Yang, X., & Yu, S. (2015). Designing a Resource Evolution Support System for Open Knowledge Communities. *Educational Technology & Society, 18* (4), 385–400.

Designing a Resource Evolution Support System for Open Knowledge Communities

Xianmin Yang[1] and Shengquan Yu[2*]

[1]Institute of Education, Jiangsu Normal University, China // [2]Institute of Modern Educational Technology, Beijing Normal University, China // yangxianmin8888@163.com // yusq@bnu.edu.cn
[*]Corresponding author

(Submitted August 24, 2014; Revised December 12, 2014; Accepted March 4, 2015)

ABSTRACT

The continuous generation and evolution of digital learning resources is important for promoting open learning and meeting the personalized needs of learners. In the Web 2.0 era, open and collaborative authoring is becoming a popular method by which to create vast personalized learning resources in open knowledge communities (OKCs). However, the essence of openness of OKCs also gives rise to concerns regarding the knowledge quality and non-orderliness of resource evolution. In this study, we design a resource evolution support system (RESS) called learning cell system (LCS) in one OKC. Two key issues, namely, the intelligent control of content evolution and the dynamic semantic associations between resources, are addressed by combining technologies of semantics, trust evaluation, rule-based reasoning, and association rule mining. One typical case is taken to illustrate the actual evolution process of learning resources assisted by RESS in LCS. The operating effect of this system shows that RESS can control content evolution and effectively build semantic associations among resources. Finally, the academic contribution to the OKCs, implications for educational practice, limitations, and future research plans are presented.

Keywords

Open knowledge community, Learning resource, Resource evolution, Evolution control, Semantic association

Introduction

With the spread of Web 2.0 ideas and technologies, open knowledge communities (OKCs) have become increasingly popular. OKCs can be used as knowledge management tools and virtual learning environments for learners. Considerable research has recently been conducted to examine the potential of OKCs in educational application scenarios. For example, some scholars investigated the effect of student-generated content (SGC) on student engagement and learning outcome in wiki-type OKCs (Li, Dong, & Huang, 2011; Wheeler et al., 2008). In OKCs, any valid user can create new learning resources and coedit existing resource content (Yang et al., 2014). The involvement of many users ensures the continuous generation and evolution of learning resources while meeting the growing personalized needs of learners. Nevertheless, a credibility issue is associated with this collaborative authoring approach (Luo & Fu, 2008). Inevitably, a large number of non-credible, poor-quality contributions exist in this type of open and unmonitored environment. These contributions hinder the effective and efficient use of OKCs in the open learning era.

The complexity of the user groups and the liberalization of production in OKCs directly result in the explosive growth and disorderly evolution of digital resources. For example, facing the pressure of trust crisis (Seigenthaler, 2005; de Laat, 2014; Sapienza & Zingales, 2012), Wikipedia is setting up an increasing number of rules, including administrative rules, fact-checking rules, real-time peer review rules, and rules that prohibit the creation of new articles by anonymous users. In addition to these rules, Wikipedia also provides manual revision and check functions based on feedback. The purpose of all these measures is to ensure the high quality and orderliness of resource evolution (Wang, 2009). How to control the content evolution of learning resources adequately has become a realistic problem that should be addressed (Yang, 2012).

In addition, resources should not be isolated from one another to promote effective and efficient learning. Along with the development of semantic web technologies, the establishment of semantic associations among learning resources brings significant value according to connectivism theory (Siemens, 2005). First, the resource retrieval burden of learners is reduced, and target resources can be efficiently located. Second, an individual resource associated with many external entities will have a higher probability of being retrieved, viewed, and improved. Currently, the semantic association among e-learning resources in OKCs is insufficient. Moreover, most associations are described by static metadata and built by manually editing hyperlinks between resources (Shao et al., 2008; Yang et al., 2013), such as the entity associations in Wikipedia and Hudong Baike (http://www.hudong.com). The integration of semantic technology and resource association has the potential to address the aforementioned problems, which can aid in the development of OKCs and open learning.

Literature review

Definition of resource evolution

Ecology is a discipline that focuses on species distribution, diversity, interactions among and between species, as well as the external environment (Begon et al., 2006). An integral ecosystem has the features of entirety, openness, dynamic equilibrium, self-organization, and sustainable evolution. The concept of evolution stems from ecology. Evolution is the change in the inherited characteristics of biological populations over successive generations. The application of evolutionary principles has occurred in different fields, such as economics, culture, computer science, and management. Evolution refers to the process of developing from simple to complex (Zhao, 2006). Development, change, and adaptation are the core ideas of evolution. Organisms change and grow gradually by interacting with and consequently adapting to their external environment. The evolution of learning resources specifically refers to the improvement and adjustment of resource content and structure (Yang & Yu, 2011). The aim is to meet the dynamic and personalized learning needs of learners. The orderly control of evolution means that an individual resource (or an entire resource population) develops in a desired direction in a quality-controlled manner.

Models of resource evolution

Resource evolution may be described according to two different models: content evolution and association evolution.

Content evolution

Content evolution refers to the constant updating and improvement of resource content through the collaborative content editing of many users (Yang & Yu, 2011). The process of content evolution is exhibited by continual content version changes.

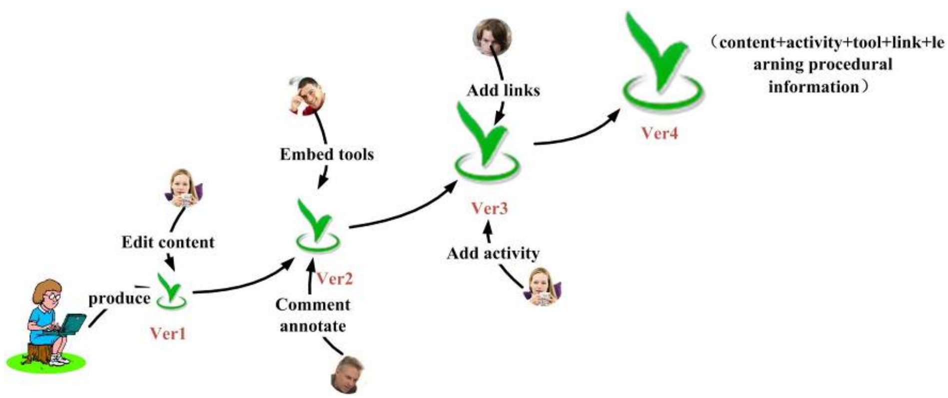

Figure 1. Content evolution of learning resources

Figure 1 shows the process of content evolution. The producer creates and publishes a learning resource and attracts collaborators to edit the content. Given that the resource is available to the public, any user can revise content, add photographs, or insert external RSS (Really Simple Syndication) resources. As the resource grows, more users come into contact with it. These users comment, add activities, and annotate entries. By harnessing this collective wisdom, the content is constantly updated. Finally, the content evolves into a high-quality learning resource that meets the learning demands of different users.

