34] Goudas T, Louizos C, Petasis G, et al. Argument extraction from news, blogs, and social media [C] // Artificial Intelligence: Methods and Applications. Greece: Springer, 2014: 287-299. [35] Florou E, Konstantopoulos S, Koukourikos A, et al. Argument extraction for supporting public policy formulation [C] // Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities. Sofia: ACL, 2013: 49-54. [36] Dusmanu M, Cabrio E, Villata S. Argument mining on Twitter: Arguments, facts and sources [C] // Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing.

Copenhagen: ACL, 2017: 2317-2322. [37] Reimers N, Schiller B, Beck T, et al. Classification and clustering of arguments with contextualized word embeddings [C] // Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence: ACL, 2019: 567-578. [38] Ferrara A, Montanelli S, Petasis G. Unsupervised detection of argumentative units through topic modeling techniques [C] // Proceedings of the 4th Workshop on Argument Mining. Copenhagen: ACL, 2017: 97-107. [39] Lawrence J, Reed C, Allen C, et al. Mining arguments from 19th century philosophical texts using topic-based modeling [C] // Proceedings of the First Workshop on Argumentation Mining. Baltimore: ACL, 2014: 79-87. [40] Sardianos C, Katakis I M, Petasis G, et al. Argument extraction from news [C] // Proceedings of the 2nd Workshop on Argumentation Mining. Denver: ACL, 2015: 56-66. [41] Li M, Gao Y, Wen H, et al. Joint RNN model for argument component boundary detection [EB/OL]. [2020-03-07]. http://arxiv.org/abs/1705.02131. [42] Petasis G. Segmentation of argumentative texts with contextualized word representations [C] // Proceedings of the 6th Workshop on Argument Mining. Florence: ACL, 2019: 1-10. [43] Trautmann D, Daxenberger J, Stab C, et al. Fine-grained argument unit recognition and classification [C] // The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20). New York: AAAI, 2020: 9048-9056. [44] Lippi M, Torroni P. Context-independent claim detection for argument mining [C] // Proceedings of the 24th International Joint Conference on Artificial Intelligence. Argentina: AAAI Press, 2015: 185-191. [45] Eckle-Kohler J, Kluge R, Gurevych I. On the role of discourse markers for discriminating claims and premises in argumentative discourse [C] // Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon: ACL, 2015: 2236-2242. [46] Stab C, Gurevych I. Identifying argumentative discourse structures in persuasive essays [C] // Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. Doha: ACL, 2014: 46-56. [47] Lugini L, Litman D. Argument component classification for classroom discussions [C] // Proceedings of the 5th Workshop on Argument Mining. Brussels: ACL, 2018: 57-67. [48] Dutta S, Das D, Chakraborty T. Changing views: Persuasion modeling and argument extraction from online discussions [J]. Information Processing & Management, 2019, 57(2): 1-14. [49] García-Gorrostieta J M, Lopez-Lopez A. Argument component classification in academic writings [J]. Journal of Intelligent & Fuzzy Systems, 208, 34(5): 3037-3047. [50] Stab C, Gurevych I. Parsing argumentation structures in persuasive essays [J]. Computational Linguistics, 2017, 43(3): 619-659. [51] Aker A, Sliwa A, Ma Y, et al. What works and what does not: Classifier and feature analysis for argument mining [C] // Proceedings of the 4th Workshop on Argument Mining. Copenhagen: ACL, 2017: 91-96. [52] Lawrence J, Reed C. Combining argument mining techniques [C] // Proceedings of the 2nd Workshop on Argumentation Mining. Denver: ACL, 2015: 127-136. [53] Dagan I, Dolan B, Magnini L B, et al. Recognizing textual entailment: Rational, evaluation and approaches [J]. Journal of Natural Language Engineering, 2009, 15(4): 1-17. [54] Cabrio E, Villata S. Combining textual entailment and argumentation theory for supporting online debate interaction [C] // Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics. Korea: ACL, 2012: 208-212. [55] Hou Y, Jochim C. Argument relation classification using a joint inference model [C] // Proceedings of the 4th Workshop on Argument Mining. Copenhagen: ACL, 2017: 60-66. [56] Galassi A, Lippi M, Torroni P. Argumentative link prediction using residual networks and multi-objective learning [C] // Proceedings of the 5th Workshop on Argument Mining. Brussels: ACL, 2018: 1-10. [57] Nguyen H, Litman D. Context-aware argumentative relation mining [C] // Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Berlin: ACL, 2016: 1127-1137. [58] Mitrović J, O'Reilly C, Mladenović M, et al. Ontological representations of rhetorical figures for argument mining [J]. Argument & Computation, 2017, 8(3): 267-287. [59] Green N L. Representation of argumentation in text with rhetorical structure theory [J]. Argumentation, 2010, 24(2): 181-202. [60] Rajendran P, Bollegala D, Parsons S. Contextual stance classification of opinions: A step towards enthymeme reconstruction in online reviews [C] // Proceedings of the Third Workshop on Argument Mining. Berlin: ACL, 2016: 31-39. [61] Green N. Manual identification of arguments with implicit conclusions using semantic rules for argument mining [C] // Proceedings of the 4th Workshop on Argument Mining. Copenhagen: ACL, 2017: 73-78. [62] Green N. Towards mining scientific discourse using argumentation schemes [C] // Proceedings of the 2nd Workshop on Argumentation Mining. Denver: ACL, 2015: 12-21. [63] Niven T, Kao H-Y. Probing neural network comprehension of natural language arguments [C] // Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence: ACL, 2019: 4658-4664. [64] Peldszus A, Stede M. Joint prediction in MST-style discourse parsing for argumentation mining [C] // Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon: ACL, 2015: 938-948. [65] Afantenos S, Peldszus A, Stede M. Comparing decoding mechanisms for parsing argumentative structures [J]. Argument & Computation, 2018, 9(3): 177-192. [66] Walker V R, Foerster D, Ponce J M, et al. Evidence types, credibility factors, and patterns or sorts of rules for weighing conflicting evidence: Argument mining in the context of legal rules governing evidence assessment [C] // Proceedings of the 5th Workshop on Argument Mining. Belgium: ACL, 2018: 68-78. [67] Al-Abdulkarim L, Atkinson K, Bench-Capon T. Statement types in legal argument [C] // Frontiers in Artificial Intelligence and Applications. Hague: IOS Press, 2016: 3-12. [68] Persing I, Ng V. Modeling argument strength in student essays [C] // Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Beijing: ACL, 2015: 543-552. [69] Wachsmuth H, Naderi N, Habernal I, et al. Argumentation quality assessment: Theory vs. Practice [C] // Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. Vancouver: ACL, 2017: 250-255. [70] Wachsmuth H, Al-Khatib K, Stein B. Using argument mining to assess the argumentation quality of essays [C] // Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. Japan: The

