A Micro-Web Involving Learning Scenario Generation with Linked Open Data for Web-Based Investigative Learning

Yasuhiro Kakinuma(✉) and Akihiro Kashihara

Department of Informatics, Graduate School of Informatics and Engineering,
The University of Electro-Communications,
1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan
y.kakinuma@uec.ac.jp, akihiro.kashihara@inf.uec.ac.jp

Abstract. In investigative learning with Web resources, learners need to define the relationship between the topic and sub-topics through the creation of learning scenario. However, the Web space is not well-structured for learning. It is accordingly difficult for the learners to construct a learning scenario by themselves. The main purpose of this paper is to build up Micro Web inheriting the characteristics of the Web, which involves automatic generation of learning scenario by means of DBpedia Japanese as Linked Open Data (LOD) for Wikipedia. Such generated scenario enables an adaptive support for the learning scenario creation.

Keywords: Micro web · LOD · Learning scenario · Investigative learning

1 Introduction

On the Web, there are currently a great number and variety of existing information resources suitable for learning. These Web resources allow learners to investigate any topics to learn, which would promote learning in a wider, deeper, and timely way [1, 2]. Such investigative learning involves navigating the Web pages in the resources to construct knowledge about a topic investigated. In recent years, investigative learning with Web resources has been introduced into primary and secondary education classes [3].

In the investigative learning process with Web resources, the learners need to reconstruct the contents to be learned by themselves since most Web resources are not well structured for learning. In addition, these resources do not always provide a learning scenario indicating the topics and the sequences to be learned. It is accordingly necessary for the learners to create their own learning scenario while investigating learning process. In investigating a topic with Web resources, the learners are expected not only to integrate and construct knowledge learned at each resource, but also to find out related topics to be further investigated that can be viewed as the sub-topics. In this way, investigative learning process involves decomposing the topic into the sub-topics, which is regarded as creating a learning scenario. Such topic decomposition would make the investigative learning process more structured. However, it is not so easy for the learners to conduct the topic decomposition concurrent with the navigation and knowledge construction process.

P. Zaphiris and A. Ioannou (Eds.): LCT 2016, LNCS 9753, pp. 344–355, 2016.
DOI: 10.1007/978-3-319-39483-1_32

We have accordingly developed a system called interactive learning scenario builder on iPad (iLSB-tablet for short), which scaffolds learning scenario creation. iLSB-tablet follows the model of Web-based investigative learning we have built up. This model expects learners to search for Web resources with search engine to navigate the Web pages and construct knowledge and to create a learning scenario. The results of the case study with iLSB-tablet suggest that the learning scenario creation can be promoted while investigative learning process, and also that it is not so easy for the learners to create more proper scenario by themselves.

The main issue addressed in this paper is how to help learners create a proper learning scenario. Our approach to this issue is to design a micro-Web, which is a micro-world for the Web. The micro-Web provides learners with a scaffold for their investigative learning where they can be supported by means of a scenario suitable for learning a topic. It also prepares a limited number and kinds of Web resources related to the topic. A challenge in this work is how to generate such learning scenario from the unstructured hyperspace provided on the Web.

In this paper, we propose a promising method of the learning scenario generation with Linked Open Data (LOD for short) of Wikipedia, which is provided by DBpedia Japanese. It is possible to use SPARQL queries to get the data (or keywords) linked to a topic, which are found in the Wikipedia pages. The procedure of generating a learning scenario involves the following four steps. In the first step, an instructor is expected to select an initial topic and its related topics to provide a topic set involving the related topics. The second step is to generate a SPARQL query with the initial topic to obtain the related data that includes a lot of related keywords. The third step is to extract the topics commonly included in the obtained data and the topic set, which correspond to the sub-topics. In the fourth step, the part-of link between the initial topic and each sub-topic is generated. Generating the SPARQL query with each sub-topic and repeating from the second step to the fourth step, a learning scenario is automatically generated.

This paper also describes a case study with the learning scenario generation, which involves several trials of generating scenarios for learning different topics. The results would suggest its effectiveness. In addition, we will discusses how to build up the micro-Web involving the learning scenario generation in detail, and how to use iLSB-tablet on the Micro-Web.

2 Model of Web-Based Investigative Learning

Let us first introduce the model of Web-based investigative learning proposed in our previous work [4]. As shown in Fig. 1, this model includes three phases, which are phase of search for Web resources, navigational learning phase, and phase of learning scenario building.