Orderly control of resource content is a core issue in the evolutionary study of learning resources. In an open environment, resource evolution is centered on Web 2.0 technologies. For quality control, manual checking is commonly used to ensure that evolution is appropriate. With the help of manual labor, the versions evolve toward a higher quality and reliability. The quality control model based on manual checking is time- and labor-intensive. Ultimately, the quality control model fails to meet the requirements of dynamic generation and infinite expansion of learning resources.

Association evolution

Association evolution is the process of establishing and enriching semantic associations, which are mainly exhibited by the expansion of the resource network (Yang & Yu, 2011).

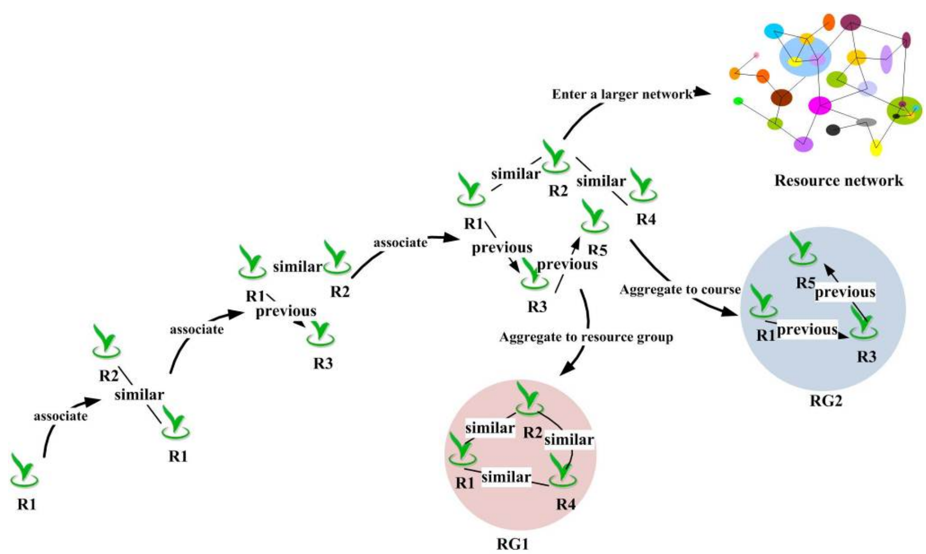

Figure 2. Association evolution of learning resources. *Note.* R, resource; RG, resource group.

Figure 2 shows the ideal process of association evolution. From the time of its inception, an individual resource strives to connect with other resources to enhance survival capability. The resource establishes an increasing network of semantic associations with other resources through manual operation or automatic discovery. Semantic relationships may include similarity, hyponymy, precursor, or equivalence. The growth of these associations provides data support for resource aggregation. With the use of specific resource aggregation mechanisms, resources that are similar will automatically form subject-based resource groups, whereas resources in semantic order will form knowledge chains with different learning priorities. Finally, any learning resource will become one node in an infinitely expanding resource network and will subsequently develop through dynamic association with other nodes.

Dynamic building, mining, and developing of the semantic associations among resources are vital and constitute the objectives of resource evolution. The final objective of building associations between resources is to promote effective learning. Associations between resources can help produce a resource network that is infinitely expandable. The technology used for the automatic building of semantic associations will overcome the limitations of manual manipulation, such as high levels of subjectivity and demands of time and labor.

Evolutionary mechanisms and technologies in OKCs

In the current situation, most current OKCs, such as Wikipedia, Hudong Baike, and Cohere (http://cohere.open.ac.uk/), still use the manual control approach to check content quality and achieve the orderly evolution of learning resources (Yang, 2012). Content version management is a commonly used technology. However, this method is time and labor consuming with unsatisfactory efficiency. To reduce manual labor and accelerate resource evolution in OKCs, some studies have been recently conducted to improve the mechanisms of resource evolution in Wikipedia with the use of trust evaluation technology. Adler et al. (2008) proposed a system that computes quantitative values of trust of Wikipedia articles. Maniu et al. (2011) studied the signed web of user trust from user interactions. The values of article trust and user trust can provide an indication of reliability. However, only a few studies have been further conducted to control content evolution automatically. In addition, Vrandecic (2009) proposed the use of semantics to check content quality automatically. Although controlling resource evolution in OKCs is a good idea, Vrandecic failed to not propose any feasbile technical solution. Therefore, a new and intelligent method is required to control content evolution automatically in OKCs.

Regarding the technologies of association evolution, some studies have been conducted in recent years in addition to the method of manually editing hyperlinks. At present, the learning object (LO) is the main resource form in the e-learning domain. Some scholars studied the association technologies of LOs, which mainly involve relationship metadata design (Ullrich, 2005; Lu & Hsieh, 2009), association representation design (Lv & Du, 2010; Shi et al., 2003), similarity measurement (Zhang et al., 2006), association path search (Li, 2005), and automatic assembly (Farrell et al., 2004). Generally, current research focusing on association technology has two shortcomings, as follows: (1) Representation of the association relationship mainly adopts static metadata description technology without considering the semantic relationship among resources and without standardized methods to describe resource association. (2) Although some researchers have considered the calculation of semantic relationships from the perspective of resource ontology, they are mainly restricted to similarity relationship measurement, ignoring the dynamic nature of semantic relationships (e.g., preorder, successor, and opposition) of other resources. Therefore, we should enhance the relationships between learning resources from the perspective of semantics in OKCs and examine a new technical solution.

Research objectives and questions

In summary, content evolution control and semantic associations of learning resources have become the two major practical issues in the development of OKCs (Yang, 2012). To address the above issues and promote the orderly evolution of learning resources in OKCs, we propose a resource evolution support system (RESS) for open knowledge communities. Two major research questions are as follows:

- How can the direction of content evolution be intelligently controlled in the process of multiuser collaborative content editing?
- How can rich semantic associations be built among learning resources in an automated manner?

Functional architecture

A learning cell system (LCS, http://lcell.bnu.edu.cn) is an ideal platform for a study of this nature because it is specifically developed to support open social learning. LCS provides an open application–programming interface for third-party developers. This platform also supports collaborative authoring and ontology management, which saves time and maximizes efficiency. In this study, we develop the RESS based on LCS. We then provide an overview of LCS and describe the system architecture design of RESS.

Overview of LCS

LCS is an OKC for open learning (Yu et al., 2015). The platform consists of six functional modules: Llearning cell (LC), knowledge group (KG), knowledge cloud (KC), learning tool (LT), learning community (LCm), and personal space (PS). Unlike the ad hoc collaboration of Wikipedia, users actively collaborate on a learning resource to achieve a common learning objective.

