Linking Scholarly Contents: The Design and Construction of an Argumentation Graph†

Ningyuan Song*, Hanghang Cheng **, Huimin Zhou *** and Xiaoguang Wang ****

*School of Information Management, Nanjing University, 163 Xianlin Road, Nanjing, China, 210093, <songny@nju.edu.cn>

School of Information Management, Wuhan University, 299 Bayi Road, Wuhan, China, 430072, **<chhang@whu.edu.cn>, ***<zhouhm@whu.edu.cn>, ****<wxguang@whu.edu.cn>

Ningyuan Song received his Ph.D. degree from Wuhan University. Currently, he is engaged in post-doctoral research at the School of Information Management, Nanjing University. His main research directions include improving the content organization form of scientific and technological publications by using semantic technology, semantic organization of digital resources in the context of digital humanity, and knowledge management and user behavior in social media.

Cheng Hanghang is a Master's graduate of the School of Information Management, Wuhan University. Research fields include semantic publishing, data mining, and knowledge graph. Relevant papers were published in international conferences and journals such as JASIST, JCDL, and ASIST Conference.

Zhou Huimin is a Master's graduate of the School of Information Management, Wuhan University. Research fields include semantic publishing and information structure of scientific papers. Relevant papers were published in international conferences and journals such as JASIST, JCDL, and ASIST Conference.

Wang Xiaoguang is Vice Dean of the School of Information Management of Wuhan University, Professor, Director of the Philosophy and Social Sciences Laboratory of the Ministry of Education for Cultural Heritage and Intelligent Computing of Wuhan University, Director of the Digital Humanities Research Center of Wuhan University, serves as the director of the Chinese Information Society of Social Sciences, the executive director of the China Audio-Video and Digital Publishing Association (CADPA), the core member of the innovation team of NSFC, head of the innovation team of Hubei Natural Science Foundation, and chief expert of the major program of the National Social Science Foundation.

Song, Ningyuan, Hanghang Cheng, Huimin Zhou and Xiaoguang Wang. 2022. "Linking Scholarly Contents: The Design and Construction of an Argumentation Graph." *Knowledge Organization* 49(4): 213-235. 53 references. DOI:10.5771/0943-7444-2022-4-213.

Abstract: In this study, we propose a way to link the scholarly contents of scientific papers by constructing a knowledge graph based on the semantic organization of argumentation units and relations in scientific papers. We carried out an argumentation graph data model aimed at linking multiple discourses, and also developed a semantic annotation platform for scientific papers and an argumentation graph visualization system. A construction experiment was performed using 12 articles. The final argumentation graph has 1,262 nodes and 1,628 edges, including 1,628 intra-article relations and 190 inter-article relations. Knowledge evolution representation, strategic reading, and automatic abstracting use cases are presented to demonstrate the application of the argumentation graph. In contrast to existing knowledge graphs used in academic fields, the argumentation graph better supports the organization and representation of scientific paper content and can be used as data infrastructure in scientific knowledge retrieval, reorganization, reasoning, and evolution. Moreover, it supports automatic abstract and strategic reading.

Received 20 November 2021; revised 21 June 2022; accepted 16 August 2022

Keywords: scientific papers, argumentation, argumentation graph, knowledge graph, knowledge evolution

† This research is funded by the Science Fund for Creative Research Groups of NSFC (71921002), the Science Fund for Creative Research Groups of Natural Science Fund of Hubei Province (2019CFA025), the General Program of NSFC (1874129), and the China Postdoctoral Science Foundation (2021M691534).

1.0 Introduction

Scientific papers are still crucial for constructing and disseminating scientific knowledge and exchanging scientific ideas and views. The increase of papers burdens reading, especially in document collections of specific research fields. To master the construction process of domain knowledge, users must repeatedly jump between different articles and quickly understand the structure and related content (Liu 2005). Strategic reading (Renear and Palmer 2009) is a recent approach to reading and understanding literature. By using computer technology to analyse and organize documents and generate new summary content, users could be supported to adopt a non-linear reading order and combine skipping and skimming to obtain necessary information and knowledge. Therefore, it is essential to build assistant tools that could support users in adopting strategic reading. This needs to help users quickly understand the current paper's structure and content and the process of knowledge construction of a research domain.

The semantic representation, annotation, and organization at the sentence level have also become fundamental to facilitating strategic reading, and research has made achievements in related aspects. The ontologies suitable for representing the semantics of statement (DEO, DoCO, AMO, etc.) (Bagnacani et al. 2015) and semantic content representation models (nanopublication, micropublication, etc.) have been developed (Groth et al. 2010; Clark et al. 2014). Meanwhile, as an essential organizational tool, the knowledge graph has been used to link scholarly data or content. Many publishing groups and research institutions have built graphs based on bibliographic data, such as SciGraph[1], Microsoft academic Graph (MAG)[2], Aminer[3] and Open Academic Graph (OAG).[4] Besides these graphs, researchers started to build up new scientific knowledge graphs based on parsing the content of scientific papers (Auer et al. 2018; Moser and Mercer 2020). These graphs can play an essential role in knowledge discovery and research trend identification, and so on. But they could not better support users' strategic reading because of the lack of representation of the content and structure of the scientific papers.

The argumentation structures within scientific papers are considered essential and fundamental for constructing scientific theory, and they provide the necessary clue for reading scientific papers(Walton and Zhang 2013). In a specific research field, different researchers tend to put forward research problems, reuse or improve research methods by arguing or citing other papers, and the various research results would be comprehensively compared in the discussion part. Thus, the argument between multiple articles could also characterize the construction process of domain knowledge.

From these perspectives, representing the argumentation structure of scientific papers among document collections, reorganizing the scientific content by using an argumentation graph, and visualizing it could present the knowledge construction process in the current research field for users, help users quickly summarize the discovery process of significant conclusions, and then assist users in strategy reading in the semantic enhancement environment.

To construct the argumentation graph of multiple scientific papers, several research questions should be solved:

- RQ1: How to represent the cross-discourse argumentation structure within the context of scientific papers.
- RQ2: How to construct the argumentation graph, and which approach should the construction work follow.
- RQ3: How to visualize the argumentation graph to represent the argumentation structure.

Aimed at addressing these problems, a data model for argumentation graph was proposed in this research. We have also built two system prototypes to construct the multiple discourse argumentation graph: an annotation platform and an argumentation graph visualization system. On this basis, 12 scientific papers related to the Technical Acceptance Model (TAM) are selected as samples to carry out the argumentation graph construction experiment. Three use cases, knowledge evolution representation, strategic reading, and automatic abstracting, are given.

2.0 Literature review

2.1 Computational modeling and visualization of arguments

Argumentation theory originated from philosophy and dialectics in the fourth century B.C.E. It has now developed into an interdisciplinary research field including philosophy, communication, linguistics, psychology, law, artificial intelligence, and computer science (Lippi and Torroni 2016). According to the definition in the Merriam-Webster Dictionary, argumentation is the act or process of forming reasons, drawing conclusions, and applying them to a case in discussion. In the 1950s, Toulmin put forward a general conceptual model to represent argument, and the model consists of six parts: *data* (or *evidence*), *warrant, backing, rebuttal, qualifier,* and *claim*. Among them, claim, warrant, and data are the core parts of the argumentation process. Backing, qualifier, and rebuttal are supplementary elements, which can be adjusted according to the specific situation(Toulmin 2003). Besides the Toulmin model of argument, Kunz and Rittel (1970) pointed out that the argumentation process should take the controversial question as the core, and in the model of Issue-Based Information Systems (IBIS), a visual structure of argument was carried out, including the components such as the *main contention*, *premises*, *co-premises*, *objections*, *rebuttals*, and *lemmas* (Kunz and Rittel 1970). Subsequently, Tweed used the general framework of reasoning (Toulmin, Rieke and Janik 1984) to further detail IBIS for knowledge representation in designing regulations and standards and proposed the microstructure between assertion and argumentation nodes (Tweed 1994), which made the IBIS representation of knowledge can describe argumentation and reasoning with a clearer definition of relations. Dung regards argumentation as a directed graph, treated the argumentation unit as a node, the relationship as an edge, and gave a formal representation of argumentation (Dung 1995). In 2001, Freeman sorted out five argument structures, namely *convergent, divergent, hybrid, linked,* and *serial*, which reveal the argumentation relationships more intuitively (Freeman 2001).

Starting from the basic theory, the thought, structure, and model of argumentation have been used in the representation of content by different domains, such as law (Mochales and Moens 2011), online information communities (Lange et al. 2008), artificial intelligence (Bench-Capon and Dunne 2007; Rahwan and Simari 2009), business management (Klein 2007), further enriches the practical application of argumentation theory (Schneider, Groza and Passant 2013).

Besides the theoretical exploration of argumentation, some research focused on visualization by utilizing graphs. The argumentation graph (or diagram) was constructed to represent the logical or process of argumentative activities, such as reasoning, inferences, debates, and cases (Van Gelder 2009). In 1988, David Kelley proposed a structural diagram to represent the different argument patterns with numbered premises and arrows indicating inferential relations, and three patterns including *serial, additive*, and *non-additive*, were recognized (Kelley 1988). In the 1990s, Tim Van Gelder developed a series of computer applications that permitted the claims to be fully stated and edited in the diagram (Van Gelder 2007). Meanwhile, the visualization of argumentation has been wildly applied in educational (Harrell 2005), academic, law, artificial intelligence (Reed, Walton and Macagno 2007), and business areas (Kirschner, Shum and Carr 2003; Okada, Shum and Sherborne 2008), showed great potential to improve the comprehensibility of the text (Peldszus and Stede 2013).

2.2 The representation of argumentation in scientific papers

From the perspective of argumentation, many research studies on knowledge representation of scientific paper argumentation have carried out formal modeling for scientific statements and explored the formal representations of the argumentation unit, argumentation relationship, argumentation process, and argumentation structure, which laid a solid theoretical basis for the research on the fine-grained organization of the scientific paper argumentation graph.

In terms of argumentation structure, research mainly focuses on the definition and representation of argumentation unit and argumentation relationship. As for the argumentation unit, Teufel put forward the Argumentation Zoning model (Teufel 1999), which defined the argumentation structure as a series of argumentation units with different functions, including *aim, contrast, basis,* and *background*, also distinguished the research of others (*other*) and own researches (*own*). Subsequently, Argumentation Zoning II (Teufel 2010) was put forward to emphasize the argumentation process, and new classes such as *CoDI, gap weak,* and *support* were designed. The model holds that the internal logic of a scientific paper is the comparison between the old and the new ideas and that it is based on the existing research and around the established scientific objectives to demonstrate new ideas and new theories. In addition, Nancy Green identified argumentation schemes in genetics research articles (Green 2010; 2015). As for argumentation relationships within scientific papers, Kirschner et al. (2015) annotated argumentation structures in the Introduction and Discussion sections of scientific articles and defined four argumentation relationships (*support, attack, detail*, and *sequence*).