In the phase of search for Web resources, learners would use the search engine such as Google with a keyword (called topic keyword) that represents an initial topic, and explore the Web resources. In the navigational learning phase, the learners would navigate across the Web resources to investigate the topic and learn the contents of the

Fig. 1. Model of Web-based investigative learning

Web resources. Such navigational learning involves constructing knowledge that they have learned from the Web resources.

In the phase of learning scenario building, the learners would find out related topics from their knowledge constructed, which can be considered as sub-topics of the initial topic. They would build the learning scenario by decomposing the topic into sub-topics and relating the topic and the sub-topics. The learning scenario building is continued until the topic decomposition does not occur anymore.

3 Learning Scenario Generation for Micro-Web

This paper proposes an automatic method to generate a learning scenario for the Micro-Web. Toward building the Micro-Web from the unstructured Web, we use LOD of Wikipedia presented by DBpedia Japanese.

3.1 Micro-Web

The Micro-Web is viewed as a micro-world [5] for the Web. It provides learners with a scaffold for their investigative learning where they can be supported by means of a scenario suitable for learning a topic. The Micro-Web could also include unstructured and unreliable Web resources in addition to the ones useful for learning.

In general, the learners have difficulties in finding out the resources necessary for investigating the topic because of the enormous ones on the Web, and often finish learning with insufficient knowledge. In contrast, the Micro-Web allows an instructor

to beforehand restrict the number of Web resources. The instructor is also allowed to choose the sub-topics to assume the learning scenario that the learners should follow in investigating the topic. The learners are accordingly allowed to construct their knowledge appropriate to the topic.

In this way, the Micro-Web provides learners with a structured micro-world for the Web, which includes the Web resources useful for their investigative learning process.

3.2 Learning Scenario Generation with LOD of Wikipedia

Our approach to automatic generation of learning scenario is to use LOD of Wikipedia provided by DBpedia.

– Linked Open Data (LOD)

LOD provides the method to publish and share the data on the Web, and link the data each other to represent Web information. It is generally under the open license.

Representative LOD services include DBpedia and Freebase [6]. In this work, we use DBpedia Japanese [7] presented by the national information science research institute [8]. This LOD processes the data of Japanese Wikipedia such as infobox, category information, images, geographical coordinate, and links to the outside Web pages. Such information is structured and converted into linked data, which is also represented with RDF (Resource Description Framework). RDF is a machine-readable data model for describing the Web information. In RDF, data is described by means of triplet form such as subject-predicate-object. More specifically, it means resource (as subject), property (as predicate), and resources (as object). The RDF triplets describing the data in LOD is obtained with query language SPARQL [9].

Figure 2 shows an example of RDF triplet description. This triplet describes a property "wikiPageWikiLink" of DBpedia Japanese as the predicate in a resource of "Global Warming" as the subject, and describes a resource of "Kyoto Protocol" as the object. The property "wikiPageWikiLink" means the link from a certain Wikipage to different one [10]. In this triplet, the predicate means that there is a link from the resource "Global Warming" to the resource "Kyoto Protocol". All resources dealt with by DBpedia Japanese have the corresponding Wikipedia pages. As shown in Fig. 2, we can search many resources including "Global Warming" (such as greenhouse gas, carbon dioxide, etc.) when we put the query about the "Global Warming" into the SPARQL endpoint [11] to get the information of relevant words.

– Procedure of the learning scenario automatic generation

Let us next explain the basic steps for the automatic generation of learning scenario from SPARQL query.

1. An instructor is expected to select an initial topic and its related topics to be learned, and to provide a topic set involving the related topics.
2. A SPARQL query with the initial topic is generated and sent to DBpedia Japanese. The related keywords are then obtained from the LOD.