COLING 2016 Organizing Committee, 2016: 1680-1691. [71] Persing I, Ng V. Why can't you convince me? Modeling weaknesses in unpersuasive arguments [C] // Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. Melbourne: International Joint Conferences on Artificial Intelligence Organization, 2017: 4082-4088. [72] Konat B, Lawrence J, Park J, et al. A corpus of argument networks: Using graph properties to analyze divisive issues [C] // Proceedings of the 10th International Conference on Language Resources and Evaluation. Portoro: LREC, 2016: 3899-3906. [73] Liebeck M, Esau K, Conrad S. What to do with an airport? Mining arguments in the German online participation project Tempelhofer Feld [C] // Proceedings of the Third Workshop on Argument Mining. Berlin: ACL, 2016: 144-153. [74] Teufel S, Moens M. Summarizing scientific articles: Experiments with relevance and rhetorical status [J]. Computational Linguistics, 2002, 28(4): 409-445. [75] Yepes A J, Mork J, Aronson A. Using the argumentative structure of scientific literature to improve information access [C] // Proceedings of the 2013 Workshop on Biomedical Natural Language Processing. Sofia: ACL, 2013: 102-110. [76] Green N. Identifying argumentation schemes in genetics research articles [C] // Proceedings of the 2nd Workshop on Argumentation Mining. Denver: ACL, 2015: 12-21. [77] Mayer T, Cabrio E, Villata S. Evidence type classification in randomized controlled trials [C] // Proceedings of the 5th Workshop on Argument Mining. Brussels: ACL, 2018: 29-34.

Author Contributions: Li Yongze was responsible for data collection and paper writing; Ou Shiyan identified the research direction and ideas, and contributed to paper revision, review, and finalization.

Editor's Note: This guideline is formulated based on the journal's positioning, nature, and development needs, combined with frontier hotspots in library, information, and archival science and important problems that need to be solved currently and in the future. It was planned and customized by the editorial board members and young editorial board members of this journal, and then organized, revised, and supplemented by the editorial department. These are the key areas (including but not limited to these topics) that this journal will focus on and report in 2020, providing reference for authors in topic selection, research, and submission to this journal.

1. Construction of disciplinary system, academic system, and discourse system for library, information, and archival science with Chinese characteristics
2. Development strategy for first-level discipline construction and integration in library, information, and archival science
3. Key issues in the formulation of library "14th Five-Year" plans
4. National literature and information resource guarantee capability and its construction
5. Information resource construction issues in the context of open science
6. Library's positioning and responsibility in national reading promotion