Ontology was also used to represent the argumentation structure; in Scholarly Ontology, Shum, Motta and Domingue (2000) defined the relationship between claims as the *causal*, *problem related, similarity, general, support/challenge,* and *classification.* Based on Toulmin's Model, the Argument Model Ontology (AMO) defined *claim, warrant, backing, qualifier, rebuttal*, etc. In addition, in the Micropublication model, classes such as *method* and *data* were described as argumentation units, and the relationships between statement and data were also defined (Clark, Ciccarese and Goble 2014).

Other studies and ontologies understand the content structure of scientific papers from semantic and pragmatic perspectives. As a W3C standard, the Ontology of Rhetorical Blocks (ORB) describes scientific papers' metadata and rhetorical structure at a coarse-grained level. Based on ORB, the Document Element Ontology (DEO) explicitly defines the *introduction, discussion, reference list, figure, and appen-*

dix. The experiment process is described in EXPO and CoreSCs, including *hypothesis, motivation, goal, object, background, method, experiment, model, observation, result, conclusion*, etc (Soldatova and King 2006; Liakata et al. 2012).

2.3 Knowledge graph for scientific papers

A knowledge graph refers to the semantic knowledge base with a directed graph structure including entities, concept nodes, and semantic relationships. It processes structured, semi-structured, and unstructured data into semantic knowledge through data extraction, processing, fusion, and other processes and uses the organizational form of knowledge to store data logically (Saruladha 2011) to support the management, retrieval, use, and understanding of knowledge (Yan et al. 2018). On the one hand, the knowledge graph organizes and represents the relations and constraints between the upper concepts through ontology; on the other hand, it extracts, organizes, and stores the entities in the data set with triples as the basic unit, thus forming a knowledge network.

The construction of the knowledge graph generally includes three processes: knowledge acquisition, knowledge fusion, and knowledge processing. Knowledge acquisition refers to extracting structured knowledge such as entities, relationships, and entity attributes from unstructured, semi-structured, or structured data (Cowie and Lehnert 1996). Knowledge fusion can eliminate the concepts of error, redundancy, and ambiguity and improve the quality of knowledge through semantic annotation, the vector space model, singular value decomposition, and other technologies such as entity disambiguation and coreference resolution (Bagga and Baldwin 1998; Turney 2001; Pedersen, Purandare and Kulkarni 2005; Han and Zhao 2009; Sen 2012). Knowledge processing is an operation of processing entities, relationships, and attributes extracted from data sets to obtain a structured and networked knowledge system.

The knowledge graph has been used widely in different areas as the data infrastructure to support tasks such as semantic search and intelligent Q&A systems. Academic knowledge graphs such as SciGraph, MAG, Aminer, OAG, and SciKGraph have been employed to evaluate and predict the influence of authors, articles, and journals, discovering author promotion trends, expert opinion search, author relationship prediction, academic information recommendation, and academic institution detection. In addition, using these knowledge graphs, some researchers have explored the construction of a domain knowledge base. Sadeghi integrated MAG and DBLP bibliographic metadata to construct the domain knowledge base of academic exchange (SCM-KG) (Sadeghi et al. 2017). In recent years, some researchers have begun to use knowledge graphs to represent the content of papers, which provides a basis for our research. Auer et al. (2018) used triples to describe the knowledge, and based on this, built up the Open Research Knowledge Graph. Moser and Mercer (2020) used the claim graph to represent the argumentation structure within biomedical literature.

As a necessary means of knowledge organization, knowledge graph construction technology and methods are relatively mature. Still, these existing academic knowledge graph does not go deep into the content structure of scientific papers. The lack of representation of the content and structure of the paper makes it very difficult for users to quickly locate and read the content of the paper using a knowledge graph. For this reason, in our study, we tend to use the knowledge graph to realize the semantic integration and reorganization of scientific paper content.

3.0 Data model design

In our previous work, Zhou et al. (2019) integrated the characteristics of scientific papers. They designed a more comprehensive ontology named the Scientific Paper Argumentation Ontology (SAO) to describe the argumentation structure of scientific papers. By referencing and reusing AMO, DEO, Core Information about Scientific Papers (CISP), EXPO, and the CoreSC concept model, SAO consists of seven core classes, 13 extension classes, and 15 relationships between the argumentation structures, as shown in Figure 1, and the contextual characteristics of scientific papers are fully considered.

This study incorporates the SAO ontology as the primary reuse ontology, referencing and reusing the Citation Type Ontology (CiTO), Document Component Ontology (DoCO), and DCMI classes and relationships to design the argumentation graph data model for scientific papers.

This study considers the basic structure of the argumentation graph of scientific papers to consist of an intra-article argumentation structure and inter-article argumentation relations. The inter-article argumentation structure is characterized using SAO, whereas the inter-part argument relations are multiplexed using the relationships defined by CiTO. The organizational model of scientific paper argumentation is shown in Figure 2. This model consists of 18 classes, 12 intra-article argumentation relationships, 16 inter-article relationships, 21 class attributes, and six relationship attributes.

3.1 Class design

A class of argumentation graph organization model is a set of argumentation graph entities with the same semantic characteristics that refer to the argumentation units in the scientific papers. There are 18 classes, including *background, claim,*

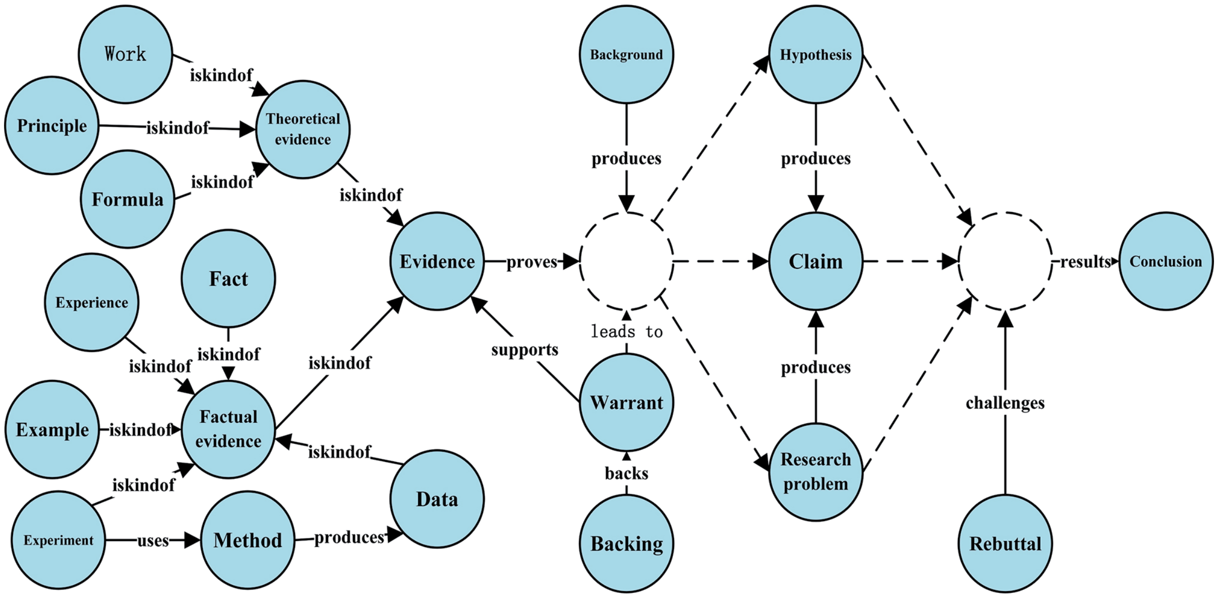

Figure 1. The SAO structure.

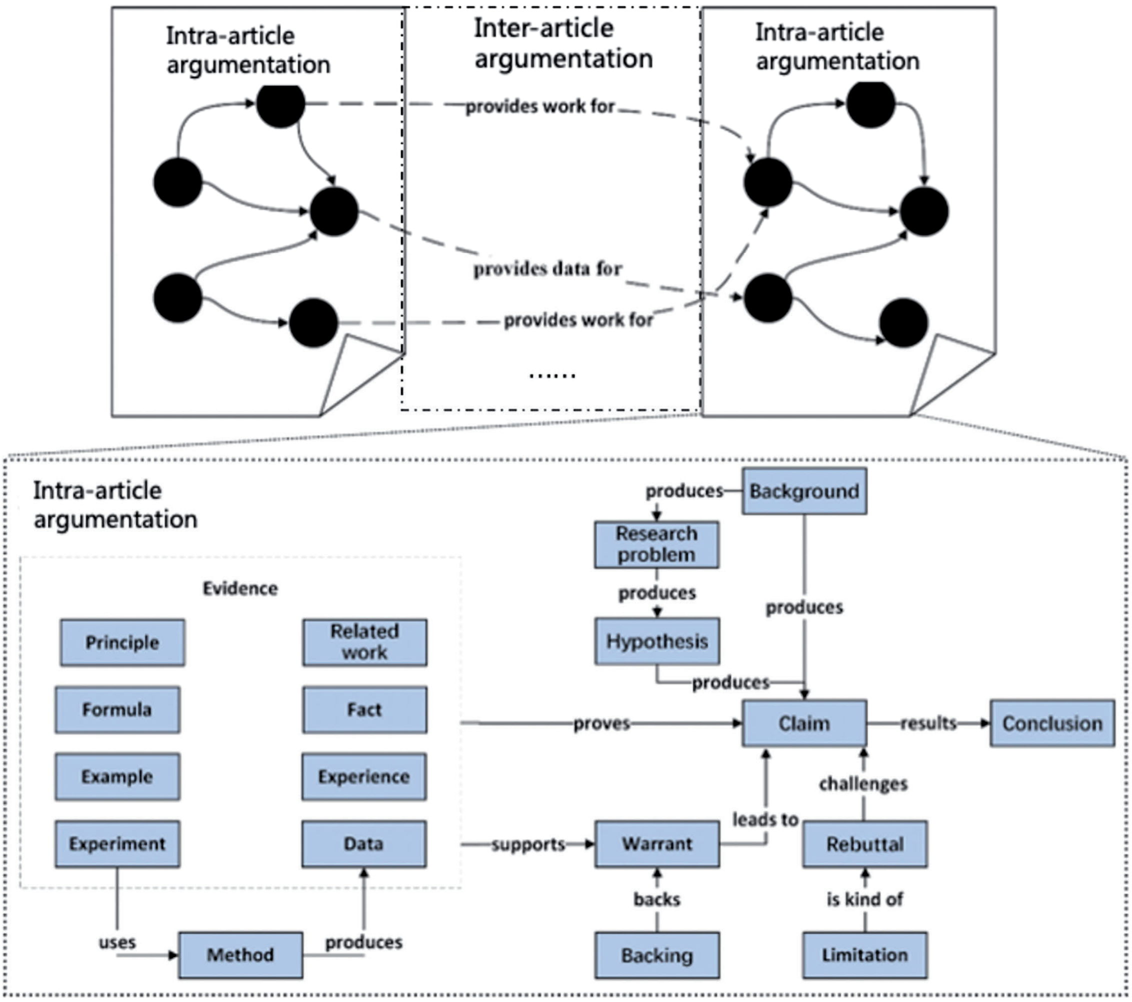

Figure 2. The basic data model of argumentation graph for scientific papers (Note: There are other 14 inter-article argumentation relationships not shown in the figure).

warrant, backing, rebuttal, conclusion, etc. The definitions of each of these classes can be seen in Table 1.

