RESEARCH ON INFORMATION RESOURCES AGGREGATION IN ACADEMIC TO SEMANTIC PUBLISHING

Iza Fajri, Shanshan Zai & Siti Aminah

Abstract
With the constant development of information and digitization, the proportion of digitization in scientific research publications is increasing day by day. On the one hand, the rapid growth of digital scientific research data and academic literature has provided many facilities for academic exchanges among scientific research users. On the basis of systematically combing the relevant theories of semantic publishing and information resource integration, this paper summarizes the current situation of information resource aggregation in academic journals and the significance of digital resource aggregation. Secondly, this paper illustrates the important role of semantic information resource integration in semantic publishing of academic journals. Taking Elsevier semantic publishing model as an example, it focuses on the resource query and resource utilization under semantic publishing. Final adoption with the comparison of web of science database and the analysis and evaluation of the results of resource aggregation verify the feasibility of the semantic based digital resource aggregation method in the digital publication of academic journals.

Keywords: *Semantic Publishing; Semantic Web, Digital Resource,* and *Aggregation elsevier*

1. INTRODUCTION

The development of today's technology is increasingly rapid, almost all areas of technology is now internet-based. One of the technologies that we currently feel the benefits is the World Wide Web or known by the name web. Web now a source of data that is very and very valuable for each user because the web collection of documents connected to each other and can be accessed via internet connection. Recent developments, which take into account the current stage of technology and the web, have led to a novel definition for electronic publications, as semantic publications "We define the term semantic publication to include anything that enhances the meaning of a published journal article. Facilitates its automated discovery, enables its linking to semantically related articles, provides access to data within the article in actionable form, or facilitates integration of data between articles"[1]

Semantic Web (SW) technologies constitute a step forward semantic retrieval and processing in computational environments. The proposal content of a Web document is no longer a matter of keyword match as in conventional computational environments since the 1960s, but instead comprises structured sets of concepts connected by precise meaning relations as in RDF (Resource Description Framework) and RDF Schema statements. Such a rich knowledge representation schema enables software agents to perform "inferences" and more sophisticated tasks based on the document content. We have chosen RDF for modeling the citation life cycles because of its advantages with respect to other formalisms. RDS is modelar; a subject of RDF triples from an RDF graph can be used separately, keeping and consistent RDF model. RDF is the main building block of the semantic web initiative, together with a set of technologies for defining RDF vocabularies like RDF schema (RDFS) and the OWL. The objective of this paper is to

propose a richer semantic content publishing model, in which scientific claims made by authors throughout articles are expressed by relations between phenomena. In the proposed model, each article, in addition to being published in and exploit SW technologies to enhance scientific communication, management, sharing, and reuse of knowledge; others aim at providing direct access to semantic content of scientific articles. Thus, there is an increasing trend in electronic publishing experiences toward formalizing the text of articles or structuring them, marking them, and identifying significant parts to facilitate more direct reading by humans, potentially by relating with Academic Journal Oriented to Semantic Publishing.

Aggregation academic journal oriented semantic publishing not only using by RDF also, can including a formal model and XML serialization syntax, the basic building block model, OWL and an Ontology for the semantic publishing. OWL is a vocabulary for describing properties and classes of RDF resources, complementing RDFS's capabilities for providing semantics for generalization hierarchies of such properties and classes. OWL has the influence of more than the years of DL research. This knowledge allowed the set of constructors and axioms supported by OWL to be carefully chosen the as to balance the expressive requirements of typically applications with a requirements for reliable and efficient reasoning support. Concepts of semantic publishing is allowed the effective categorization and comparison of different approaches used in mapping RDB data to RDF.

2. LITERATURE OF REVIEW

2.1 Process of Semantic Publishing

Process of Semantic Publishing is concern two steps;

a) The alignment of the ontological instances created during semantic authoring to the (final and clean) textual content of the scientific publication, and
b) The embedding of the resulted semantic metadata into the final publication document, thus creating by a semantic publication. As we discuss, both or step are dependent on the final document format, and raise a series of interesting technical challenges, especially with regard to the creation of pointers for the textual counterparts of the SALT ontological instances. Associated with two above, mentioned environments, we chose as resulting document formats PDF and DOC. Consequently, we developed two particular solutions for creating pointers, as specializations and instances of the pointer concept from the document Ontology.

2.2 Advantage of Semantic Publishing

Therefore, the key to finding solutions to ease this problem of information overload is to take advantage of the support provided by electronic publications. Such advantages are two-fold:

a) It gives direct access to the textual content of the publication, thus allowing one to process is and, subject to the actual electronic document format.
b) It may offer the possibility of storing and linking enriched information about the publication, or the publication's content.

Furthermore, these two elements enable additional mechanisms, which when adapted and integrated, can lead to partial solutions to the information overload problem.