LC is a resource entity that can be a lesson or a knowledge point. Each LC contains not only content but also learning activities, semantic information, and generative information. The LC can introduce related assistant LTs to support learning. Each KG consists of LCs on related subjects. For instance, a course can be a KG, and each lesson or knowledge point in the course can be an LC. When users access KG, they can find all of the LCs related to the course. The KC aggregates multiple KGs. Different KGs are connected via semantic relationships. In a KC, users can easily find all of the KGs related to their subject. The LT assembles all of the personalized learning gadgets, which can be used by LCs, KGs, PSs, and LCms. An LCm is a collective learning environment in which community members collaborate and share with one another. For instance, members can publish a notice, initiate a discussion, share resources, and initiate learning activities. In addition to LCs, all users have their own personalized learning environment (PLE). In the PLE, users can post basic personal information, manage (create, collaborate, and subscribe to) LCs and KGs, and select recommended learning resources.

All users can edit and improve the resource content within the framework of collaborative editing (Figure 3), which is crucial for sustaining the evolution of content. To ensure security during evolution, a content version management function was implemented in LCS.

Figure 3. Screenshot of collaborative content editing

LCS resources are organized using ontology technology. LCS includes three different ontologies, namely, knowledge, user, and context ontology. An ontology management module was developed based on the JENA framework. Unlike traditional ontologies created by experts, LCS adopted an open and collaborative technique to construct discipline ontologies. Any user from a particular disciplinary background was permitted to create the corresponding discipline ontology. In addition to manual construction, some core discipline concepts were extracted automatically from resource contents to enrich the ontology base. External ontology bases (e.g., FOAF (http://www.foaf-project.org/), vCard (http://www.foaf-project.org/)) could be quickly imported, and the ontologies generated in LCS could be exported to standardized web ontology language files for sharing with other systems. Ontology refining was implemented to remove outdated and unqualified concepts and properties automatically, as well as to guarantee the quality of ontologies.

System architecture design

The functional architecture of RESS in LCS is shown in Figure 4. Technologies, such as Java 2 Platform, Enterprise Edition, ontology, trust evaluation model, rule-based reasoning, association rule mining (ARM), and Flex were adopted. The three main modules are content evolution control, semantic association, and visual presentation of the resource evolutionary path.

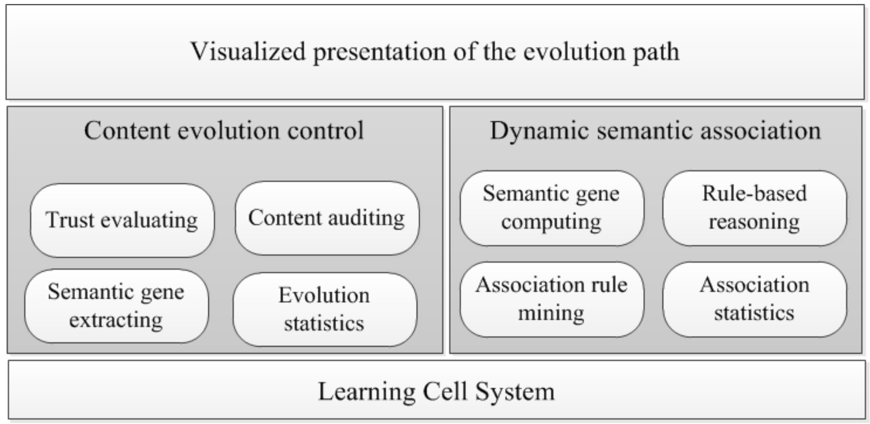

Figure 4. Functional architecture of the RESS

Content evolution control

The core function of this module is to take intelligent control of the direction of resource evolution by automatically reviewing new versions of content. Semantic gene abstraction and trust computing underpin evolution control. The semantic gene is the basic information unit representing the meaning conveyed by the resource content. Formally, the semantic gene appears as a set of concepts in domain ontologies with assigned weights and as a set of semantic associations among these concepts. The trust evaluation model is commonly used for assessing the trust level of any entity involved in network interaction. Trust evaluation is modeled on trust relationships in society. By effectively integrating the semantic gene and user trustworthiness, the system could automatically conduct content version auditing, with the process stored in system logs (Figure 5). The statistical function of evolution control could help the system administrator in mastering the overall effect of intelligent content evolution control.

Figure 5. Logs of content evolution control

Dynamic demantic association

Dynamic semantic association differs from webpage hyperlinks, which have no semantic association. The module automatically establishes standardized semantic relationships among resources. Association relationships among resources can vary and need to be updated and developed with each change of content. Lu and Hsieh (2009) identified 15 new semantic relationships based on the metadata in SCORM (Sharable Content Object Reference

Model). Using these relationships, 32 semantic relationships were used for building resource associations. The system administrator has rights to manage all of the relationships. Learning resource associations can be automatically established, and the results of resource associations can be analyzed statistically and presented visually. The system provides three ways to build semantic associations among resources using semantic gene computing technology, rule-based reasoning technology, and ARM technology. Figure 6 shows the semantic association network of LCs in one KG.

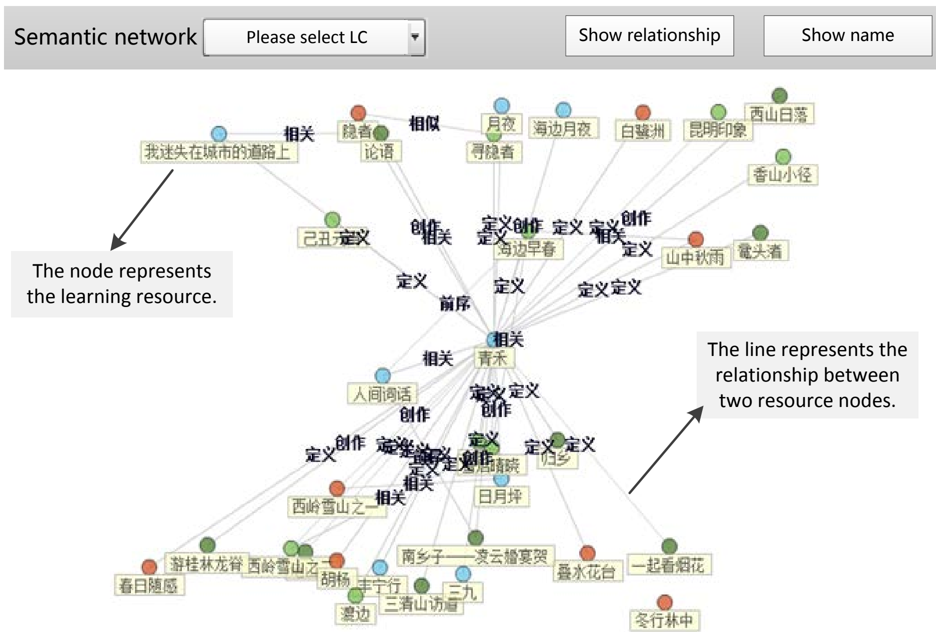

Figure 6. Semantic associations among learning resources

Visual oresentation of the evolutionary path

Flex technology was used to visualize the resource evolutionary path and to present the process of resource evolution more intuitively. Learners could view the history of resource improvement and progress, which may be helpful for learning. The changes of resource content and associations at different times were presented visually (Figure 7).