3.2 Relationship design

A relationship of the organization model of the argumentation graph contains certain argumentation semantics which connect the class or entity nodes. Based on SAO ontology and Dublin Core Element Set (DCT), this study designs 12 intra-article argumentation relations; 18 inter-article argumentation relations are also designed by reusing CiTO ontology.

3.2.1 Intra-article Argumentation Relation

Using the class definitions, 12 intra-article argumentation relationships of the argumentation graph for scientific papers were designed (Table 2). This study reuses nine of the argu mentation relations defined by SAO and redefines the meaning of relations according to the different organization methods and application scenarios. In addition, in the design of the argumentation graph, the discourse and formal structure of scientific papers must be considered when representing the intra-article relationships. From the perspective of discourse structure, the coherence of scientific papers includes argumentation relationships between different text spans and the relationships among text, paragraphs, and sentences, so the relations *is_part_of* and *is_replaced_by* are adopted from DC to represent the relations other than argumentation relations. Meanwhile, from the perspective of formal structure, the contents of scientific papers include tables, pictures, formulas, code, and other non-textual contents (Bishop 1999). These non-textual contents also play a vital role in the argumentation process. The design of intra-article relations needs to map and associate the explanatory text and the non-textual argumentation units, so *is_format_of* was also defined.

Class	Definition
Background	The initial description that states the purpose and goals of the subsequent text.
Claim	The acceptable thesis that the author tries to confirm in the argumentation, is also the core assertion of scientific paper argumentation from the perspective of full text.
Warrant	The bridge between evidence and claim which ensures that the claim can be derived from evidence rationally.
Backing	The additional supporting statement provided to strengthen the authority in response to the challenge of justifiable reasons.
Rebuttal	Some aspects of the assertion are refuted and explained.
Conclusion	The conclusion and opinion about the results are summarized, and inferred, corresponding to the claim.
Research problem	A statement of the main issues studied, and the objectives intended to be achieved.
Hypothesis	The assumption of an unproven phenomenon or fact.
Related work	The materials (principles, documents, formulas, etc.), which are cited to demonstrate a certain statement, to explain the correctness or error.
Principle	The principle of objective existence.
Formula	Expressed by mathematical symbols, illustrate the relationship between variables
Data	Factual data and explanation and analysis of data.
Experiment	One of the basic methods of scientific research, the scientific process of discovering, testing hypotheses or proving known facts.
Experience	Knowledge or skill acquired from many practices.
Fact	Objective facts that do not depend on people's subjective consciousness.
Example	A representative instance used to help illustrate or prove a situation or statement.
Method	Methods used in experiments or research and methods used to obtain experimental data.
Limitation	The deficiency of the research.

Table1. Classes and definitions of Argumentation Graph Data Model.

Relation	Start Node	End Node	Reference
is part of	Any argumentation units	Argumentation units that summarize the content of start unit	DCT
is replaced by	Any argumentation units	Argumentation units with the same semantic as the start unit	
is format of	Any argumentation units with the format of Figure, Table etc.	Textual content corresponding with figure, table etc.	
supports	Related Work Formula Principle Fact Example Experiment Data Experience	Warrant	SAO
proves	Related Work Formula Principle Fact Example Experiment Data Experience	Claim	
challenges	Rebuttal	Claim	
produces	Background	claim	
		Research Problem	
	Research Problem	Hypothesis	
	Hypothesis	claim	
	Experiment	Data	
backs	Backing	Warrant	
results	Claim	Conclusion	
leads to	Warrant	Claim/Hypothesis	
uses	Experiment	Method	
is kind of	Limitation	Rebuttal	

Table 2. Intra-article relations of argumentation graph.

3.2.2 Inter-article argumentation relation

The new knowledge within scientific papers inherits and develops the knowledge in the cited article. Therefore, the argumentation graph of scientific papers should also consider the relationships among articles, that is, it needs to model the inter-article argumentation relationships across articles. The citation of scientific papers has the semantics of recognition, reference, inheritance, amendment, refutation, and criticism, which can be used to express the argumentation relations among documents.

In this research, CiTO was adopted to represent the inter-article relations. Referring to the definition rules of CiTO rhetorical reference relations, 11 inter-article argumentation relations were also defined according to the types of ending nodes. The definitions of each relation are shown in Table 3.

Relation	Start Node	End Node	Reference
gives background to	Argumentation units	Background	CiTO
provides conclusions for		Conclusion	
provides data for		Data	
provides method for		Method	
provides assertion for		Claim	
provides example for		Example	Newly defined relationships
provides experience for		Experience	
provides experiment for		Experiment	
provides formula for		Formula	
provides principle for		Principle	
provides work for		Work	
provides backing for		Backing	
provides warrant for		Warrant	
provides rebuttal for		Rebuttal	
provides research problem for		Research problem	
provides hypothesis for		Hypothesis	

Table 3. The inter-article relations of argumentation graph.

3.3 Attribution design

3.3.1 Class attributions

Based on the DCT, SAO ontology and DoCO ontology, four categories of attributions and related value of classes were designed in this research, including basic attributions, article attributions, context attributions and location attributions, with a total of 21 attributions (Table 4). The Basic attributions provide the basic information descriptions of argumentation units, and are *id, name, content, author, date, createDate, identifier, language, format* and *formatSource*. The Article attributions include the metadata information (*articleID, articleTitle, Subject, Type,* and *Source*) of the argumentation units. These attributions, inherited from scientific papers, are designed to associate scientific papers, journals, and other entities. The Context attributions describe the implicit knowledge of argumentation units. They consist of *knowledgeType*, *tendency* and *certainLevel*, mainly adopted from SAO ontology. The Position attributions consist of *position, chapter,* and *chapterLabel*. They record the physical and logical position of the argumentation unit, which can be used for content positioning and reorganization.

3.3.2 Relation attributions

In this study, to accurately express the argumentation relationships of scientific papers, six relation attributions and their values were designed, including id, name, type, createDate, startNode and endNode. Their details are listed in Table 5.

4.0 Tools for the construction and visualization of argumentation graph

4.1 Annotation platform

To better implement the semantic annotation of the content of scientific papers, the authors of this study designed and developed the *Semantic Annotation Platform for Scientific Papers* (SAPSP). This platform contains three modules: system management, ontology management, and article management. In this study, we use SAPSP as an annotation tool to realize the argumentation units and their attributions in the argumentation graph. The annotation interface is shown in Figure 3.

Category	Attribution	Definition	Type	Value	REF
Basic attribution	id	Unique identifier of argumentation units	Long	Entity ID automatically assigned by database	
	name	The name of classes	String	18 class names	
	content	Text content of argumentation units	String	Argumentation unit Content	DCT
	author	Author of argumentation units	String	Author list split by ';'	
	date	Publication date of argumentation unit	Long	UNXI timestamp	DCT
	createDate	Creation date of argumentation unit	Long	UNXI timestamp	DCT
	identifier	Digital resource address of argumentation units	String	DOI	DCT
	language	The language of argumentation units	Sting	EN, CH, etc.	DCT
	format	The type of non-textual argumentation units	String	Figure, Table, Video, Audio, etc.	SAO
	formatSource	Digital resource address of non-textual units	String	URL	
Article attribution	articleId	Unique identifier of article	Long	DOI	
	articleTitle	Title of article	String	Title	DCT
	subject	Subject field of article	String	Biomedical, LIS, etc.	DCT
	type	Types of writing	String	Review, qualitative, quantitative, etc.	DCT
	source	the address of the digital resource of the article	String	DOI	DCT
Context attribution	knowledgeType	Knowledge type of argumentation units	String	Investigation, Observation, Analysis, General, Other	SAO
	tendency	Emotional attitude towards units	String	Positive, Neutral, Negative	SAO
	certainLevel	Degree of certainty	String	Low, Mid, High	SAO
Position attribution	position	Physical position of units within article	Int	Start position of argumentation unit	
	chapter	Logical position of units within article	String	Chapter number	Doco
	chapterLabel	The chapter title of argumentation units	String	Chapter title	Doco

Table 4. Class attributions of argumentation graph.

Attribution	Definition	Type	Value
id	Unique identifier ID of the relation instance	Long	Automatically assigned relationship ID
name	Name of relation	String	28 relation names
type	Type of relation	String	Article (intra-article relation) Citation (inter-article relation)
createDate	Creation date of argumentation relation	Long	UNIX timestamp
startNode	Id of start node	Long	id
endNode	Id of end node	Long	id

Table 5. Relation attributions of argumentation graph.

4.2 Visualization systems for argumentation graph

The node-link graph is the most widely used method for visualizing the knowledge graph. It represents the knowledge graph as interconnected nodes, points, or circles to represent nodes and directed edges connecting nodes to represent the relationship between them. The node-link graph not only provides a macro-overview of the network structure, but it also represents the microstructure of the knowledge graph effectively.

In this study, based on the d3.js front-end visualization framework, the Visualization System for Argumentation Graphs (VSAG) was developed. Using different node colors to distinguish class types and different arrow colors to distinguish relationship types, the argumentation process, and structure of a scientific paper can be clearly expressed. Moreover, the display of the argumentation graph can be zoomed in and dragged, so the microstructure can also be viewed. This system supports advanced retrieval in different dimensions, such as paper, author, discipline, or entity class. The argumentation process can also be retrieved while the node and relationship information is viewed. Figure 4 provides an example of the argumentation graph display of a single scientific paper.

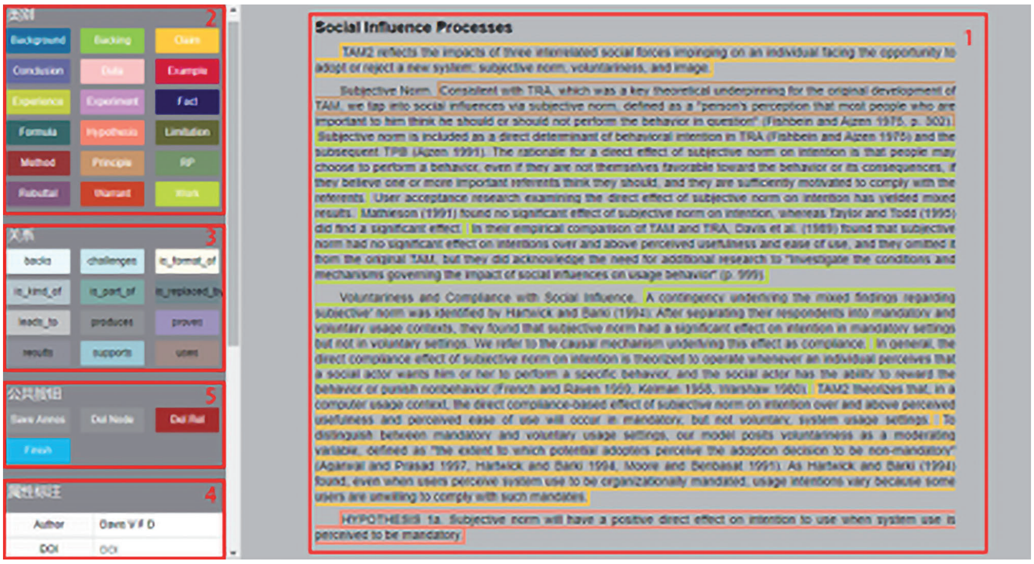

Figure 3. Annotation interface of the SAPSP. 1) article display area; 2) class selection area; 3) relationship selection area; 4) attribution tagging area.