Making the discourse structures trapped within the publication's content explicit/ this leads to a direct crystallization and exchange of ideas from the authors to the readers, and consequently to knowledge creation, thus enabling the electronic publications to surpass their current use. However although made explicit, such knowledge (i.e…the implicit discourse structure) is dependent on the degree of formalization. In the case of the rhetorical and argumentation discourses, it can be a couple of keywords, or a weakly structured text, both possibly including direct references of other publications. In order to take full advantage of semantic publications, we need tailored Web platforms that enable richer and more meaning full retrieving and browsing capabilities. These will, in turn, show the benefit of using semantic publication and, consequently, increase the incentive for both the authors and the publishing companies for creating them.

2.3 Research on Information Resources Organization of Academic Journals

Benefit of Semantic Publishing A system that lets you dynamically interconnect content objects such as tags, concepts, terms and many others, is a game changer. By aggregating relevant and timely content automatically, such technology minimizes the cost of manual content management and makes the best use of content assets. Not only does it do all the heavy-lifting of organizing and categorizing information, but also opens countless opportunities for efficient content repurposing, reuse and relevant delivery. With dynamically generated pages readers are being served exactly the content they are interested in, when and where they need it.

The dynamic of Semantic Publishing platform can:

[1] Automatically create relationships between Things (people, places, events etc.)
[2] Automatically add tags, concepts and topics to content
[3] Dynamically aggregate topic and profile pages
[4] Create custom content streams
[5] Reuse and repurpose content, based on variables
[6] Provide tailored recommendations
[7] Give related content and data as the user types

Recent developments in Web technology can be used for semantic enhancement of scholarly journal articles, by aiding publication of data and metadata and providing 'lively' interactive access to content. Such semantic enhancements are already being undertaken by leading STM publishers, and automated text processing will help these enhancements become affordable and routine. Publisher, editor, and author all have primary roles in that process; an incremental approach is needed. Publication of data and metadata to the Web make possible added-value 'ecosystem services'; semantic publishing will bring substantial benefits to scholarly communication.

2.4 Research on Semantic Publishing

The Royal Scienty of Chemistry's prize winning project prospect relies heavily upon the domain expertise of its editors. Its editorial production system manager, Richard Kidd said to a great extent succes is due to outr technical editors who have developed ne skilss to judge the meaning and context of terms as they're exprienced highly qualified chemical scientiest, we feel that this is a great applications of their skills and knowledge.

Pragmatics as the youngest sub disciples of linguistic science grew and evolved from antisintaxist tendencies, social-critical tendencies, philosophical traditions, and

ethnomethodologically traditions. The first tendency to reject the view of Chomsky syntax, that in the study of the central language is syntactic, and that phonology, morphology, and semantic is a mere peripheral. But two Chomsky, George Lakoff and Robert Roos disciples disagree with their teacher, and they think syntax is not everything because in the practice of language usage, it is not uncommon for us to encounter language errors, neither syntax neutrality nor semantic inaccuracies, nevertheless communication goes on. Meanwhile, the tendency of ethnomethodology examines how members of a speech community organize and understand their speech activities in everyday life, not on the aspects of programmability, but how the wearer understands what they are saying even if the speech is grammatically not complete. Concerning the activity of Levinson [2] states that basically the presuppositions in taking a number of conclusions are not based on semantic factors in the narrow sense, but rather on highly sensitive contextual factors. Another linguist [3], suggests that the concept of presupposition only concerns an important semantic relationship called entailment.

The meaning of a speech can not only be determined by the lingual factors that make up the speech, but can also be determined by nonlingual factors. Determining the meaning of a speech based on lingual factors can be studied from the lingual forms that make up it. But the determination of the meaning of a speech based on the usual nonlingual factors varies greatly depending on the underlying speech situation. The first is the object of semantic study while the second is the object of pragmatic study. The definition indicates that there are pragmatic constraints in the use of a sentence such that such sentences are only reasonably used if assumed in the context that the propositions addressed by presupposition-triggers are true. So it can be said that telling a sentence that wrong presupposition will only produce speech that is not fair or appropriate. Unlike the semantic view that states the statement of a sentence can never be said to be true or false. Many scholars have expressed their opinion regarding the definition of this pragmatic presupposition which all of them use their respective perspectives. But according to Levinson [4] we do not really need to be interested in studying the definitions proposed by experts but must think about how to provide a model that can precisely predict presupposition behavior. Levinson's concept of fairness or appropriateness still needs further study.

2.5 Related *Concepts of Semantic Publishing*

The relation of meaning is the semantic relation which exists between one language unit and another. Language unit here can be a word phrase or sentence; and that semantic relation can express the same meaning. The conflicts of meaning, meaninglessness, meaninglessness or excess meaning. Relations of meaning are usually discussed issues called synonyms, antonyms, polysomic, homonymy, hyponymy, ambiguity, and redundancy. Basically the principle of meaning relation there are four types, namely:

a) The principle of contiguity. The principle of contiguity is the principle that explains that some words can have the same or similar meaning. This principle can lead to a meaning relation called synonymy.
b) The principle of collation. The principle of complementation is the principle that explains that the meaning of one word contrary to the meaning of another word. This principle can lead to a meaning relation called antonymy.
c) The overlaping principle. Theoverlaping principle is the principle that explains that

one word has a different meaning or words that sound the same but contain different meanings. This principle can lead to a meaning relation called homonym and polysomic.

d) The inclusion. The principle of inclusion is the principle that explains that the meaning of a word includes several meanings of other words. This principle can lead to a relation of meaning called hyponymy.