7. Theory and practice of library space services
8. Performance evaluation and management of embedded subject services
9. Citizen science, scientific literacy, and pan-information literacy
10. Models and capabilities of library services for undergraduate education
11. Theory and practice of library cultural inheritance and cultural education
12. Library publishing and publishing services
13. Function and practice of library science communication in the new media era
14. Strategy and tactics of library marketing and promotion
15. Practice and theory of library pan-cooperation research
16. Construction and innovative services of library alliances under national regional development strategies
17. Informatics issues in cyberspace governance
18. Capability and effect evaluation of intellectual property information services
19. New technologies and methods in information analysis
20. Standardization and evaluation of intelligence services
21. Research and practice of digital humanities and digital scholarship
22. Application of artificial intelligence in library, information, and archival science
23. Library intelligent services and smart services
24. Research on metadata development models in open data ecosystems
25. Open scientific data behavior and its model construction
26. Data resource construction and data librarian capability building
27. Information organization and knowledge organization in the big data era
28. Scientific data management and services
29. Academic achievement monitoring and subject competitiveness analysis
30. Theory and methods of intelligence computing (computational intelligence)
31. Quality and effectiveness evaluation of intelligence analysis services
32. Relationship between intelligence research and think tank research
33. Theory and methods for analyzing science and technology frontiers
34. Health informatics under the Healthy China 2030 strategy
35. Human-computer interaction behavior and service model innovation
36. Mechanism of library, information, and archival science in new think tank construction
37. Theory and methods of intelligent information services
38. Digital public cultural resources, services, and system construction
39. Government information resource management and development in the data era
40. Governance strategies for digital archives ecosystem
41. Theory and system of archival data governance
42. Application and evaluation of government data open platforms
43. Organization, protection, and development of archival information resources from the perspective of social memory
44. Industrial development and utilization of ethnic documentary heritage

45. Education models and talent cultivation capabilities in library, information, and archival science

Figures

Free Q&A:

$$Y_{\text{Answer\_Length}} = \beta 0 + \beta 1 X_{\text{Question\_Length}} + \beta 2 X_{\text{Factual}} + \beta 3 X_{\text{Suggestion}} + \beta 4 X_{\text{Subjective\_Recommendation}} + \beta 5 X_{\text{Objective\_Recommendation}} + \beta 6 X_{\text{Self\_Expression}} + \beta 7 X_{\text{Psychology}} + \beta 8 X_{\text{Game}} + \beta 9 X_{\text{Medicine}} + \beta 10 X_{\text{Question\_Pablish\_Time}} \quad \text{(Formula 1)}$$

$$Y_{\text{Answer\_Count}} = \beta 0 + \beta 1 X_{\text{Question\_Length}} + \beta 2 X_{\text{Factual}} + \beta 3 X_{\text{Suggestion}} + \beta 4 X_{\text{Subjective\_Recommendation}} + \beta 5 X_{\text{Objective\_Recommendation}} + \beta 6 X_{\text{Self\_Expression}} + \beta 7 X_{\text{Psychology}} + \beta 8 X_{\text{Game}} + \beta 9 X_{\text{Medicine}} \quad \text{(Formula 2)}$$

$$Y_{\text{Evaluation\_Count}} = \beta 0 + \beta 1 X_{\text{Question\_Length}} + \beta 2 X_{\text{Answer\_Length}} + \beta 3 X_{\text{Factual}} + \beta 4 X_{\text{Suggestion}} + \beta 5 X_{\text{Subjective\_Recommendation}} + \beta 7 X_{\text{Self\_Expression}} + \beta 8 X_{\text{Psychology}} + \beta 9 X_{\text{Game}} + \beta 10 X_{\text{Medicine}} + \beta 11 X_{\text{Question\_Publish\_Time}} + \beta 12 X_{\text{Answer\_Publish\_Time}} \quad \text{(Formula 3)}$$

Paid Q&A:

$$Y_{\text{Answer\_Length}} = \beta 0 + \beta 1 X_{\text{Historical\_Answer\_Count}} + \beta 2 X_{\text{Price}} + \beta 3 X_{\text{Question\_Length}} + \beta 4 X_{\text{Factual}} + \beta 5 X_{\text{Suggestion}} + \beta 6 X_{\text{Subjective\_Recommendation}} + \beta 7 X_{\text{Objective\_Recommendation}} + \beta 8 X_{\text{Self\_Expression}} + \beta 9 X_{\text{Psychology}} + \beta 10 X_{\text{Game}} + \beta 11 X_{\text{Medicine}} + \beta 12 X_{\text{Question\_Publish\_Time}} \quad \text{(Formula 4)}$$

$$Y_{\text{Evaluation\_Count}} = \beta 0 + \beta 1 X_{\text{Historical\_Answer\_Count}} + \beta 2 X_{\text{Price}} + \beta 3 X_{\text{Question\_Length}} + \beta 4 X_{\text{Answer\_Length}} + \beta 5 X_{\text{Factual}} + \beta 6 X_{\text{Suggestion}} + \beta 7 X_{\text{Subjective\_Recommendation}} + \beta 8 X_{\text{Objective\_Recommendation}} + \beta 9 X_{\text{Self\_Expression}} + \beta 10 X_{\text{Psychology}} + \beta 11 X_{\text{Game}} + \beta 12 X_{\text{Medicine}} + \beta 13 X_{\text{Question\_Publish\_Time}} \quad \text{(Formula 5)}$$

Figure 6: Figure 7

Source: ChinaXiv — Machine translation. Verify with original.