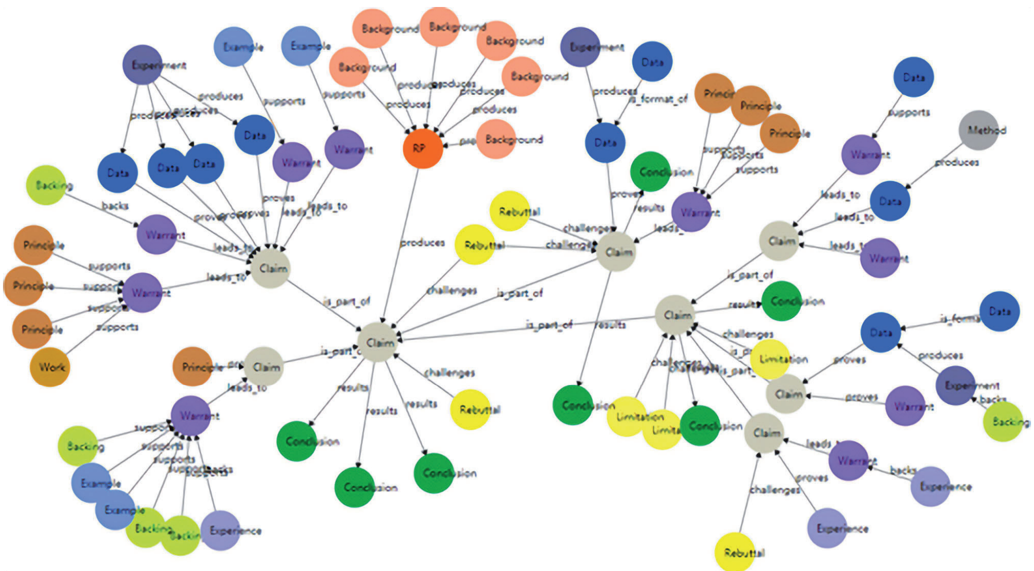

Figure 4. The example of the visualization of augmentation graph.

https://doi.org/10.5771/0943-7444-2022-4-213

5.0 Construction experiment of argumentation graph for scientific papers

5.1 Sample collection

To better represent and reflect the characteristics of the argumentation graph, this study chose the technology acceptance model (TAM) and the theory of planned behavior (TPB) as research topics. Scientific papers with these two topics were selected.

In this study, 12 scientific papers with a time span from 1989 to 2012 were selected to create an annotated sample, and SAPSP was used to annotate the text, create metadata of these 12 scientific papers, and generate HTML formatted documents as the corpus for construction experiments. Details of the papers are listed in Table 6.

5.2 Construction approach

Based on the argumentation graph data model, the construction approach follows these steps: 1) semantic annotation of scientific papers, 2) argumentation graph storage using Neo4J, and 3) graph visualization.

5.2.1 Semantic annotation

The semantic annotation of scientific papers is based on the data model of the argumentation graph. In this study, Protégé was used to formally represent the data model. The ontology file was imported into the annotation system to guide and constrain semantic annotation.

The semantic annotation process follows these steps: 1) argumentation units' annotation, 2) intra-article relation-

Number	Publishing year	Author	Title	Citation
A1	1989	Fred D. Davis	Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology	50030
A2	1991	Icek Ajzen	The Theory of Planned Behavior	76440
A3	1995	Shirley Taylor; Peter Todd	Understanding Information Technology Usage: A Test of Competing Models	9544
A4	1995	Shirley Taylor; Peter Todd	Assessing IT Usage: The Role of Prior Experience	3444
A5	2000	Viswanath Venkatesh; Fred D. Davis	A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies	18526
A6	2003	Viswanath Venkatesh; Michael G. Morris; Gordon B. Davis; Fred D. Davis	User Acceptance of Information Technology: Toward a Unified View	27582
A7	2006	Paul A. Pavlou; Mendel Fygenson	Understanding and Predicting Electronic Commerce Adoption: An Extension of the Theory of Planned Behavior	2659
A8	2006	Mun Y. Yi; Joyce D. Jackson; Jae S. Park; Janice C. Probst	Understanding information technology acceptance by individual professionals: Toward an integrative view	1172
A9	2007	Chun-Der Chen; Yi-Wen Fan; Cheng-Kiang Farn	Predicting electronic toll collection service adoption: An integration of the technology acceptance model and the theory of planned behavior	234
A10	2008	Viswanath Venkatesh; Hillol Bala	Technology Acceptance Model 3 and a Research Agenda on Interventions	4810
A11	2009	Ming-Chi Lee	Factors influencing the adoption of internet banking: An integration of TAM and TPB with perceived risk and perceived benefit	1595
A12	2012	Viswanath Venkatesh; James Y. L. Thong; Xin Xu	Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology	5412

Table 6. Samples for argumentation graph construction experiments.

ship annotation, 3) inter-article relationship annotation, 4) attribution annotation, 5) annotation verification. Using annotation and visualization systems, the scientific paper argumentation graph was constructed using the process shown in Figure 5.

The argumentation units, intra-article argumentation relations and attributions were annotated by SAPSP annotation system. To annotate the inter-article argumentation relationships, the citations between papers were recorded in an Excel table in the format of *starting argument unit ID – ending argumentation unit ID – inter text relationship name*.

5.2.2 Argumentation graph storage

A graph database, which is based on the graph model, can accurately describe entities, relationships, and attributes. It is one common way to store and retrieve large-scale knowledge graphs. Compared with Resource Description Framework (RDF) storage mode, graph databases are becoming more popular and are the fastest growing type of database.

The SAPSP annotation system employs the Neo4j database. The database data can be operated through the application programming interface, which directly stores argumentation units, argumentation relationships, and their attributions. In this research, a Python script was used to automatically read the Excel file of t he annotation results of inter-article argumentation relations. The node and relationship insertion command lines executed by the annotation system are shown in Figures 6 and 7, respectively.

5.3 Construction results

In the experiment, an argumentation graph with 1,262 nodes was constructed, and Figure 8 shows the distribution of the classes in the resulting argumentation graph. *Work*, *data*, *claim*, *warrant*, and *conclusion* are the most important nodes within the graph and make up the highest proportion of nodes.

A total of 1,818 argumentation relations, including 1,628 intra-article relations and 190 inter-article relations, were also annotated in this experiment. Figure 9 shows the distribution of intra-article relations and inter-article relations. The top three intra-article argumentation relations are *proved* (22.97%), *supports* (18.06%), and *is part of* (17.01%), and the top three inter-article relations are *work for* (33.86%), *provide principles for* (14.21%), and *provide assertion for* (12.63%).

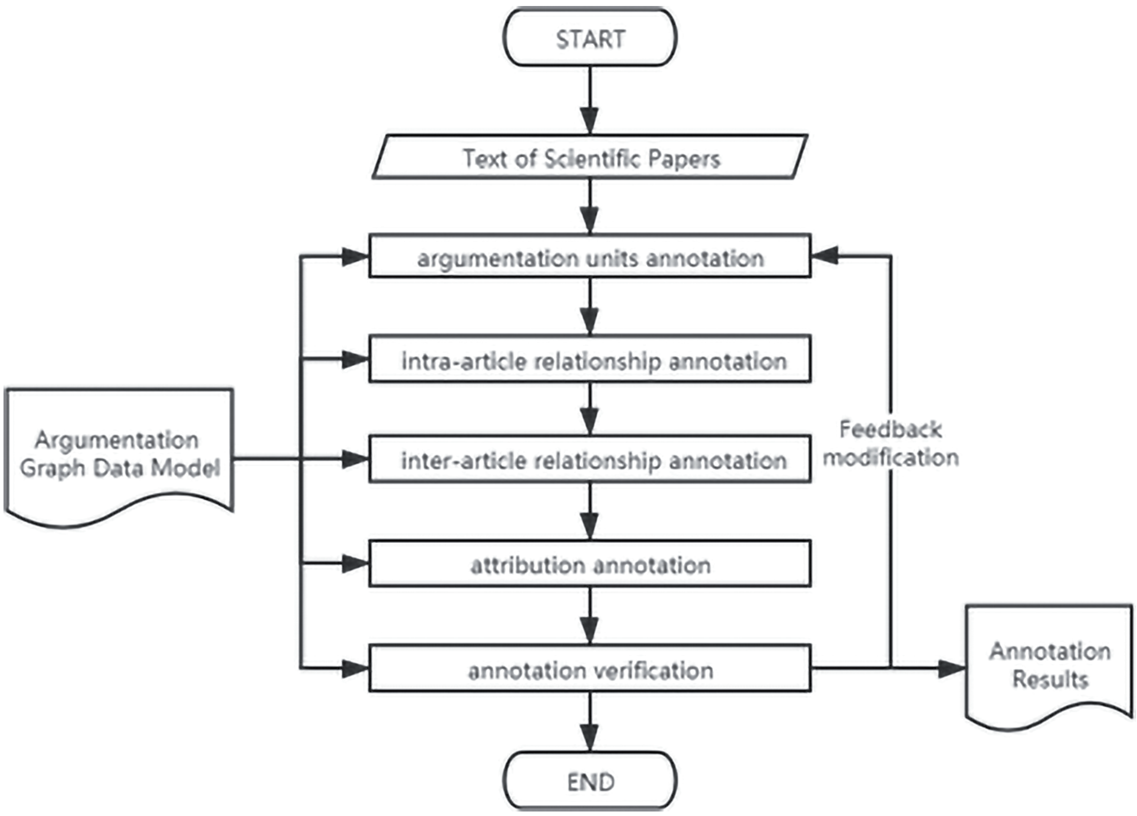

Figure 5. The process of scientific papers' semantic annotation.

```
CREATE(n:Fragment{name:'Hypothesis',  content:  'HYPOTHESIS  1a.
Subjective norm will have a positive direct effect on intention to
use when system use is perceived to be mandatory. ', author:' Davis
V F D', date: 950371200, createDate: 1577289600, identifier:' https://
dh.whu.edu.cn/argumentation/ 1581766261095826', language:'English',
format:'text', formatSource:'', articleId:42, articleTitle:'A Theoretical
Extension of the Technology Acceptance Model: Four Longitudinal
Field Studies', subject:'情报学', type:'研究类', source:'https://
www.jstor.org/stable/2634758', knowledgeType:'Investigation',
tendency:'Neutral', certainLevel:'Mid', position:7404, chapter:'2.1',
chapterLabel:' Social Influence Processes '}) return n
```

Figure 6. Node insertions commend.

```
Match (n:Fragment),(m:Fragment) WHERE id(n)=139 AND id(m)=184
CREATE (m)-[r:relationship{name: ' produces', type:'Neutral',
createDate:1577289600, startNode: 139, endNode: 184}]->(n) RETURN r
```

Figure 7. Relation insertions commend.