2.6 Characteristics of semantic Publishing

The characteristics of semantic are actually determined by identity, normative property, and behavior. Some behaviors are normative default behavior; this behavior can be overridden by higher level protocols. However, in the absence of such protocols, the behavior must be considered to follow the semantic character. Properties character combinations and canonical ordering behavior cannot be replaced by higher level protocols. The purpose of this constraint is to ensure that the sequence of combining marks in the text and the result of normalization is predicted. Each aspects of the meaning of a word is seen in a characteristic pattern of normality and semantic abnormality in a grammatically appropriate context. There were seven types of meaning according to Leech:

a) The conceptual meaning: logical, cognitive, or denotative content. Example : HOUSE
b) The connotative meaning: which is communicated to what language refers to.
c) The stylistic meaning: which is communicated from the social circumstances concerning the use of language.
d) The affective meaning: expressed from the feelings and behaviors of the speaker or author.
e) The meaning of reflection: conveyed through association with another meaning of the same expression.
f) The meaning of colocative: which is conveyed through association with the word that tends to occur in other word sphere.
g) The thematic significance: communicated in a way in which the message is arranged on the basis of sequence and pressure.

The semantic publishing is a publication that examines the meaning of analysis and is formal and reviewed according to the idea, whereas the digital publishing is one part of a cryptographic method that uses hash functions to mark a data. This signature is useful to know that the data sent actually comes from the sender and does not change. The semantic-based scientific semantic search system is a scientific search engine that uses metadata as a bibtex entry as a search criterion. In such systems, digital signatures are used to identify data sources and prevent duplication of data. Bibtex data, obtained from public data service providers such as CiteSeer, Google Scholar, etc. as well as from proprietary data, will be converted to metadata (in this case RDF) and also processed using hashing and encryption functions to create digital signatures. The results of this processing then put together into the form of RDF data which will be used in the repository metadata.

2.7 Evant methods for Semantic Publishing

a) RDF. RDF or Resource Description Framework on Semantic Web. RDF is the

standard model for data exchange on the web. According to Liang Yu in his book A Developer's Guide to the Semantic Web, RDF is a language for representing information about resources in the world; RDF is a framework for representing information contained on the web (RDF Concepts); RDF is the main goal of language to represent information contained within the web (RDF Syntaxt, RDF Schema). As the RDF notes have been explained earlier, that RDF can represent information about both resource resources in the world and resources that are on the web. RDF represents that information into a sentence, just as humans would describe something. RDF describes something in a sentence consisting of a subject, an object, and a predicate. The subject, object and predicate are the main componets of RDF. In RDF, the subject is a resource or something that is usually described as a URI address, whereas predicate is the property of the resource that connects the subject with the object. While the object is a URI or literal. These three components are usually commonly called triple. Thus, it can be said that the RDF contains the required information. To save the RDF is needed one is a triple store. In other words, a triple store is a storage place for RDF.

b) ***Semantic Web.*** Semantic web or semantic web is one of the developments in web applications. According to language, semantic web has meaning of web which has meaning. In other words, semantic web is a web application that has a certain knowledge base so that it can be said semantic web has more intelligent properties than the previous web. One example of semantic web is the web can recommend something to the user in accordance with their respective user interests. Thus, it could be that when multiple people access the same web address, the content or content of the web page will not be the same. One example of semantic web is igoogle. When we first open igoogle, we will be asked to fill the location where we are, after that we just go to igoogle home page. Thus, it may be that the content of the igoogle home page will vary according to the pre-loaded location, be it from news content, air temperature, and more. As explained above, Semantic Web or web semantics was first coined by Tim Berners-Lee in 2001. Semantic Web is often referred to as web version 3.0. In making semantic web apps is not as easy as creating a regular web app. To create a semantic web there are several technologies that need to be studied, such as RDF, ontology, RDF query, RDF Store. Actually Semantic Web consists of 2 words that each - each has a quite different understanding there are: 1) Web: the intended web here is a wide computer network that is WWW (World Wide Web); and 2) Semantics: can be interpreted as the science of meaning or meaning, the one of the three levels of language analysis; phonology, grammar and semantics.