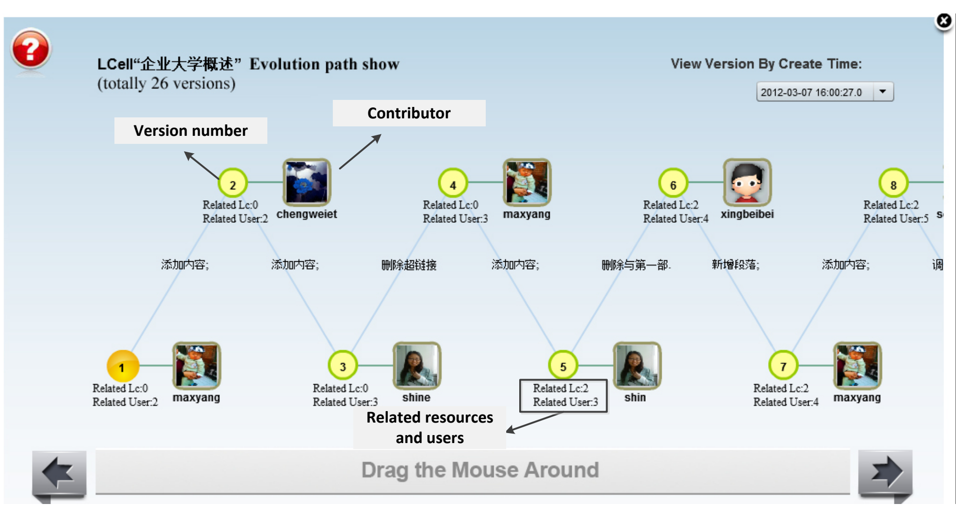

Figure 7. Visual presentation of the resource evolutionary path

Implementation technologies

Corresponding to the aforementioned specific research questions, we propose two solutions by combining technologies of semantics, trust evaluation, rule-based reasoning, and ARM.

Intelligent control of content evolution

The intelligent control method of content evolution is designed based on two assumptions, as follows:

Assumption 1: The evolution of content centers on specific subjects, with strong semantic associations usually linking content before and after changes are made.

The content evolution of resources usually shows a clear direction and is linked to the development of surrounding specific knowledge structures (i.e., the semantic gene). The content introduces a particular subject, which is then updated to gradually enrich and perfect the subject. These "before and after" changes to content usually have strong semantic associations. As such, the newly added content and semantic gene are likely to be semantically similar.

Assumption 2: The behavior of highly trusted users is reliable and inclined to well-meaning content editing.

The trustworthiness of a user derives from a trust evaluation model and is computed by analyzing various interaction data. As a user accumulates desirable behavior, their trustworthiness improves. Once the trustworthiness of a user exceeds a certain value, most operations of that user are deemed reliable and their content editing will be automatically accepted. Figure 8 shows the technical framework for implementing the intelligent control of content evolution based on the semantic gene and the trust evaluation model.

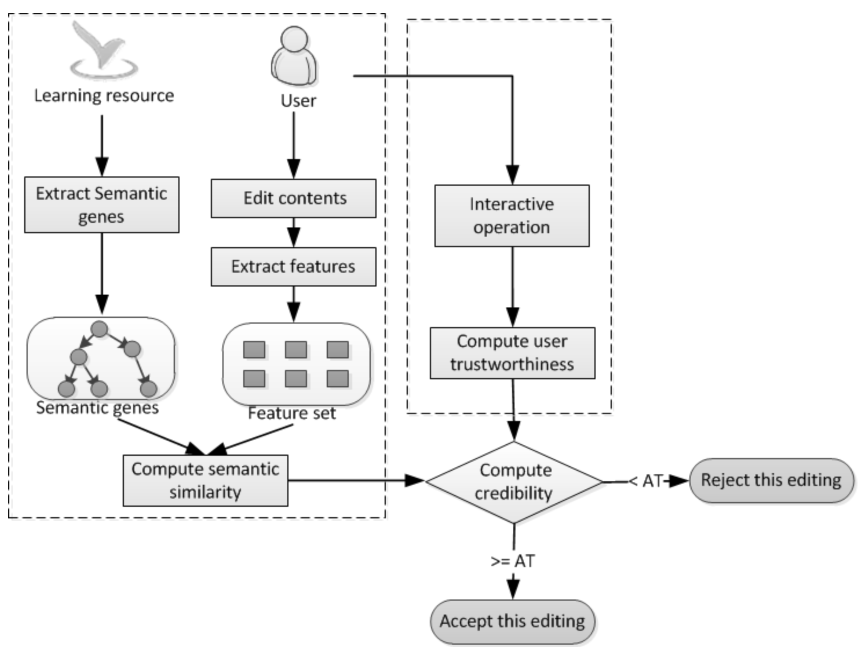

Figure 8. Procedure of intelligent control of content evolution. *Note.* AT, acceptance threshold.

The core idea is to compute the credibility of content edits by integrating two types of information. The first type is the semantic similarity between the semantic gene of the current learning resource and the feature set of the newly added content. The second type is to compute the trustworthiness of the content editor using the trust evaluation model based on key interactive operational data. If the computed credibility level of content editing exceeds the preset threshold, then the content edit will be automatically accepted; if not, it will be rejected. The creator of the resource has rights to modify the audit results manually to guarantee the quality of evolution control. The system will automatically accept or reject the content editing and immediately send e-mail notifications to the editor and manager. This feedback mechanism could be helpful for expediting resource evolution.

The semantic gene can be represented through concept aggregation, involving the core concept and relationship between concepts. Weighting may be given based on a description of ontology. The semantic gene can be presented using the following three-element format: $SG = < CS, WS, RS >$, where CS is the aggregation of the core concept, $CS = \{C_1, C_2, C_3,\ldots, C_n\}$; WS is the weighted aggregate of concept item, $WS = \{W_1, W_2, W_3,\ldots, W_n\}$; W_i is the weighting of C_i; and RS is the relationship aggregation among core concepts, $RS = \{R_1, R_2, R_3,\ldots, R_n\}$. Each relationship is presented using a resource description framework (RDF) triple < *Subject*, *Predicate*, *Object* > of domain ontology, where $R_1 = < Concept_1, Relationship, Concept_2 >$. In this study, $Concept_1$ and $Concept_2$ may not be contained in the CS and could be concepts from other domain ontology bases. *Relationship* is the concept relationship abstracted from the domain ontology base.