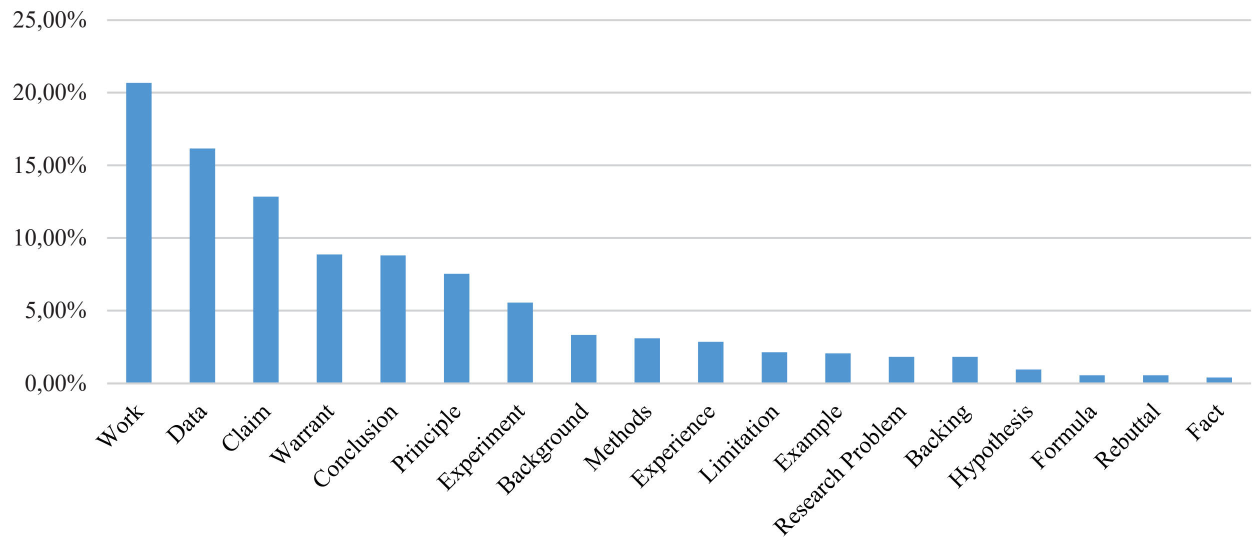

Figure 8. Class distribution.

Using the VSAG visualization system, the argumentation graph from the TAM and TPB research area was visualized and is shown in Figure 10.

6.0 Use cases of argumentation graphs

This study explores the use of the argumentation graph to realize the functions of automatic abstracting, strategic reading, and knowledge evolution representation. The aim is to explore the potential application scenarios of argumentation graphs preliminarily.

6.1 Representation of knowledge evolution

The argumentation graph of scientific papers provides a method for organizing and representing the knowledge construction process of scientific papers from the perspective of argumentation. Describing and analyzing the topological structure of this graph can help to sort out the evolution path of knowledge. This would allow users quickly understand the knowledge construction process in the current field.

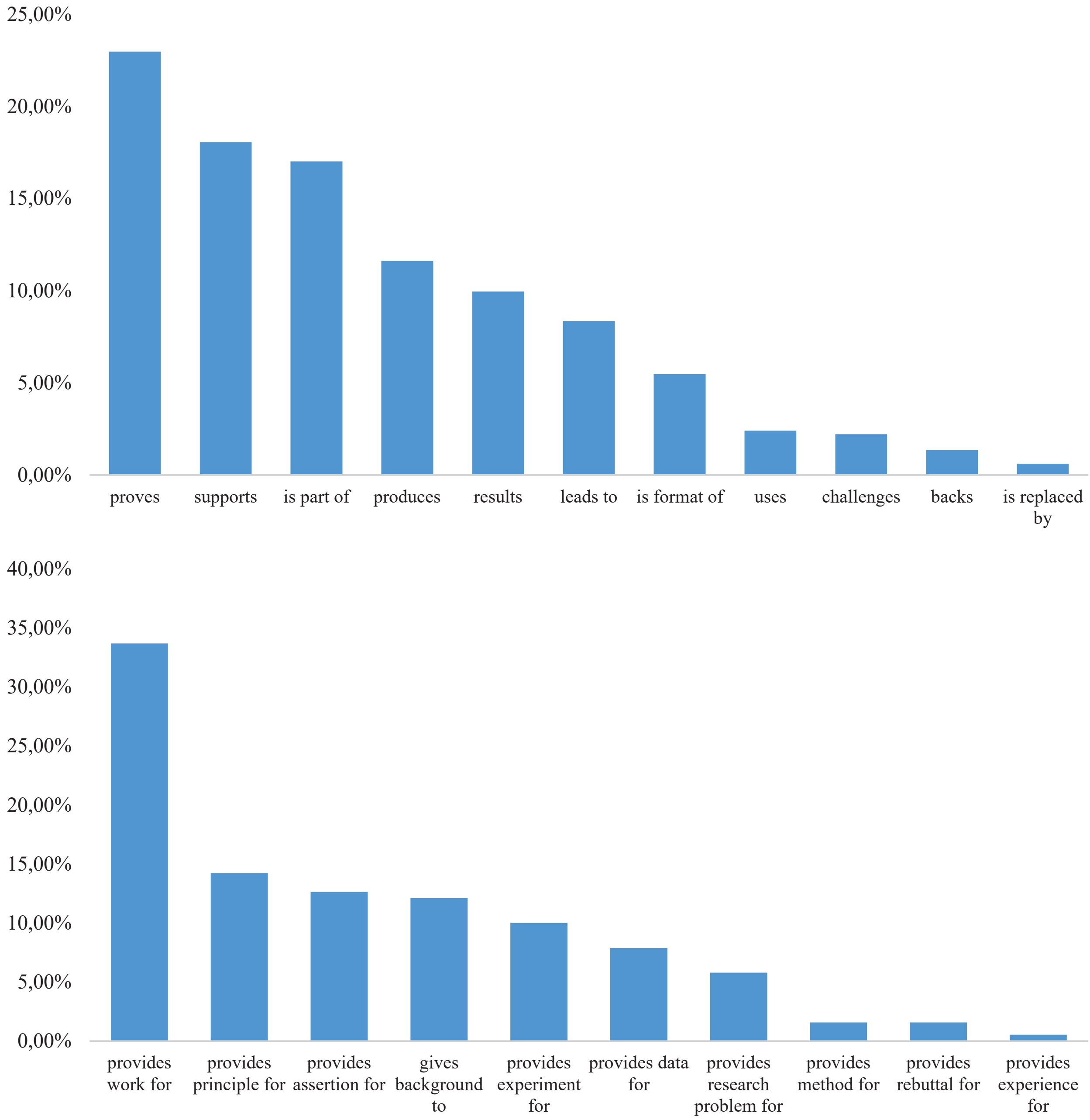

Figure 9. (a) Distribution of intra-article relations (b) Distribution of inter-article relations.

6.1.1 Evolution of single argumentation unit

The evolution of argumentation units objectively describes the argumentation process and how scientific thoughts are constructed. Figure 10 provides an example of this evolution. As shown in Figure 11, the *claim* of TAM is the initial argumentation unit, and it can be used to support the *claim* of TAM3 as the *experience,* while it also "*provides the assertion for*" the *claim* of TAM2. The theory of TAM2 has been used as the *background* and *work* in TAM3, and "*provides the assertion for*" the *claim* from TAM3. Thus, the evolution of the argumentation units from TAM to TAM2 and then to TAM3 can be represented and visualized using the argumentation graph.

6.1.2 Evolution among discourse

Applying the argumentation graph to the knowledge evolution among scientific papers can objectively describe the inter-article argumentation relationship between different sci-

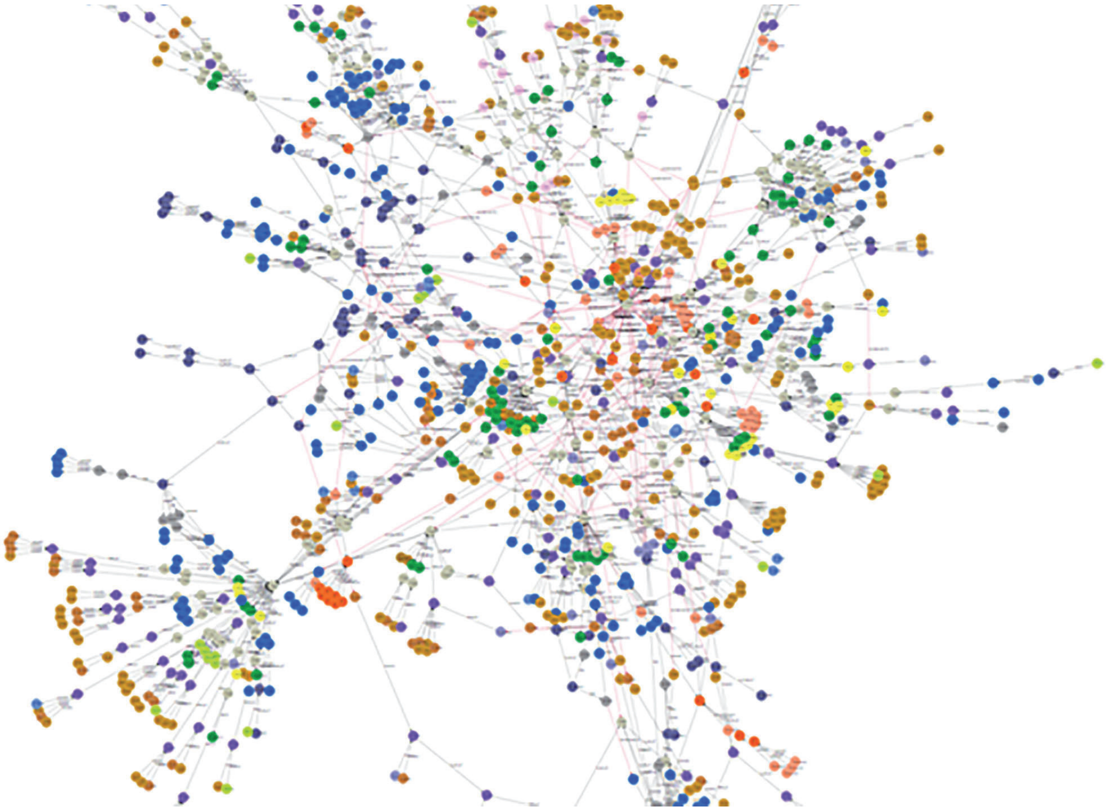

Figure 10. The argumentation graph.