c) **Metadata.** Metadata is structured information that describes, explains, discovers, or at least makes information easy to recover, use, or manage. Metadata is often referred to as data about data or information about information. This metadata contains information about the contents of a data used for the purpose of file / data management that later in a database. If the data is in text form, the metadata is usually a description of the name of the field, the length of the field, and the field type: integer, character, date, etc. For the type of image data (image), metadata contains information about who the photographer, when shooting, and

camera settings at the time of shooting. Another one for the type of data in the form of a collection of files, its metadata are file names, file types, and the name of the administrator (administrator) of those files. Metadata in principle has a function similar to the catalog in the library that is: Represents or reperesentations of a document or source of information Facilitator to make the information source easy to find by using relevant criteria, Identify the source, Grouping sources that have the similarities, Distguishes sources that have no resemblance and Provides information about source location. Metadata comes in many shapes and flavors, carrying additional information about where a resource was produced, by whom, when was the last time it was accessed, what is it about and many more details around it. Similar to the library cards, describing a book, metadata describes objects and adds more granularity to the way they are represented. Three main types of metadata exist: descriptive, structural and administrative. 1) Descriptive metadata adds information about who created a resource, and most importantly what the resource is about, what it includes. This is best applied using semantic annotation; 2) Structural metadata includes additional data about the way data elements are organized – their relationships and the structure they exist in; 3)Administrative data provides information about the origin of resources, their type and access rights.

Metadata underlies every digital object and is critical to the way they are managed, organized and used. When created and handled properly, metadata serves the clarity and consistency of information. Metadata facilitates the discovery of relevant information and the search and retrieval of resources. Tagged with metadata, any digital object can be automatically associated with other relevant elements and thus easy to organize and discover. This helps users make connections they would not have made otherwise.ith metadata you can; 1) Search resources by all kinds of criteria; 2) Identify various resources; 3) Collect resources by topic; and Trace resources.

3. RESEARCH METODHOLOGY

The concept of ontology derives from philosophy, and the most representative definition of ontology is on the concept of sharing agreement, including the concept of shared conceptual framework for domain knowledge modeling and representation protocol communication protocol interoperability and domain specific theory. In the context of knowledge organization and sharing, ontology is used to define and express the form of words. The Web architecture proposed by Tim Berners Lee, the ontology layer plays a key role. Ontology can provide a unified concept description standard for resources of different types and different sources, therefore it can realize semantic annotation of resources and semantic interoperability between resources and different ontologies aggregated, which is of great utility to solve the problem of digital resource island, and then achieve multi-level and all-round depth integration of digital resources. In knowledge organization, the main function of ontology is to build an Ontology Conceptual Model which is based on resource set, and through semantic annotation of resources, the model is organized into interconnected knowledge networks, consequently it can demonstrate knowledge structure of resources in clear way, and could achieve better knowledge retrieval.

4. RESULT AND DISCUSSION

4.1 Resource Query in the Environment of Semantic Publishing - Clinical Key Platform cases study of elsevier semantic publishing

Elsevier, the world's leading STM publisher, publishes a rich collection of papers every year, accounting for about a quarter of the publishing market each year. Elsevier is a leading STM information products service provider in the world. It has carried out many practices in semantic publishing and has provided a lot of experience in semantic publication of academic journals. Clinical Key is a medical information platform launched by Elsevier in 2012. It has the largest medical information resource base in the world, and can provide rich, one-stop access to resources. It integrates 12 major categories of resources such as journals, books, surgical videos, pictures, videos, clinical trials, treatment guidelines, and drug monographs to facilitate users to access a comprehensive and comprehensive range of resources without boundaries on the same platform. The Clinical Key platform has created the EMMeT, Elsevier Medical Classification, to provide deep indexing of vast amounts of medical content. Content under Clinical Key is indexed by Elsevier Medical Taxonomy (EMMeT) to support semantic search, helping to reduce the amount of time users repeatedly screen in mass search results and provide the most relevant information more efficiently.

EMMeT is Elsevier's proprietary IP classification and is the core technology for implementing Clinical Key's intelligent content and retrieval. Based on the UMLS, EMMeT integrates with many of the current standardized medical taxonomies and thesauri, including the MeSH Medical Thesaurus and ICD International Classification of Diseases. This taxonomy provides ontological support for Elsevier's integration and retrieval of medical resources, thereby establishing a semantic fusion, well-structured and highly organized resource system. Core concept vocabulary can deeply describe the metadata in medicine and reveal the semantic relationship between them, and can provide users with more intelligent retrieval services. Clinical Key enables intelligent clustering and intelligent analytics for intelligent retrieval. With the input of the search term, the system will automatically give the search suggestion, including the disease, disease staging, examination and treatment, and treatment drugs (Figure 1). Users can sort the results according to the needs of screening, Clinical Key support according to the type of resource, literature time and medical specialist to filter the search results, but also according to the clinical workflow presentation related search results to narrow the scope of the search to help users find the most relevant search Results (Figure1.2).