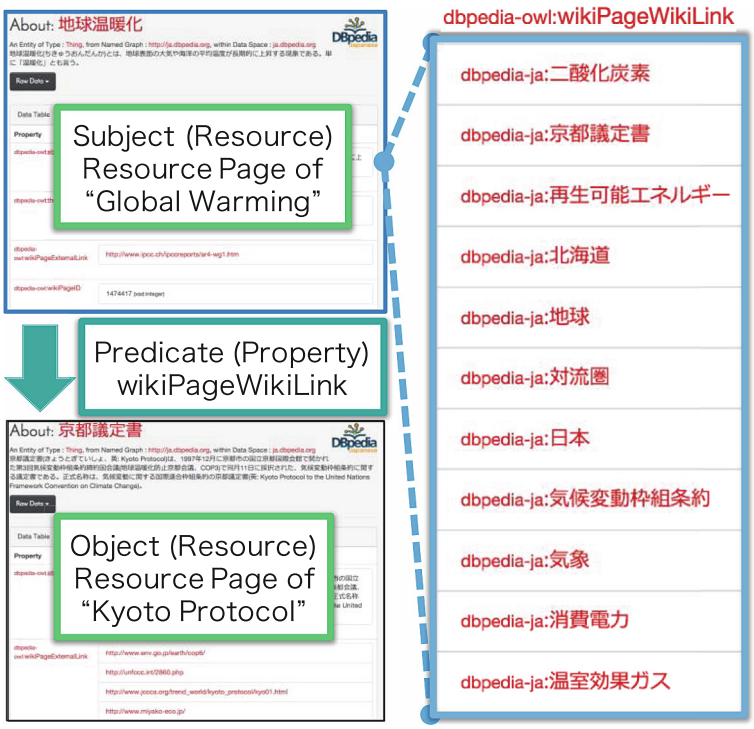

Fig. 2. Example of RDF description

3. The topics commonly included in the obtained keywords and the topic set are extracted as the sub-topics.
4. The part-of link between the initial topic and each sub-topic is generated.

The SPARQL query with each sub-topic is continuously generated, and the process from the second step to the fourth step is repeated. Then, a learning scenario is automatically generated.

As shown in Fig. 3, for example, consider the case where the instructor selects “Global Warming” as an initial topic. He/she also prepares “Greenhouse Effect Gas”, “Kyoto Protocol”, “Carbon Dioxide”, etc. as related topics that the learners should investigate about the “Global Warming”. These related topics will be included in an automatic generated scenario. Then, a SPARQL query about “Global Warming” is generated, and the Wikipedia resources linked from “Global Warming” are obtained as the related keywords from the LOD by means of the property “wikiPageWikiLink”. If there are common keywords included in the related topics and obtained keywords, they are extracted as the sub-topics. In Fig. 3, “Greenhouse Effect Gas” and “Abnormal Weather” are extracted as the sub-topics of “Global Warming”. In the same manner, the next SPARQL queries including “Greenhouse Effect Gas” and “Abnormal Weather” are generated and the related keywords are obtained from the LOD. As shown in Fig. 3, the sub-topics for “Greenhouse Effect Gas” are detected as “Kyoto Protocol”

and “Carbon Dioxide”. Repeating these processes, we can obtain a topic tree structure representing a learning scenario about “global warming”.

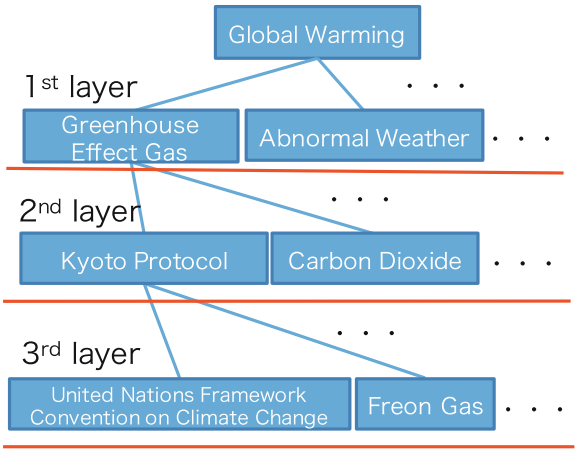

Fig. 3. Example of topic keywords acquisition every hierarchy

– Framework of the Micro-Web

Figure 4 shows a framework of the Micro-Web involving the automatic learning scenario generation. In this framework, an instructor selects a topic and related topics necessary for the automatic scenario generation. The learning scenario is generated with the LOD provided by the DBpedia Japanese.