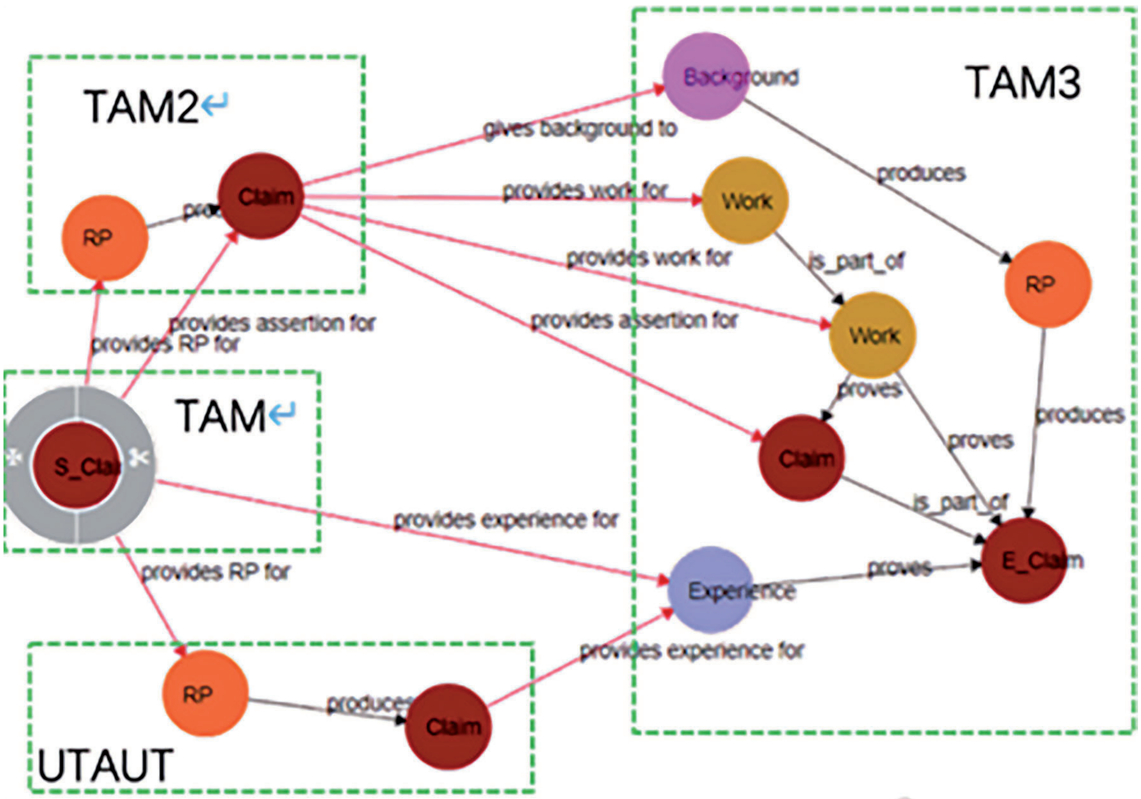

Figure 11. Example of argumentation units' evolution.

entific papers. As shown in Figure 12, the paper titled "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology" proposed the TAM theory (Davis 1989). These claims provide research *background* and *research problems* for TAM2 theory, reflecting the most common purpose of citations in citation behavior: to agree with others' opinions. In addition, the *data* generated in the TAM are used as the experimental scale of TAM2 theory, and its conclusion provides data support for TAM2 to argue TAM2's claim. This shows that researchers try to use additional argumentation units to support their own claims.

6.1.3 Evolution of claims

The evolution of claims represents the knowledge constructions within a specific field. Based on the argumentation graph, the system selects the claim type nodes and relations to generate a claim graph and subsequently uses an incremental color change to represent the evolution of a claim (Figure 13). A darker color indicates an earlier claim, and the overall evolution could be observed through the color change. Figure 12 shows the dynamic evolution of "TAM-TAM2-TAM3," so the evolution path of one specific theory can be revealed. The

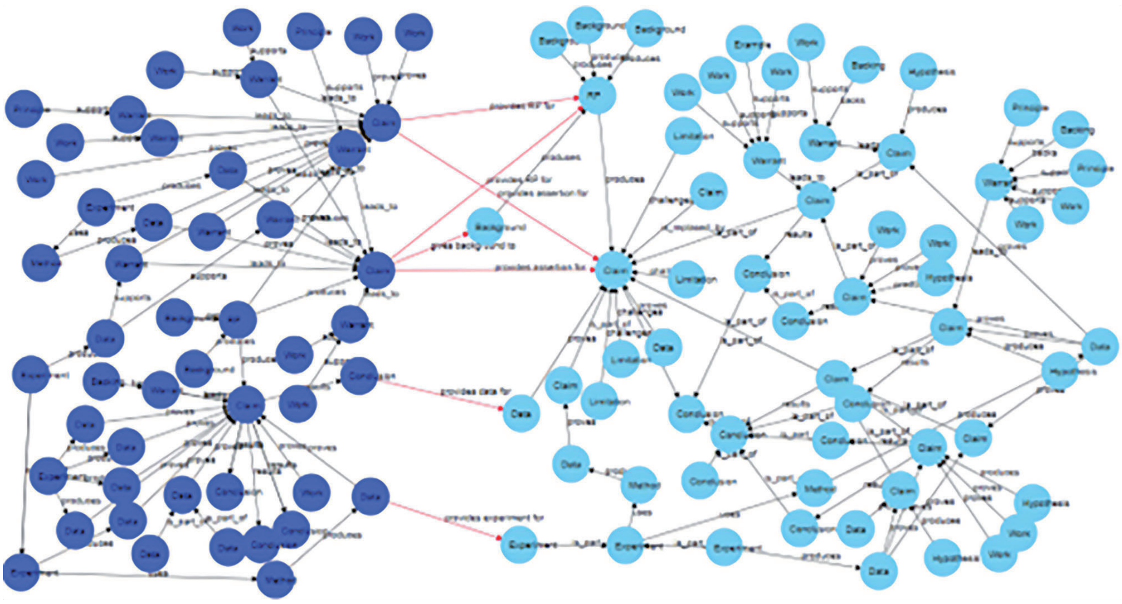

Figure12. Example of knowledge evolution between discourse

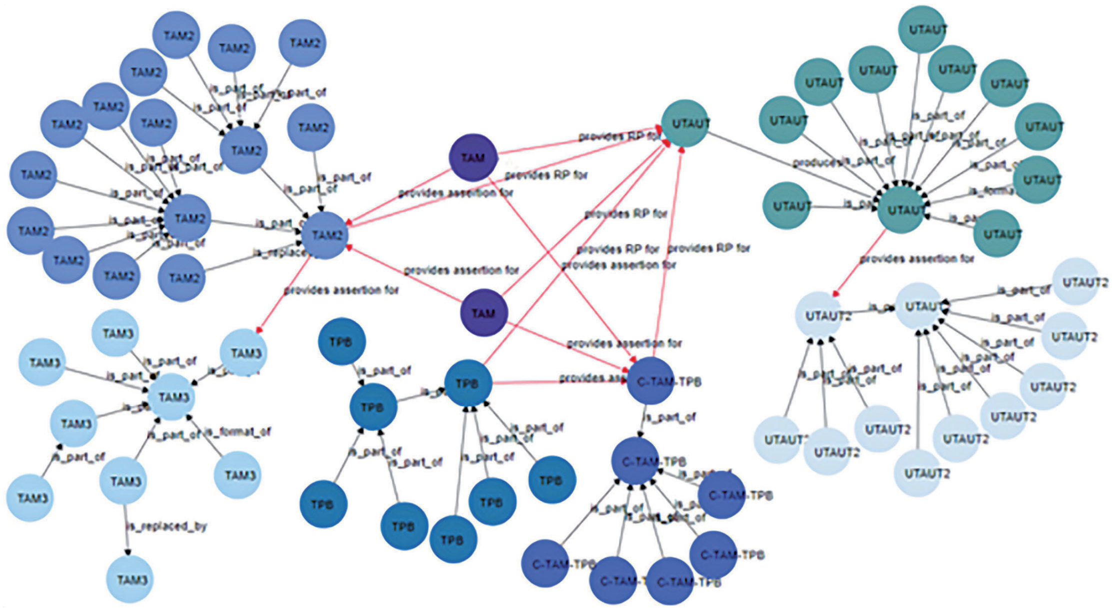

Figure13. Example of claim evolution.

process of integration and development of different theories is also represented. As shown in Figure 13, Shirley Taylor and Peter Todd developed the C-TAM-TPB theory by integrating TAM and TPB (Taylor and Todd 1995). Venkatesh *et al.* further developed the UTAUT theory by integrating TAM, TPB, C-TAM-TPB, TAM2, and other theories (Venkatesh, Thong and Xu 2021).

6.2 Automatic abstracting

The document summary around the argument can help scientific research users quickly understand the relevant supporting materials around the current argumentation. In the document collection of a specific field, the summary of multiple scientific papers can help users quickly master different discussions on the same scientific problem or claim.

In this study, the argumentation graph of scientific papers is applied to the automatic abstracting of scientific papers. Based on the results of argumentation structure mining (Song et al. 2019), the linear organization rules of automatic abstracting are constructed, shown in Figure 14.

As the core of scientific paper argumentation, the claim is the concentrated embodiment of scientific paper innovation, and it is also the core of scientific researchers' retrieval and reading. Therefore, the VSAG *visualization system* uses the claim as the starting point. The related argumentation units were selected and organized into abstract text by referring to the attributions of *text position* and *chapter number* of each argumentation unit. An example of using an argumentation graph to abstract text is given in this study, as shown in Figure 15.

6.3 Strategic reading of scientific papers

Strategic reading (Renear and Palmer 2009) is a recent approach to reading and understanding literature. The realization of knowledge evolution representation and automatic abstracting makes it possible to support users' strategic reading with the argumentation graph.

Figure 16 provides an example of how the visualization of the argumentation graph can be used for strategic reading. The system supports researchers, who input the *ID* and *Title* information of the article by visualizing the argumentation graph of that article and forming a hierarchical list of claims, which can be used as the basis of navigation or retrieval of scientific paper content. Researchers can select the *claim* content that needs to be read in detail and click the "claim" navigation button. The argumentation graph around this claim is displayed on the left, and the abstract text is displayed on the right.

In addition, the VSAG visualization system has a navigation function, and the corresponding text content on the right is highlighted when a node of the graph is selected. Researchers can choose the argumentation units that need to be read so that skimming and skipping can be used to read claim abstracts, improving the efficiency of knowledge reading and understanding.

7.0 Conclusion and future work

This paper proposed a method to construct an argumentation graph of scientific papers and a cross-discourse argumentation graph data model and two systems (SAPSP and VSAG). A construction experiment was also carried out in this study using the full text of 12 articles. The use cases demonstrate how the integration and evolution of theories can be represented using the argumentation graph.

Compared with the existing knowledge graphs in the academic field, our argumentation graph could better represent the structure and content of scientific papers and reveal the construction process of domain knowledge, which could be beneficial to assisting users' strategic reading. On

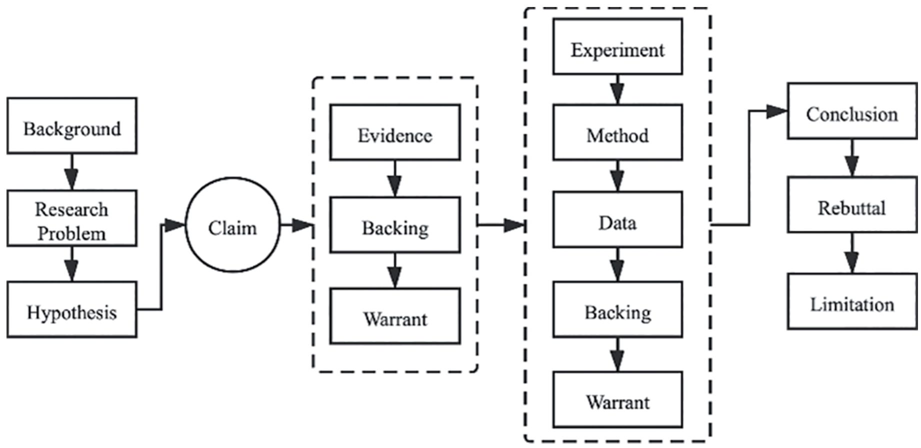

Figure14. Linear organization rules for automatic abstracting.