Figure 1Elsevier ClinicalKey Platform Search Interface

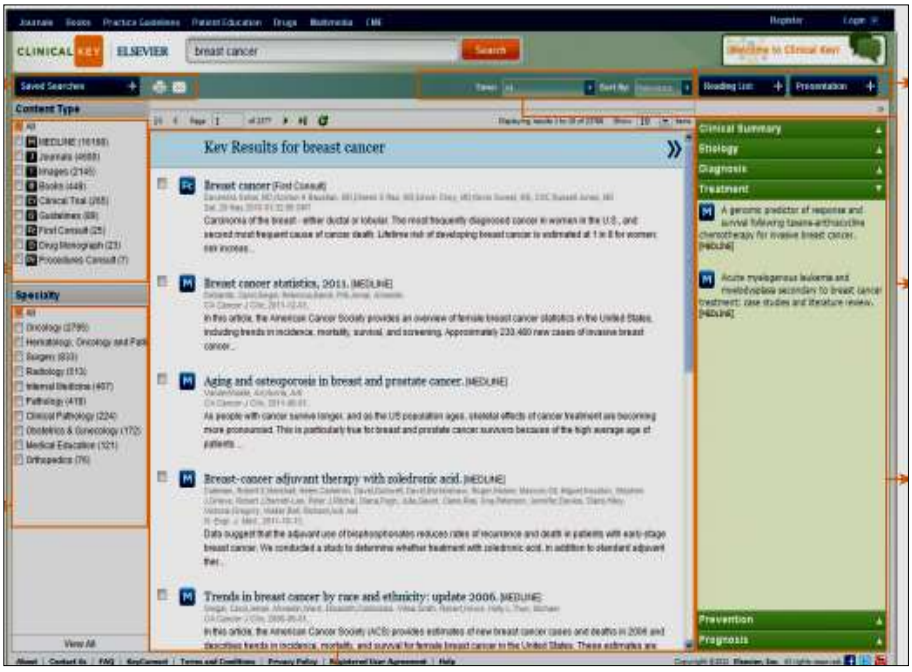

Figure1.2 Elsevier Clinical Key platform search interface

4.2 Elsevier Clinical Key platform search interface

Elsevier launched the "Article of the Future" program in 2009 in response to changes in the current academic environment and user habits. In 2010, the project was first implemented by CELL Press; in 2013, Elsevier improved Journal articles, navigation bar and articles online readable, Science Direct online articles all support HTML format. Combining the characteristics of different disciplines and the characteristics of the user's research literature, the "Article of the Future" project is currently in business management, electrochemistry, materials science, mathematics and computer science, paleontology, tropical diseases and parasitology, psychology and recognition The seven disciplines of knowledge science construct related prototypes. Elsevier believes that the ultimate direction of semantic publishing is highly automated intelligent content rich in

semantic knowledge, to provide scientific research users a better platform for the exchange of scientific information. The "Article of the Future" project aims to: ① provide users with a good online reading experience by providing a friendly interface for online reading; ② help authors to share more relevant information through relevant semantic links, for example, experimental data; ③ Link the article with the source information and present it in the relevant information to enhance the added value of the information.

First of all, Elsevier has done a lot of design on page presentations. The articles presented under the "Article of the Future" project have a three-page layout, including the navigation bar, main content area, and right-hand column. The navigation bar on the left supports navigation of the content of the article, the user can choose the way the chart thumbnails and no thumbnails. The main content area shows the content of the essay. The right-hand column complements the main content area and provides more context-related content based on the characteristics of the article content. In the aspect of the content of the article, the core argument of the article is extracted to form the integration of the research highlights, after the author's information and before the abstract of the article. Support video dynamic demonstration of research results, interactive charts, interactive tree diagrams, mathematical formulas, etc., and embedded Google Maps as additional access to geospatial data. The "Article of the Future" project also creatively provides a short video summary that allows authors to express their main research content and findings informally.

In contextual integration of the article, Elsevier correlates author information, open source data sets, terminology concepts, and reference links, related PPTs, source code software environments and related algorithms, integrates 3D views, experimental flow charts, experimental setups Equipment and other information. In the Article of the Future project, the internal knowledge units of journals can establish links with external knowledge organizations, and can also link the knowledge units within the articles so as to achieve a better organization of the overall structure of the articles and facilitate the reading by scientific users and the recognition of the contents of the articles know. For example, in the article Scholarly leadership of the study of leadership, by clicking on the author's name, the author's contact information, organization, personal homepage and other information appear on the right side of the Page. For user-friendly reading, the Article of the Future project also links the text of the introduction text to the graphic itself. Links with external knowledge organizations and links to semantic elements within journal articles are not only closer to users' reading habits, but also show the structure of journal papers more clearly and help scientific users to obtain information more easily and quickly.

4.3 Features of Elsevier Semantic Publishing

Elsevier's publishing model of semantic publishing has changed the traditional publication of academic periodical linear publishing methods, make full use of semantic tagging and hyperlinks, etc., to provide users with a good reading experience and the best way to browse, Elsevier The semantic publishing model presents the following features:

1. one-stop integrated search. Elsevier semantic publishing platform integrates comprehensive and authoritative subject information, rich variety of resources, such as Clinical Key provides a variety of resources, including medical books, medical

journals, expert comments, etc., covering all clinical specialties, users can The same platform without boundaries to obtain a comprehensive and comprehensive rich resources, no need to repeatedly check the information obtained in different medical information base to enhance the efficiency of search and utilization.