To abstract a semantic gene from learning resource content, the resource entity should first be structurally represented. The entity can be represented with four parts, as follows: $Res = < Title, Tag, Content, SemanticData >$, where *Title* is the title of the resource, *Tag* is the tag added to the resource, *Content* is the detailed content of the resource, and *SemanticData* is the ontology-based semantic description information attached to the resource. *Title*, *Tag*, *Content*, and *SemanticData* are the four main sources for abstracting the semantic gene and have different relative weighting for representing the core content of a resource. Weighting aggregation can be presented as WT = $\{WT_1, WT_2, WT_3, WT_4\}$, where WT_1 is the weighting of *SemanticData*, WT_2 is weighting of *Title*, WT_3 is weighting of *Tag*, and WT_4 is weighting of *Content*. After resolving the source and weighting, the characteristic item abstraction technology from the web data mining domain is combined with a series of characteristic words (core concepts) extracted by the domain ontology base. Then, these words are mapped to ontology and stored in CS aggregation. Each characteristic is given a different weighting value using a preset characteristic evaluation function (see Formula (1)). These weighting values are placed in WS aggregation. Finally, using the JENA framework, the semantic relationship of these characteristic words in the domain ontology base is abstracted and placed in RS aggregation.

$$FE(t) = \log\left(CF\left(c, SemanticData\right)\times WT_1 + CF\left(c, Title\right)\times WT_2 + CF\left(c, Tag\right)\times WT_3 + CF\left(c, Content\right)\times WT_4\right) \quad (1)$$

In Formula (1), $CF(c, x)$ represents the frequency of occurrence of concept c in x, $x\in\{SemanticData, Title, Tag, Content\}$. Notably, the initial semantic gene comes mainly from the basic concepts extracted from the title of the learning resource. Along with the development of resource content, its semantic gene will become more accurate and rich.

The text features of new content are extracted using the method proposed by Ray and Chandra (2012). The process involves the adoption of a key word set using statistical methods to calculate the weighting of characteristic words via the characteristic evaluation function. Ultimately, a key word set with appropriate weighting is produced. In contrast to the semantic gene of a resource, the text characteristic of newly added content lacks RS. The semantic similarity calculation of two separate texts requires the use of a vector space-based cosine algorithm (Jin, 2009).

In addition to the semantic gene, the trust evaluation model is another core element in the content evolution control technical framework. Many trust evaluation models have been developed for network communication and electronic business (Jones & Leonard, 2008; Li & Wang, 2011; Denko et al., 2011). Considering the different interactions in OKCs, some researchers have studied trust evaluation methods adapted for OKCs (Rowley & Johnson, 2013; Javanmardi et al., 2010; Moturu & Liu, 2009). However, these studies considered the disadvantages, such as inadequate influencing factors, and separated the treatment of trustworthiness for users and resources. Based on this research, we developed a two-way interactive feedback model (Yang et al., 2014). The model has two core components, namely, resource trustworthiness (RT) and user trustworthiness (UT). The model is based on more interaction data, considers the interrelation between RT and UT, and better represents the features of interpersonal trust. In this study, we limit ourselves to a description of the computing method of user trustworthiness. More details can be found in the study of Yang et al. (2014).

UT is described by four elements: $UT = \{UT_{res}, UT_{col}, UT_{fri}, UT_{rev}\}$, where UT_{res} is the trustworthiness component for a user calculated from the resources that he/she created, UT_{col} is calculated from his/her interaction with other users,

UT_{fri} is based on friendship relationships between this user and other community members, and UT_{rev} is a component calculated from his/her editing history in the community. Equation (2) is used to calculate the user trustworthiness by combining UT_{res}, UT_{col}, UT_{fri}, and UT_{rev}. The relative importance of the four components is described by a user weight set: $UW = (UW_1, UW_2, UW_3, UW_4)$ ($\sum UW_i = 1$).

$$UT = UW_1 \times UT_{res} + UW_2 \times UT_{col} + UW_3 \times UT_{fri} + UW_4 \times UT_{rev} \quad (2)$$

Semantic association of learning resources

The abundant semantic association among learning resources can enhance the interconnection among resource entities, improve the frequency of browsing or content editing for each resource, and promote the rapid evolution of a resource. The abundant semantic association can also provide the data foundation from which to develop dynamic aggregation leading to resource groups with larger size and internal logic relationships.

In this study, we propose the learning resource dynamic semantic association technical framework shown in Figure 9. The framework implements dynamic semantic association in association evolution. Semantic annotation is conducted on learning resources using the concepts and attributes in a knowledge ontology base to generate a semantic information space that contains a range of standardized semantic description information.

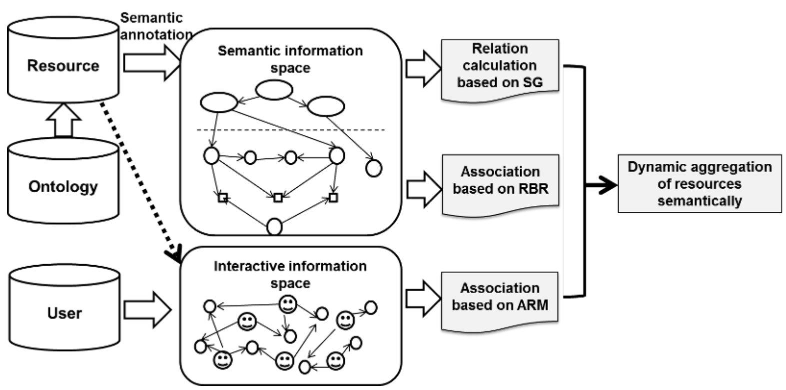

Figure 9. Semantic association and aggregation in association evolution. *Note.* SG, semantic gene; RBR, rule-based reasoning; ARM, association rule mining.

To build semantic associations among resources automatically, we used semantic genes to compute the types of relationships among resources. This method requires the use of semantic dictionaries, such as HowNet (Dong et al., 2007) or WordNet (Miller, 1995). To calculate the similarity relationship based on the semantic gene, we first calculated the similarity between two concepts using semantic dictionaries and domain ontology. Then, we set similarity weighting values in combination with weighting values of concepts in the semantic gene. Finally, we determined whether the similarity relationship between two resources met the similarity threshold. The semantic similarity between two concepts is calculated via the 3D-SIM method proposed by Benwusiyi et al. (Wu & Wu, 2010).

Second, we adopted rule-based reasoning technology to distinguish specific associations based on the existing associations in a range of preset rules, which can be dynamically enriched and changed. JENA is a Java framework for building semantic web applications and provides an inference engine to assist in developing reasoning applications. The built-in inference engine of JENA was used to compile the association rules in this study. First, we compiled various association rules that were saved in an inference rule base. Then, the JENA inference engine extracted the rule from the rule base and linked the rule to the ontology model. Finally, the dominant resource association aggregation was outputted. Before adopting the JENA framework to achieve the dominant resource association based on rule inference, two critical processes were completed, as follows: (1) Data were saved in a JENA-supported ontology model via the RDF triple form. (2) Various association rules were compiled subject to the rule form defined by the JENA inference engine. The inference engine binds these rules and conducts inference using the ontology model to derive a new ontology model.