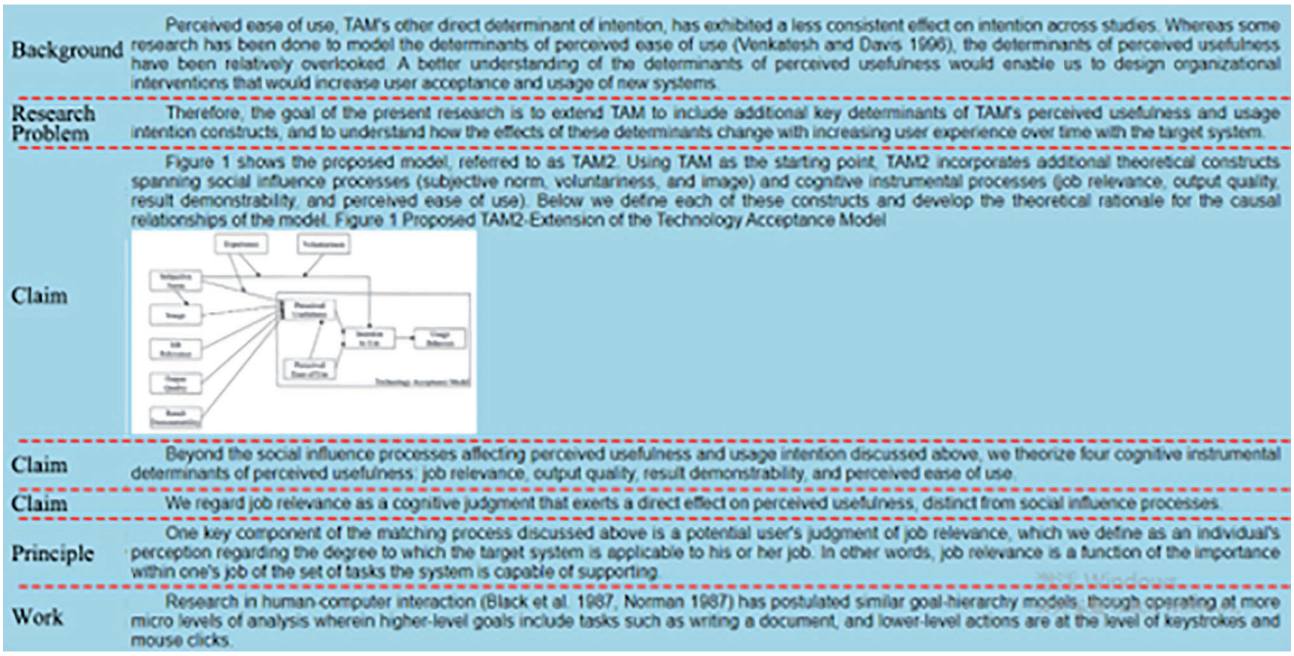

Figure 15. Examples of automatic abstracting based on argumentation graph.

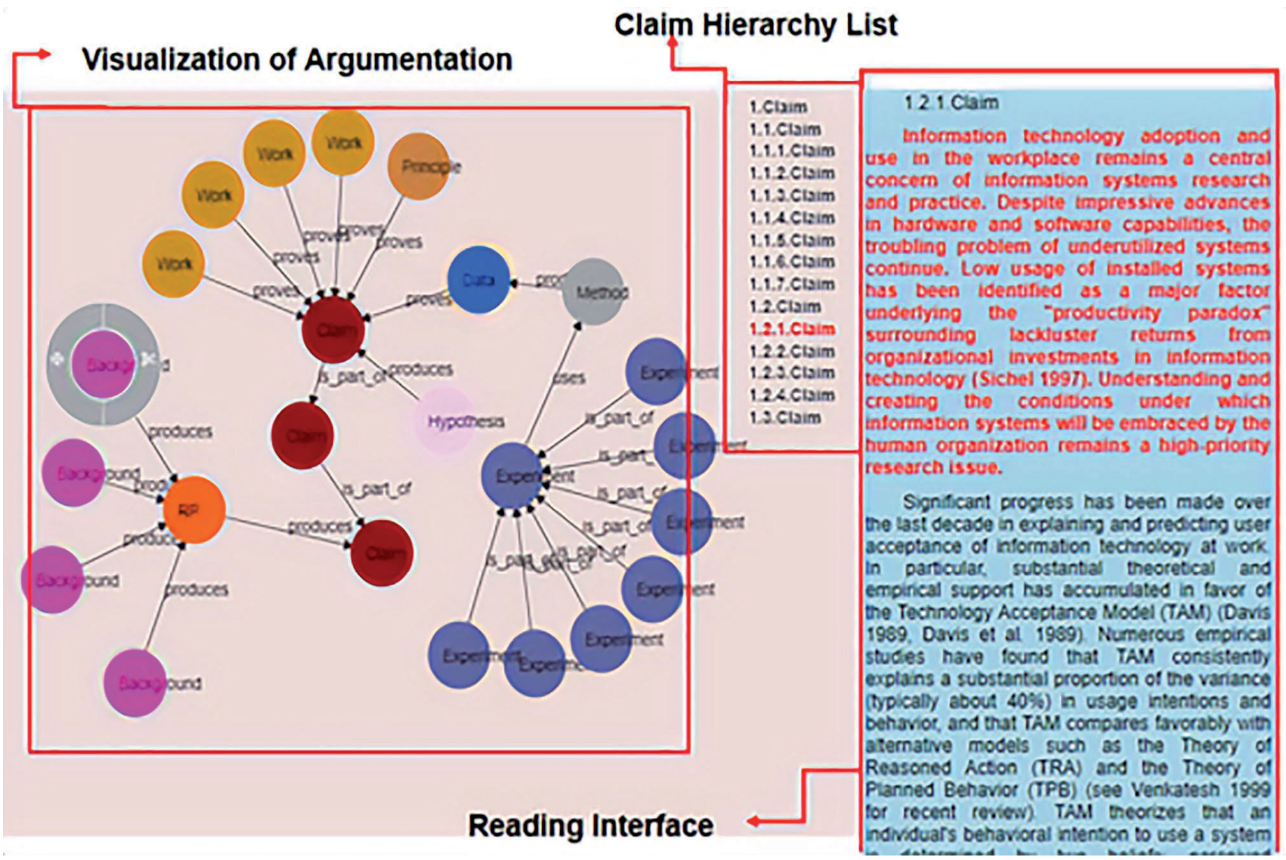

Figure 16. Example of strategic reading based on argumentation graph.

the one hand, with the representation and visualization of argumentation structure, through automatic summarization, users can quickly sort out the content and knowledge related to the current argument for strategic close reading or scanning. On the other hand, the argumentation graphs show the potential in representing the construction of domain knowledge, which can play a role in clue presentation to support users to explore the knowledge from a specific field.

However, some critical issues should be noted. Although an argumentation graph with 1,262 nodes and 1,818 relations was constructed, there were some deficiencies in the application scenarios in big data environments. It is difficult to use machine learning technology to realize the automatic extraction of argumentation units and their relationships. Considering the complexity and high cost of the graph construction process, it is essential to find a more efficient construction method.

The experiment based on 12 scientific papers of information science verified that the organization model has good usability in information science. Because of the differences in the argumentation process of scientific papers in different research fields, evaluating the model in other areas is an issue to address in the future.

This study explored the application of the argumentation graph to facilitate users' strategic reading, but its performance in these applications has not yet been evaluated. It will be an essential research direction to analyze the performance of the argumentation graph from the user's perspective.

Notes

1. https://www.springernature.com/cn/researchers/scigraph
2. https://www.microsoft.com/en-us/research/project/microsoft-academic-graph/
3. https://www.aminer.cn/
4. https://www.openacademic.ai/oag/

References

Auer, S., V. Kovtun, M. Prinz, A. Kasprzik, M. Stocker and M. E. Vidal. 2018. "Towards a Knowledge Graph for Science." In *Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics*, edited by Akerkar Rajendra et al. New York: ACM, 1–6. https://doi.org/10.1145/3227609.3227689

Bagga, A. and B. Baldwin. 1998. "Entity-Based Cross-Document Co-Referencing Using the Vector Space Model." In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*. Stroudsberg, PA: ACL, 79–85. https://doi.org/10.3115/980845.980859

Bagnacani A., P. Ciancarini, A. Di Iorio, A. G. Nuzzolese, S. Peroni and F. Vitali. 2015. "The Semantic Lancet Project: A Linked Open Dataset for Scholarly Publishing." In *Proceedings of Knowledge Engineering and Knowledge Management* (EKAW 2014). Lecture Notes in Computer Science Vol. 8982. Cham: Springer. https://doi.org/10.1007/978-3-319-17966-7_10.

Bench-Capon, T. J. M. and P. E. Dunne. 2007. "Argumentation in Artificial Intelligence." *Artificial Intelligence* 171: 619-41. https://doi: 10.1016/j.artint.2007.05.001

Bishop, A. P. 1999. "Document Structure and Digital Libraries: How Researchers Mobilize Information in Journal Articles." *Information Processing and Management,* 35: 255-79. https://doi: 10.1016/S0306-4573(98)00061-2

Clark, T., P. N. Ciccarese and C. A. Goble. 2014. "Micropublications: A Semantic Model for Claims, Evidence, Arguments and Annotations in Biomedical Communications." *Journal of Biomedical Semantics* 5: 28. https://doi: 10.1186/2041-1480-5-28

Cowie, J. and W. Lehnert. 1996. "Information Extraction." *Communications of the ACM* 39: 80-91. https://doi: 10.1145/234173.234209

Davis, F. D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly* 13: 319-40.

Dung, P. M. 1995. "On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and N-Person Games." *Artificial Intelligence* 77: 321-57. https://doi: 10.1016/0004-3702(94)00041-X

Freeman, J. B. 2001. "Argument Structure and Disciplinary Perspective. *Argumentation* 15: 397-423. https://doi: 10.1023/A:1012022330148

Green, N. L. 2015. "Identifying Argumentation Schemes in Genetics Research Articles." In *Proceedings of the 2nd Workshop on Argumentation Mining*. Denver, CO: Association for Computational Linguistics, 12-21. https://doi.org/10.3115/v1/w15-0502

Green, N. L. 2010. "Representation of Argumentation in Text with Rhetorical Structure Theory." *Argumentation* 24:181-96. https://doi.org/10.1007/s10503-009-9169-4

Groth, P., A. Gibson and J. Velterop. 2010. "The Anatomy of a Nanopublication." *Information Services and Use* 30: 51-6. https://doi: 10.3233/ISU-2010-0613

Han, X. and J. Zhao. 2009. "Named Entity Disambiguation by Leveraging Wikipedia Semantic Knowledge." In *Proceedings of the 18th ACM Conference on Information and Knowledge Management*. New York: ACM, 215-24. https://doi: 10.1145/1645953.1645983

Harrell, M. 2005. "Using Argument Diagramming Software in the Classroom." *Teaching Philosophy* 28: 163-77. https://doi:10.5840/teachphil200528222

Kelley, D. 2013. *The Art of Reasoning: An Introduction to Logic and Critical Thinking*. New York: W. W. Norton.