2. Support semantic retrieval to achieve accurate retrieval of matching results. EMMeT has the core concepts and vocabulary to deep description and indexing of medical data to reveal the semantic relationship during the semantic search to support semantic matching can accurately match the search results to reduce the number of scientific research users repeatedly find time in massive results for users provide one of the most relevant answers.
3. Presents the results simple and intuitive, user-friendly. Elsevier semantic publishing platform under the platform user interface is simple and intuitive, friendly and friendly. Personalized search function, customized function, PPT production tools and mail sharing function to further save the user's search time, improve the efficiency of information sharing, user-friendly information exchange and sharing.
4. To provide intelligent knowledge. Articles in journals can be linked with external resources, and can also be used to link knowledge elements within an article, which not only improves browsing experience, but also helps users to obtain relevant information and knowledge in all aspects.

4.4 The Digital Resource Aggregation Model of Academic Periodicals for Semantic Publishing

Based on the problems faced in the current digital process of academic periodicals and the advantages and value of the digital resource aggregation oriented by semantic publishing, this paper constructs a digital resource aggregation model for the academic periodicals oriented to semantic publishing. This model is based on the requirement of semantic publishing on the aggregation of digital resources, aiming at solving the problems in the publishing of academic periodicals and realizing the resource aggregation from three different levels. The first layer is based on the ontology of digital resource aggregation, the second layer is based on the linked data of the digital resource aggregation, and the third layer is based on the access and application of linked data.

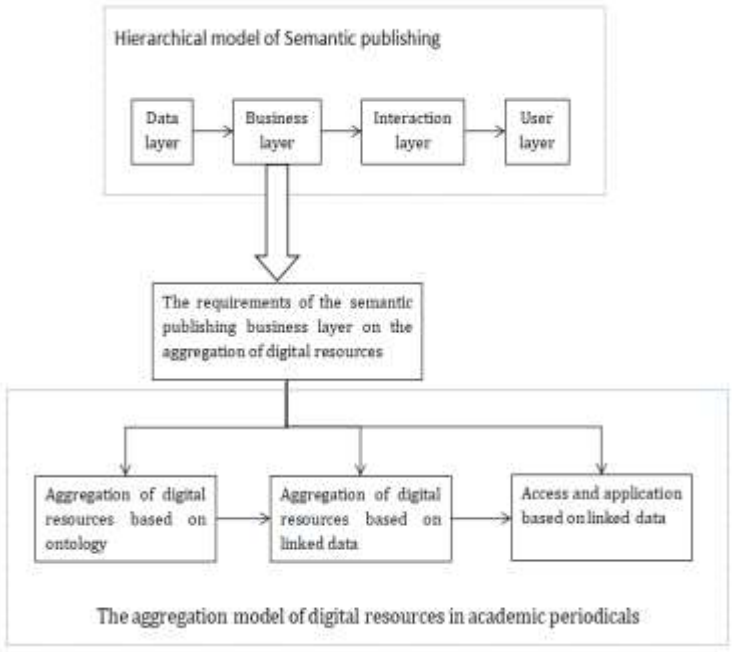

Figure 3.1 Semantic publishing oriented digital resource aggregation model for academic journals

4.5 Hierarchical Model of Semantic Publishing

The data layer is the basis of the framework of the semantic publishing system, which is used to store large amounts of structured, semi-structured and unstructured documents and data. In semantic publishing, data-layer data types are more diverse. In addition to the traditional single document as a storage unit of the literature database, a large number of complete knowledge expression and reuse potential of the "knowledge unit" will be stored separately as a separate knowledge base, such as raw data, tables, pictures, abstracts, conclusions, citations and so on. In the field of knowledge service, people always hope to change the organization structure of knowledge through effective knowledge organization to free access, free reorganization and free use of knowledge, so that those changes can help improve the effect of knowledge service

On the basis of data layer, the business layer mainly realizes the recognition of finer granularity of knowledge unit, and the Automatic links of Knowledge units at all levels, so as to build a large and dense knowledge network. These links include the links between the literature level, such as the link between the current paper and the relevant literature;The links between the various parts of the paper, such as"As shown in the figure" or "See Formulas" --the link to the diagram or formula in the article; The links between the Knowledge Unit and the literature, such as the link between the reference literature and the full text of the cited document; And the links between the knowledge Unit and the external explanatory resource, such as the link between the conceptual term and the external knowledge base and the Ontology library. Interactive layer is the foreground display and user interaction interface of semantic Publishing system, which is mainly used to instruct users to discover, read and use the resources in the station. The user layer is primarily used to identify user needs, build user interest networks, and provide personalized services. In the semantic publishing hierarchy model, we will use

Web log mining to identify and analyze the information needs and web behavior characteristics of each user, and then construct the interest map, recommend users with similar interests and resources that users may be interested in, and improve the utilization level and performance of the system and resources.