An invisible association is difficult to detect by human eye. The ARM technology, which is an important and frequently used data mining technology, was introduced to determine these potential associations. The ARM can obtain better results than association rules based on domain ontology (Babashzadeh et al., 2013). In this study, we improved the traditional ARM algorithm by adding semantic gene-based constraints, which can enhance the efficiency and precision of mining association rules. In this study, we improved the a priori algorithm by considering the minimum supporting degree (min_supp), minimum confidence (min_conf), and minimum semantic correlation (min_semrel) required to restrict the creation of association rules to improve the efficiency and accuracy of ARM. We proposed a semantic constraint-based association rule mining algorithm (semantic constraint a priori) (Yang et al., 2013). The minimum semantic correlation refers to the minimum similarity between entities contained in frequent items, where min_semrel can be calculated via the semantic gene of the resource. Through min_semrel, many insignificant candidate items can be filtered to improve the efficiency of the algorithm and, ultimately, create an association rule of higher quality.

Results of system operation

The RESS was released and ran from 1 February 2012. Then, we analyzed the evolution process of learning resources in LCS aided by RESS, as well as the effects of content evolution control and semantic association of resources.

Demonstration of resource evolution

The LC entitled “Overview of Corporate Universities” was selected as a typical case to depict the evolution process. The reasons for this choice were as follows: (1) Corporate universities are currently one of the “hottest” topics in the field of educational technology. As such, many learners may be engaged in improving the LC. (2) The LCS already contains a number of resources related to corporate universities. Therefore, building resource associations and observing the effect of resource associations would be easy.

The LC was created by user maxyang on 10 February 2012. By 6 March 2012, 26 formal versions had been generated. Version 1 contained basic contents related to corporate universities, including their origin, definition, and features. Altogether, 15 users edited the contents of the LC 68 times until Version 26 was produced. Specific revisions included adding new paragraphs, adjusting character styles, embedding photographs and activities, formatting paragraphs, and changing content. Figure 10 shows the result of the intelligent control of the evolution of the LC. At 16:58 on 29 February 2012, user noteexxx added content related to the subject of love. As this deviated significantly from the semantic gene of the LC, the system automatically rejected this edit and notified user noteexxx and the creator of this resource by e-mail. At 19:52 on 29 February 2012, user 李山 (Li Shan) added photographs of corporate universities. As 李山 was a user with a high trust level, the system automatically accepted his contribution.

By 6 March 2012, Version 26 had evolved significantly from the original version, with richer and more precise content. Some new topics and sections had been supplied, such as “Domestic and Foreign Well-known Corporate Universities” and “Reference Materials.” With the aid of Flex technology, the entire evolutionary process was visually tracked, as shown in Figure 7.

In addition to content evolution, resource associations also kept evolving from the outset. When the LC was created, it was not connected with any other resources. As contents evolved, Version 22 had connected with 8 LCs and 16 users. When the LC evolved to Version 26, it was associated with 12 LCs and 29 users. Figure 11 shows the changes in the numbers of resource associations at different evolutionary stages.

The general process of LC evolution is described as follows: (1) A user creates the LC and invites collaborators to edit contents. (2) More users are attracted to participate in viewing, editing, collecting, subscribing, and commenting on the LC. (3) Different versions of resource contents are generated by different users and checked automatically by the intelligent control program. (4) Content versions are constantly updated to achieve the evolution of resource contents. (5) As the contents evolve, the LC slowly and continually connects with other LCs and users through the dynamic semantic association program.

Revised Time	Version	Contributor	Auditor	Edit reason	CheckResult	CheckType	Note
2012-3-7 15:59	View	star	maxyang	图片居中	Accept	Manual	
2012-3-5 17:01	View	star		xxx插入参考资料;	Refuse	Auto	*Too big semantic differences*
2012-3-5 16:59	View	star		d插入参考资料;	Refuse	Auto	*Too big semantic differences*
2012-3-5 0:39	View	panna		更新内容，进化测试	Refuse	Auto	*Too big semantic differences*
2012-3-4 16:38	View	贾勇		添加内容;	Refuse	Auto	*Too big semantic differences*
2012-3-4 16:25	View	maxyang		添加学习活动;	Accept	Noneed	
2012-3-4 16:13	View	maxyang		新增参考资料;	Accept	Noneed	
2012-3-4 16:12	View	maxyang		添加企业大学建立步骤的[illegible]	Accept	Noneed	
2012-3-4 16:08	View	maxyang		[illegible]	[illegible]	Noneed	
2012-3-4 16:05	View	maxyang		调整结构	Accept	Noneed	
2012-3-1 10:38	View	noteexxx		xxxno change	Refuse	Auto	[illegible]
2012-3-1 10:32	View	邢惠娜		no change	Refuse	Auto	[illegible]
2012-3-1 10:28	View	邢惠娜	maxyang	删除首行缩进	Accept	Manual	
2012-2-29 21:37	View	maxyang		删除多余回车	Accept	Noneed	
2012-2-29 21:36	View	maxyang		删除多余回车	Accept	Noneed	
2012-2-29 19:53	View	李山		添加内容;	Accept	Auto	*Reliable operation of high credible user*
2012-2-29 19:52	View	李山	maxyang	插入图片;	Accept	Manual	*Reliable operation of high credible user*
2012-2-29 19:42	View	李山		添加内容;	Accept	Auto	
2012-2-29 16:58	View	noteexxx		添加爱情....	Refuse	Auto	*Too big semantic differences*
2012-2-29 16:49	View	maxyang		增加1条参考资料	Accept	Noneed	

This revision is accepted automatically.

This revision is rejected automatically.

Figure 10. Screenshot of intelligent control of the evolutionary process

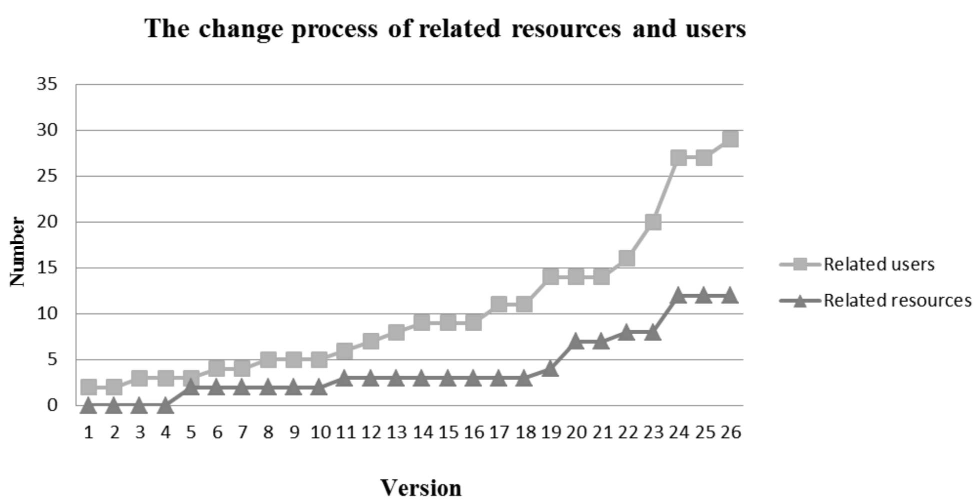

Figure 11. Process of association evolution of the LC

Effects of content evolution control

LCS has a log of resource evolution control, which can record the editorial time, reason, editor, audited result, and audit method. In addition, the audit method can be conducted manually or automatically. In a manual audit, the managers of the LCS conduct the checks. An automatic audit uses intelligent control methods.