Kirschner, C., J. Eckle-Kohler and I. Gurevych. 2015. "Linking the Thoughts: Analysis of Argumentation Structures in Scientific Publications." In *Proceedings of the 2nd Workshop on Argumentation Mining*. Denver, CO. Association for Computational Linguistics, 1-11 https:/doi: 10.3115/v1/w15-0501

Kirschner, P. A., S. J. Buckingham-Shum and C. S. Carr (Eds.). 2012. *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making*. London: Springer Science & Business Media. https:/doi: 10.1007/978-1-4471-0037-9

Klein, M. 2007. "Achieving Collective Intelligence Via Large-Scale On-Line Argumentation." In *Proceedings of the 2nd International Conference on Internet and Web Applications and Services (ICIW'07).* IEEE, 58. https://doi:10.1109/ICIW.2007.13

Kunz, W. and H. W. J. Rittel. 1970. *Issues as Elements of Information Systems*. Berkeley, CA: Institute of Urban and Regional Development, University of California.

Lange, C., U. Bojars, T. Groza, J. G. Breslin and S. Handschuh. 200). "Expressing Argumentative Discussions In Social Media Sites." In *Proceedings of CEUR Workshop. Social Data on the Web (SDoW2008) at the 7th International Semantic Web Conference.* CEUR-WS, 31-42.

Liakata, M., S. Saha, S. Dobnik, C. Batchelor and D. Rebholz-schuhmann. 2012. "Automatic Recognition of Conceptualization Zones in Scientific Articles and Two Life Science Applications." *Bioinformatics* 28: 991-1000. https://doi.org/10.1093/bioinformatics/bts071

Lippi, M. and P. Torroni. 2016. "Argumentation Mining: State of the Art and Emerging Trends." *ACM Transactions on Internet Technology* 16(2): Article 10, 1-25. https://doi.org/10.1145/2850417

Liu, Z. 2005. "Reading Behavior in the Digital Environment: Changes in Reading Behavior Over The Past Ten Years." *Journal of Documentation* 61: 700-712.

Mochales, R. and M. F. Moens. 2011. "Argumentation Mining." *Artificial Intelligence & Law*, 19:1-22. https://doi: 10.1007/s10506-010-9104-x

Moser, E. and R. E. Mercer. 2020. "Use of Claim Graphing and Argumentation Schemes in Biomedical Literature: A Manual Approach to Analysis." In *Proceedings of the 7th Workshop on Argument Mining*. Association for Computational Linguistics, 88–99. https://www.aclweb.org/anthology/2020.argmining-1.10

Okada, A., S. J. B. Shum and T. Sherborne (Eds.). 2008. *Knowledge Cartography: Software Tools and Mapping Techniques.* New York: Springer. doi:10.1007/978-1-84800-149-7

Pedersen, T., A. Purandare and A. Kulkarni. 2005. "Name Discrimination by Clustering Similar Contexts." In *CICLing. International Conference on Computational Linguistics & Intelligent Text Processing*, edited by Alexander Gelbukh. Lecture Notes in Computer Science, vol 3406. Berlin; Heidelberg: Springer, 226-37. https:/doi: 10.1007/978-3-540-30586-6_24

Peldszus, A. and M. Stede. 2013. "From Argument Diagrams to Argumentation Mining in Texts: A Survey." *International Journal of Cognitive Informatics & Natural Intelligence 7*: 1-31.

Rahwan, I. and G. R. Simari (Eds.). 2009. *Argumentation in Artificial Intelligence.* New York: Springer. https://doi: 10.1007/978-0-387-98197-0.

Reed, C., D. Walton, D. and F. Macagno. 2007. "Argument Diagramming in Logic, Law and Artificial Intelligence. *Knowledge Engineering Review* 22: 87-109. doi:10.1017/S0269888907001051

Renear, A. H. and C. L. Palmer. 2009. "Strategic Reading, Ontologies, and the Future of Scientific Publishing." *Science* 325(5942): 828-32. https:/doi: 10.1126/science.1157784

Sadeghi, A., C. Lange, M. E. Vidal and S. Auer. 2017. "Integration of scholarly communication metadata using knowledge graphs." In *Research and Advanced Technology for Digital Libraries. TPDL 2017*, edited by J. Kamps, G. Tsakonas, Y. Manolopoulos. L. Iliadis, and I. Karydis. Lecture Notes in Computer Science vol 10450. Cham: Springer, 328-41. https:/doi: 10.1007/978-3-319-67008-9_26

Saruladha, K. 2011. "A Comparative Analysis of Ontology and Schema Matching Systems." *International Journal of Computer Applications* 34(8): 14-21.

Schneider, J., T. Groza and A. Passant. 2013. "A Review of Argumentation for the Social Semantic Web. *Semantic Web* 4: 159-218. https:/doi: 10.3233/SW-2012-0073

Sen, P. 2012. "Collective Context-Aware Topic Models for Entity Disambiguation." In *Proceedings of the 21st Annual Conference on World Wide Web*. https:/doi: 10.1145/2187836.2187935

Shum, S. B., E. Motta and J. Domingue. 2000. "ScholOnto: An Ontology-Based Digital Library Server for Research Documents and Discourse." *International Journal on Digital Libraries* 3: 237-48. https:/doi: 10.1007/s007990000034

Soldatova, L. N. and R. D. King. 2006. "An Ontology of Scientific Experiments." *Journal of the Royal Society Interface* 3: 795-803. https:/doi: 10.1098/rsif.2006.0134

Song, N., H. Cheng, H. Zhou and X. Wang. 2019. "Argument Structure Mining in Scientific Articles: A Comparative Analysis." In *Proceedings of 2019 ACM/IEEE Joint Conference on Digital Libraries (jcdl).* ieee, 339-40. doi: 10.1109/jcdl45914.2019

Taylor, S., and P. Todd. 1995. "Assessing IT Usage: The Role of Prior Experience." *MIS Quarterly* 19: 561-70.

Teufel, S. 1999. "Argumentative Zoning: Information Extraction from Scientific Text." *Unpublished PhD thesis University of Edinburgh*. https://doi: 10.1.1.105.485

Teufel, S. 2010. "The Structure of Scientific Articles: Applications to Citation Indexing and Summarization." *Computational Linguistics* 38: 443-5. https://doi.org/10.1162/COLI_r_00102

Toulmin, S., R. Rieke and A. Janik. 1984. *An Introduction to Reasoning*. 2nd edition. New York: Macmillan.

Toulmin, S. E. 2003. *The Uses of Argument*. Updated edition. Cambridge: University Press. https://doi.org/10.1017/CBO9780511840005

Turney, P. D. 2001. "Mining the Web for Synonyms: PMI-IR versus LSA on TOEFL." In *Machine Learning: ECML 2001*, edited by L. De Raedt and P. Flach. Lecture Notes in Computer Science vol 2167. Berlin; Heidelberg: Springer, 491-502. https://doi: 10.1007/3-540-44795-4_42

Tweed, C. 1994. "Intelligent Authoring and Information System for Regulatory Codes and Standards." *International Journal of Construction Technology* 2: 53-63.

Van Gelder, T. 2009. "What is Argument Mapping?" https://timvangelder.com/

Van Gelder, T. 2007. "The Rationale for Rationale." *Law, Probability and Risk* 6: 23-42. doi:10.1093/lpr/mgm032

Venkatesh, V., J. Y. L. Thong and X. Xu. 2012. "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology." *MIS Quarterly* 36: 157-78.

Walton, D. and N. Zhang. 2013. "The Epistemology of Scientific Evidence." *Artificial Intelligence & Law* 21: 173-219. https://doi: 10.1007/s10506-012-9132-9

Yan, J., C. Wang, W. Cheng, M. Gao and A. Zhou. 2018. "A Retrospective of Knowledge Graphs." *Frontiers of Computer Science* 12: 55–74. https://doi: 10.1007/s11704-016-5228-9

Zhou, H., N. Song, H. Cheng and X. Wang. 2019. "Argument Ontology for Describing Scientific Articles: A Statistical Study Based on Articles from Two Research Areas." In *Proceedings of the Association for Information Science and Technology* 56: 855-7. https://doi: 10.1002/pra2.204.

Appendix I Annotation results

Argumentation units	Nodes Num.	Proportion in total nodes	Sentences Num.	Proportion in total sentences
Work	261	20.68 %	530	9.98 %
Data	204	16.16 %	969	18.25 %
Claim	162	12.84 %	229	4.31 %
Warrant	112	8.87 %	171	3.22 %
Conclusion	111	8.80 %	234	4.41 %
Principle	95	7.53 %	185	3.48 %
Experiment	70	5.55 %	519	9.77 %
Background	42	3.33 %	222	4.18 %
Methods	39	3.09 %	118	2.22 %
Experience	36	2.85 %	107	2.01 %
Limitation	27	2.14 %	51	0.96 %

Argumentation units	Nodes Num.	Proportion in total nodes	Sentences Num.	Proportion in total sentences
Example	26	2.06 %	46	0.87 %
Research Problem	23	1.82 %	53	1.00 %
Backing	23	1.82 %	69	1.30 %
Hypothesis	12	0.95 %	22	0.41 %
Formula	7	0.55 %	11	0.21 %
Rebuttal	7	0.55 %	19	0.36 %
Fact	5	0.40 %	6	0.11 %
All	1262	100 %	5311	67.5 %

Table 1. The annotation results of classes.

Relationship Types	Argumentation Relationships	Numbers	Proportion of relationship
inner-article relationships (1628)	proves	374	22.97 %
	supports	294	18.06 %
	is part of	277	17.01 %
	produces	189	11.61 %
	results	162	9.95 %
	leads to	136	8.35 %
	is format of	89	5.47 %
	uses	39	2.40 %
	challenges	36	2.21 %
	backs	22	1.35 %
	is replaced by	10	0.61 %
intra-article relationships (190)	provides work for	64	33.68 %
	provides principle for	27	14.21 %
	provides assertion for	24	12.63 %
	gives background to	23	12.11 %
	provides experiment for	19	10.00 %
	provides data for	15	7.89 %

Relationship Types	Argumentation Relationships	Numbers	Proportion of relationship
	provides research problem for	11	5.79 %
	provides method for	3	1.58 %
	provides rebuttal for	3	1.58 %
	provides experience for	1	0.53 %
	provides conclusions for	0	0 %
	provides example for	0	0 %
	provides formula for	0	0 %
	provides backing for	0	0 %
	provides warrant for	0	0 %
	provides hypothesis for	0	0 %

Table 2. Annotation results of relations.