4.6 The requirements of the Semantic Publishing Business Layer on the Aggregation of Digital Resources

In the semantic publishing hierarchy model, the data layer is mainly used to store all kinds of data, and to realize user's free access and utilization of information and knowledge through the establishment of knowledge base. The business layer emphasizes the deep processing and integration of the data resources in order to realize the semantic recognition of the knowledge unit and the semantic links between them, and to promote the automatic discovery and acquisition of knowledge by constructing knowledge network. Through the technology of information visualization, the interactive layer presents the content of the semantic annotation and related articles to the users in a vivid, easy and understandable way, and helps them understand the complex knowledge content. The user layer is to reveal the user's interest and to carry out personalized recommendation service based on the analysis of the user's browsing and utilization of resource records, and the use of data mining and algorithm recommendation technology.

First of all, the semantic publishing requires the digital resource aggregation to realize the structural description of the periodical content and the semantic annotation of the internal knowledge Unit. Academic periodicals from different sources often adopt different metadata standards, which brings about semantic and heterogeneous problems for resource aggregation, and it is difficult to realize the semantic interoperability between resources, and the problem of Information Island is serious. Semantic publishing requires semantic technology to provide a unified concept description standard for different types of resources in order to realize the structural description of resources and semantic annotation of knowledge units and semantic interoperability between resources, and solve the semantic heterogeneity of periodical content.

Secondly, semantic publishing requires the realization of the interrelated resources and enrich the content of the publication. On the traditional Internet, the links between digital objects is mainly realized by hyperlinks, which are mostly static and difficult to update. Semantic publishing can realize the rich, multi-level and automatic interrelation between digital objects by means of semantic technology. This association includes not only the relationship between periodicals and periodicals, but also the relationship between the internal knowledge units, the knowledge unit and the external explanatory resources. By creating links, semantic associations are established between different sources and types of data, making them seamless, open-ended, and enriching the knowledge of the publication. Finally, semantic publishing requires the construction of a knowledge network. Based on the description of resource content structure and the semantic association between resources, a large and extensive knowledge network is built through the technology of linked data, which can get rid of the isolated state, realize the rich semantic association between resources and external resources.

4.7 Aggregation of Digital Resources Based on Linked Data

Linked Data is a way to publish and link various kinds of data information, presented by Berners in 2006. Its purpose is based on the current web to create a network of data that joins all the natural, social and spiritual worlds, with a computer-understandable description of everything and the relationship between things and things, building up close links between data. So that anyone can use the Internet to efficiently find, use and share this information and knowledge [5] Berners proposed four principles to be followed for publishing linked data:[6] First, the linked data publication should use the URI as the identification of anything; Second, by using the HTTP and URI, anyone can achieve access to the data, and third, through (RDF, SPARQL) standards to provide users with access to information channels; The forth, the linked data publication should provide as many related URIs as possible. Linked data can play a key role in building a network of data. Compared with the traditional network, the data network based on the linked data can provide the users with more accurate and richer information resources and better meet the needs of the users to obtain the resources. The function of linked data in the aggregation of digital resources mainly embodies the following three points:

a) Links different sources, different formats of data to achieve resource aggregation and association. Linked data is the key technology of constructing data network, which can play a powerful role in resource aggregation. It can make a resource in a different database, or data in the consolidated database that cannot be associated and interoperable because of a different format type, into a seamless whole by creating a link to implement semantic association, while also enabling the association with external resources.
b) Reveals the link between resources effectively. To publish the digital resource information of academic periodicals in the form of linked data, we can access resources directly through the HTTP protocol, and use RDF links to access other related resources to reveal the interrelationship between resources.
c) Establishes the link between resources and enriches the knowledge content of resources. Through the linked data, it can provide the technical framework for the academic periodicals to connect more and more external open resources, therefore it can greatly enrich the knowledge content of the periodical. Thus, the publication of academic periodicals can not only embody the contents of the periodical itself, but also realize the connection of the data resources in the semantic level. This also satisfies the requirements associated with semantic publishing.

According to the principle of association data, in the second layer of the digital resource aggregation model of the academic periodicals oriented to semantic publishing, we will use the URI address that the HTTP protocol can dereference to name all the resources, and then describe the digital resources on the basis of the metadata ontology model. The SKOS semantic description is used to convert other resources into a unified RDF format for resource description. By connecting various related resources through RDF, the aggregation of different knowledge units in digital resources of academic periodicals and the links with external knowledge organizations are realized.