We randomly selected the one month log data on content editing with a total of 3,938 edit records to check the effect of resource content evolution. A total of 497 edits required auditing, including 340 automatic reviews by the intelligent control procedures (70.98%) and 139 instances of manual auditing (29.02%). Figure 12 shows that the intelligent control program reduced the workload of content auditing by approximately seven tenths.

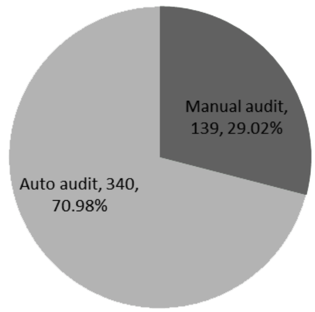

Figure 12. Percentage of manual audit and automatic audit

Moreover, we investigated the accuracy of the evolving content via intelligent control. One hundred fifty editing records were selected randomly by the program and stored in an Excel datasheet. A Ph.D. student was invited to log on to LCS to evaluate the accuracy of each record. Results showed that, of the 150 edits, 124 were evaluated as correct, yielding an accuracy rate of 82.67%. Focusing on the incorrect audits, we observed that most of these included multimedia contents, such as videos, animations, and pictures. These edits are difficult for the system to deal with. With the growth of registered LCS users, user groups will become larger and user actions may be more difficult to predict. The accuracy rate of intelligent control should be further analyzed, and the methods should be updated based on the results of studies such as this one.

Effects of resource association

The two ways to establish associations between LCs are manual editing and automatic building. The automatic building of resource association applies the methods of dynamic semantic association, including rule-based ratiocination, semantic gene-based calculating, and association rule-based data mining (affair of favorites or subscribes).

We developed a statistical function of resource association to monitor the evolution of the overall resource association. The result showed that the resource total (RT) number of resources in LCS was 3,775 and the association total (AT) of resources in LCS was 3,557. The average association degree (AAD; AT/RT) was 0.91, which means that each LC had an average semantic association of 0.91 with other LCs. This ratio implies that the semantic association between resources is not high.

Further analysis shows that 2,918 LCs have large differences in content and do not have semantic association with each other at all. Current resource associations in LCS are mainly distributed in 857 LCs (approximately 22.7%) that form a partial resource association net. The partial AAD number is 4.15. The distribution of resource association in LCS seems to follow a Pareto principle (80:20 rule), that is, over 80% of the total associations are concentrated in 20% of the resources.

Table 1. Statistics of resource association

	Manual	SG	RBR	ARM
Total number	570	783	2,156	672
Percentage	13.6%	18.7%	51.6%	16.1%

Table 1 shows the statistical result of resource association by different methods. The percentage of manual association is 13.6%, whereas that of automatic association is 86.4% (semantic gene-based calculating = 18.7%, rule-based ratiocination = 51.6%, association rule-based data mining = 16.1%). The results indicate that the associations built by users are few. Most of the associations are automatically created through dynamic semantic association, which plays an important role in the association evolution of the LCS resource group.

In addition, the accuracy of resource association also requires consideration. Considering the entire dynamic semantic association method, we examined the accuracy of the automatically established resource association. One hundred fifty resource association records were randomly selected by the program and stored in Excel datasheets. Two Ph.D. students were invited to evaluate the accuracy of 75 resource associations each and note their evaluations in the datasheets. A kappa consistency test revealed a result of 0.81 ($p < 0.01$). The result shows that the precision of automatically established associations is 71.33%. This finding indicates that the semantic association method has a high reliability and that most of the associations automatically established between resources are accurate. However, the monitoring result shows that the semantic gene-based similarity relationship calculation has disadvantages in performance efficiency and time demands. Therefore, the algorithm needs further testing and optimization.

Conclusion

In this study, we develop the RESS, which mainly addresses two issues in OKCs, namely, the intelligent control of content evolution and the semantic association. Integrated application of semantic features of resources based on ontology and a trust evaluation model is feasible for automatically checking content edits in the effort to achieve intelligent control of content evolution. Semantic association is an important part of resource evolution. Through relation calculations based on the semantic gene, rule-based reasoning, and association rule mining with semantic constraint, semantic associations could be dynamically built in association evolution.

This study generally makes two evident academic contributions to the area of OKCs. First, a novel RESS, which may shed light on the improvement of various OKCs, is proposed. Second, the integration of technologies consisting of semantics, trust evaluation, rule-based reasoning, and association rule mining creates an effective and novel method by which to implement the intelligent control of content evolution and the semantic association among resources. This approach can be applied to other OKCs that feature collaborative content editing. The limitation of this work is that no comparison is conducted between the effects of traditional methods and those of the new methods proposed in this study in relation to addressing problems of content evolution control and semantic association.

Moreover, the results have some implications for the applications of OKCs in educational scenarios. First, RESS can be used to promote the generation of course content in the activities of SGC. Second, teachers can build the semantic network of knowledge points of courses assisted by the semantic association technology proposed in this study. Third, when searching knowledge in OKCs, teachers and students should consider the credibility of learning resources and related users.

In the future, we plan to track the process of resource evolution continually in LCS, as well as to accumulate more data to assess the evolution effect. In addition, comparison studies will be conducted to assess the actual effectiveness of the methods proposed in this study.

Acknowledgments

This work was sponsored by a project funded by the National Social Science Foundation of China for the Youth Scholars of Education (Grant Number: CCA130134).

References

Adler, B. T., Chatterjee, K., De Alfaro, L., Faella, M., Pye, I., & Raman, V. (2008). Assigning trust to Wikipedia content. In *Proceedings of the 4th International Symposium on Wikis* (Article No. 26). New York, NY: ACM.

Babashzadeh, A., Daoud, M., & Huang, J. (2013). Using semantic-based association rule mining for improving clinical text retrieval. In G. Huang et al. (Eds.), *Health Information Science* (pp. 186-197). Heidelberg, Germany: Springer Berlin Heidelberg.

Begon, M., Townsend, C. R., & Harper, J. L. (2006). *Ecology, from individuals to ecosystems* (4th ed.). Oxford, UK: Wiley-Blackwell.

de Laat, P. B. (2014). From open-source software to Wikipedia: “Backgrounding” trust by collective monitoring and reputation tracking. *Ethics and Information Technology, 16*(2), 157-169.

Denko, M. K., Sun, T., & Woungang, I. (2011). Trust management in ubiquitous computing: A Bayesian approach. *Computer Communications, 34*(3), 398-406.