4.8 Access and Application Based on Linked Data

This layer is used to realize the browsing and querying of digital resources of academic periodicals, and the user can get concise and clear query results through linked data, and then extend the links to other linked data resources to realize seamless interface between resources. After the data aggregation and the linked data are released, the data set can be browsed through the appropriate data browser and the RDF link between the data is used to guide the user through the different data sources. The semantic search engine can also be used to retrieve the necessary information resources from the linked data network. Span thought that linked data was not a completely new technology, its main technical foundation was URI, HTTP and RDF, and so on. On this basis can develop the linked data release, storage and other tools.[7] These tools include D2R server, Virtuoso Universal Server, Pubby, SIOC exporters, etc. When a large number of linked data is published on the network, the browsing and retrieval queries can be implemented by means of the linked data browser and the linked data search engine. Through the linked data Browser, the user can provide the navigation between resources. With the help of RDF links, users can freely switch between different data sources and realize the whole data network browsing, improve browsing efficiency. The user's query of linked data can be realized by using the linked data search engine, and the linked data and relevant information are crawled in the network. In addition, in linked data access, it is possible to combine its browsing and retrieval techniques to provide more intelligent data services to users.

5. CONCLUTION

The semantic publishing is a publication that examines the meaning of analysis and is formal and reviewed according to the idea, whereas the digital publishing is one part of a cryptographic method that uses hash functions to mark a data. This kind of information service cannot solve the problem that there exists a large number of polysemous words in the domain and cannot help the users to carry out deeper knowledge discovery and semantic mining. For example, in querying the literature on "coronary heart disease," the database only matches articles that contain "coronary heart disease", while articles which contain synonyms for other coronary heart disease (such as "coronary artery disease" or "coronary artery spasm") are failed to be retrieved. But in fact these documents are also what users need. This keywords-based matching method not only reduces the recall of information retrieval, but also reduces the efficiency of information retrieval.

REFERENCES

[1] Shotton D, Portwin K, Klyne G, Miles A (2009). "*Adventures in Semantic Publishing: Exemplar Semantic Enhancements of a Research Article*". PLoSComputBiol 5(4): e1000361.

[2] Daedalus, "*Journal of the American Academy of Arts and Sciences; Proceedings of the American Academy of Arts and Sciences*", 125 (4).

[3] A. Crystal, J. Greenberg, (2006). "*Relevance criteria identified by health information users during web searches, Journal of the American Society for Information Science and Technology",* 57 (10).

[4] Achille Felicetti., Hubert Mara (2008). "*Semantic Maps and Digital Islands: Semantic Web technologies for the future of Cultural Heritage Digital Libraries*". Workshop Proceedings. Spain, p.51-62.

[5] Liu Wei. linked data: "*Concept, Technology and Application Prospect"* [J]. University Library Journal, 2011, 29 (2): 5-12.
[6] Tim Berners-Lee. Linked data[EB/OL].http://www.w3.org/DesignIssues/ LinkedData.html,2016-03-21.
[7] Pan Youneng, Zhang Yue. "*Research and Application of Linked Data*" [J]. Information science, 2011 (1): 124-130.
[8] Borgman, C. L. (in press). "*From Gutenberg to the Global Information Infrastructure: Access to Information in the Net-Worked World"*. Cambridge, MA: MIT Press
[9] Borys Omelayenko, (2008). "*Porting Cultural Repositories to the Semantic Web. Workshop Proceedings".* Spain, p.14-25.
[10] Dave Reynolds., Paul Shabajee (2004). "*Semantic Information Portals"*. New York, USA.
[11] David Hawking., Justin Zobel, (2007). "*Does Topic Metadata Help With Web Search?. Journal of the american society for information science and technology"*. p.613-628.
[12] Eero Hyvonen., Eetu Makela, (2005), "*Museum Finland : Finnish Museum on The Semantic Web. Helsinki Institute for InformationTechnology* (HIIT)", Finland.
[13] Ganesh D. Bhatt (Ganesh D. "*Bhatt is an Assistant Professor in the Department of Information Science and Systems",* Morgan State University, Baltimore, Maryland, USA.).
[14] Georgia, D.Solomou., Dimitrios A, Koutsumitropoulos, T.S.P (2008). "*Semantics-Aware Querying of Web-Distributed RDF (S) Repositories"*. Workshop Proceedings. Spain, p. 39-50.
[15] Hildebrand, M., Ossenbruggen, (2007), "*An Analysis of Search-based User Interactionon the Semantic Web"*. University of Technology. Amsterdam
[16] http://dx.doi.org/10.1371/journal.pntd.0000228.x001
[17] https://doi.org/10.1371/journal.pcbi.1000361
[18] http://www.stm-assoc.org/brussels-declaration/
[19] http://ilmuweb.net/semantic-web/web-semantik/
[20] http://ilmuweb.net/semantic-web/rdf-resource-description-framework-pada-web-semantik/
[21] http://dy06cool.blogspot.com/2012/03/lebih-dalam-mengenai-web-semantik.html
[22] http://buset.staff.gunadarma.ac.id/Downloads/files/15813/semantic-web.pdf