Dong, Z. D., Dong, Q., & Hao, C. L. (2007). Theoretical findings of HowNet. *Journal of Chinese Information Processing, 21*(4), 3-9.

Farrell, R. G., Liburd, S. D., & Thomas, J. C. (2004). Dynamic assembly of learning objects. In *Proceedings of the 13th International World Wide Web Conference on Alternate Track Papers & Posters* (pp.162-169). New York, NY: ACM press.

Javanmardi, S., Lopes, C., & Baldi, P. (2010). Modeling user reputation in wikis. *Statistical Analysis and Data Mining: The ASA Data Science Journal, 3*(2), 126-139.

Jin, X. Q. (2009). *Chinese text similarity algorithm research based on semantic similarity* (Unpublished doctoral dissertation). Zhejiang Industrial University, Zhejiang, China.

Jones, K., & Leonard, L. N. (2008). Trust in consumer-to-consumer electronic commerce. *Information and Management, 45*(2), 88-95.

Li, Y. Y. (2005). *Management and utilization of semantic learning resource* (Unpublished doctoral dissertation). Institute of Computing Technology, Chinese Academy of Science, Beijing, China.

Li, Y., Dong, M., & Huang, R. (2011). Designing collaborative e-learning environments based upon semantic Wiki: From Design models to application scenarios. *Educational Technology & Society, 14*(4), 49-63.

Li, Z. Y., & Wang, R.C. (2011). Dynamic secure trust management model for P2P e-commerce environments. *Journal on Communications, 32*(3), 50-59.

Lu, E. J. L., & Hsieh, C. J. (2009). A Relation metadata extension for SCORM content aggregation model. *Computer Standards & Interfaces*, *31,* 1028-1035.

Luo, Z. C., & Fu, Z. Z. (2008). The Impact analyze of some external forces on the Wikipedia order process. *Knowledge of Library and Information Science, 3,* 28-33. Retrieved from http://d.wanfangdata.com.cn/Periodical/tsqbzs200803007

Lv, Q. C., & Du, H. (2010). Design of learning management system based on knowledge map. *Silicon Valley,* 8, 57-58. Retrieved from http://d.wanfangdata.com.cn/Periodical/guig201008047

Maniu, S., Abdessalem, T., & Cautis, B. (2011). Casting a web of trust over Wikipedia: an interaction-based approach. In *Proceedings of the 20th International Conference Companion on World wide web* (pp. 87-88). New York, NY: ACM.

Miller, G. A. (1995). WordNet: A Lexical database for English. *Communications of the ACM, 38*(11), 39-41.

Moturu, S. T., & Liu, H. (2009). Evaluating the trustworthiness of Wikipedia articles through quality and credibility. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration* (Article No. 28). New York, NY: ACM.

Ray, S., & Chandra, N. (2012). Domain based ontology and automated text categorization based on improved term frequency - Inverse document frequency. *International Journal of Modern Education and Computer Science*, 4, 28-35.

Rowley, J., & Johnson, F. (2013). Understanding trust formation in digital information sources: The Case of Wikipedia. *Journal of Information Science, 39(*4), 494-508.

Sapienza P., & Zingales, L. (2012). A Trust crisis. *International Review of Finance, 12*(2), 123-131.

Seigenthaler, J. (2005). A False Wikipedia “Biography”. *USA Today*. Retrieved from http://usatoday30.usatoday.com/news/opinion/editorials/2005-11-29-wikipedia-edit_x.htm

Shao, G. P., Yu, S. G., & Guo, L. (2008). Promotion of e-learning resources management by semantic web. *Journal of Jiangsu Radio and TV University, 19*(5), 23-26. Retrieved from http://d.wanfangdata.com.cn/Periodical/jsgbdsdxxb200805006

Shi, Y. D., Zhang, S. Y., & Xiang, C. (2003). The Representation and association among knowledge points in network-based courses. *Journal of Zhejiang University (Engineering Science), 37*(5), 508-511. Retrieved from http://d.wanfangdata.com.cn/Periodical/zjdxxb-gx200305003

Siemens, G. (2005). Connectivism: A Learning theory for the digital age. Retrieved from http://er.dut.ac.za/handle/123456789/69

Ullrich, C. (2005). The Learning-resource-type is dead, long live the learning-resource-type. *Learning Objects and Learning Designs,1*(1), 7-15.

Vrandecic, D. (2009). Towards automatic content quality checks in semantic wikis. *Social semantic web: where web,* 2.

Wang, D. D. (2009). Quality assurance system under the mode of Wikipedia self-organization. *Library Theory and Practice*, 8, 21-24. Retrieved from http://d.wanfangdata.com.cn/Periodical/tsgllysj200908008

Wheeler, S., Yeomans, P., & Wheeler, D. (2008). The Good, the bad and the wiki: Evaluating student-generated content for collaborative learning. *British Journal of Educational Technology, 39*(6), 987-995.

Wu, S. Y., & Wu, Wu, Y. Y. (2010). The Similarity calculation of Chinese and English words based on WordNet. *Journal of Zhengzhou University (Natural Science Edition)*, (2), 66-69. Retrieved from http://caod.oriprobe.com/articles/29446892/Chinese_and_English_Word_Similarity_Measure_Based_on_Chinese_WordNet.htm

Yang, X. M. (2012). The Research of resource evolution in the ubiqutious learning environment (Unpublished doctoral dissertation). Beijing Normal University, Beijing, China.

Yang, X. M., & Yu, S. Q. (2011). The Construction of learning resource evolution model in ubiquitous learning environment. *China Educational Technology*, *9*, 80-85. Retrieved from http://d.wanfangdata.com.cn/Periodical/zgdhjy201109015

Yang, X. M., Yu, S. Q., & Zhang, F. (2013). The Design and implementation of learning resource's dynamic and semantic association. *China Educational Technology*, *1*, 70-75. Retrieved from http://d.wanfangdata.com.cn/Periodical/zgdhjy201301012

Yang, X. M., Qiu, Q.,Yu, S. Q., & Tahir, H. (2014). Designing a trust evaluation model for open-knowledge communities. *British Journal of Educational Tec*hnology*, 45*(5), 880-901.

Yu, S. Q., Yang, X. M., Cheng, G., & Wang M. J. (2015). From learning object to learning cell: A Resource organization model for ubiquitous learning. *Educational Technology & Society*, *18*(2), 206-224.

Zhang, Q., Zhang, X., & Liu, J. R. (2006). Semantic-based integration of SCORM learning resources. *Journal of Huazhong University of Science and Technology (Nature Science Edition), 42*(17), 296-298. Retrieved from http://d.wanfangdata.com.cn/Periodical/jsjgcyyy200617013

Zhao, D. H. (2006). The Scientific significance of evolution. *Seeking Truth, 33*(2), 27-31. Retrieved from http://d.g.wanfangdata.com.cn/Periodical_qsxk200602004.aspx