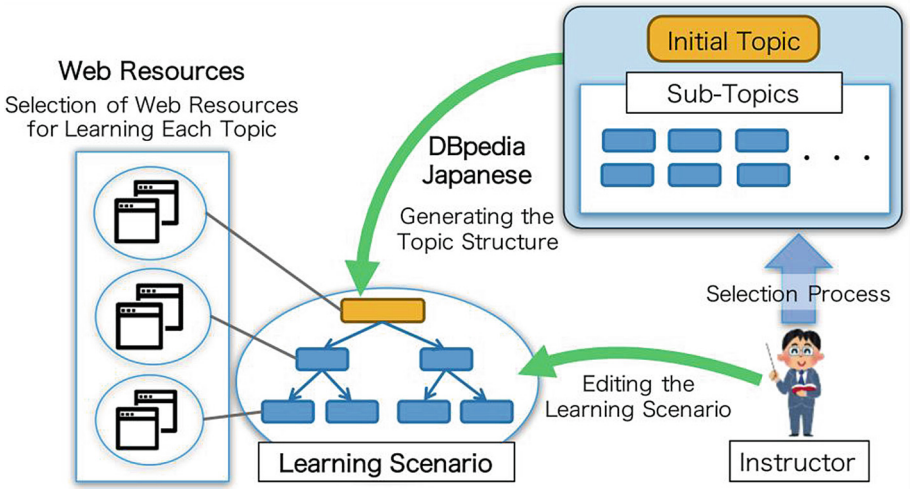

Fig. 4. Framework of the Micro-Web

If the generated scenario structure does not satisfy the instructor, he/she could edit it to re-generate the appropriate scenario. After the scenario is decided, the instructor could integrate the Web resources, which are contributory to investigative learning with all the topics included in the scenario.

3.3 Topic Keyword Extraction by SPARQL Query

Figure 5 shows an example of a query that queries the SPARQL endpoint of DBpedia. The resource (subject) in this example is represented as "dbp:global warming", and the property (predicate) corresponds to "dbp-owl:wikiPageWikiLink". The resource (object) to be extracted is represented as "?thing1", which is a variable in the query. The results of the query are stored in this variable. The "prefix" in the query is used for shortening the description of the URI. The names of the resources stored in "?thing1" are shown in the property "rdfs:label". The resources is also stored in the variable "?thing2".

Default Data Set Name (Graph IRI)
http://ja.dbpedia.org

SPARQL Query

Query Text

```
PREFIX dbp: <http://ja.dbpedia.org/resource/>
PREFIX dbp-owl: <http://dbpedia.org/ontology/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT *
WHERE {
        dbp:地球温暖化 dbp-owl:wikiPageWikiLink ?thing1.
        ?thing1 rdfs:label ?thing2.
}
```

Fig. 5. An Example of SPARQL query

Figure 6 shows an output of the query in Fig. 5, which includes the URLs obtained. After the SPARQL query is inputted, the JSON format of the keyword data is extracted as the output.

Figures 7 and 8 show examples of learning scenario generation. In Fig. 7, 16 topics out of 20 related ones are extracted as the sub-topics from the initial topic "Global Warming". 12 topics of these 16 topics are also extracted at the first layer, and there are only the 3 sub-topics that have the sub-topics at the second or third layers. The remaining 4 related topics are not extracted in the scenario.

In Fig. 7, most related topics are extracted at the first layer because there are the links between these topics and "Global Warming" in Wikipedia. In order to avoid such topic decomposition, we need to update the learning scenario generation method. We currently consider the update as follows. First, we will beforehand classify the related topics into several groups. We will then refine the automatic scenario generation method so that one keyword is extracted from each group at the same layer.

thing1	thing2
http://ja.dbpedia.org/resource/11世紀	"11世紀"@ja
http://ja.dbpedia.org/resource/12月29日	"12月29日"@ja
http://ja.dbpedia.org/resource/1850年	"1850年"@ja
http://ja.dbpedia.org/resource/1860年	"1860年"@ja
http://ja.dbpedia.org/resource/1896年	"1896年"@ja
http://ja.dbpedia.org/resource/1900年	"1900年"@ja
http://ja.dbpedia.org/resource/1979年	"1979年"@ja
http://ja.dbpedia.org/resource/1988年	"1988年"@ja
http://ja.dbpedia.org/resource/1992年	"1992年"@ja
http://ja.dbpedia.org/resource/19世紀	"19世紀"@ja
http://ja.dbpedia.org/resource/1世紀	"1世紀"@ja
http://ja.dbpedia.org/resource/2001年	"2001年"@ja
http://ja.dbpedia.org/resource/20世紀	"20世紀"@ja
http://ja.dbpedia.org/resource/7月6日	"7月6日"@ja
http://ja.dbpedia.org/resource/9月29日	"9月29日"@ja
http://ja.dbpedia.org/resource/アル・ゴア	"アル・ゴア"@ja
http://ja.dbpedia.org/resource/メタン	"メタン"@ja
http://ja.dbpedia.org/resource/二酸化炭素	"二酸化炭素"@ja
http://ja.dbpedia.org/resource/京都議定書	"京都議定書"@ja
http://ja.dbpedia.org/resource/再生可能エネルギー	"再生可能エネルギー"@ja

Fig. 6. Example of SPARQL query output (HTML Format)

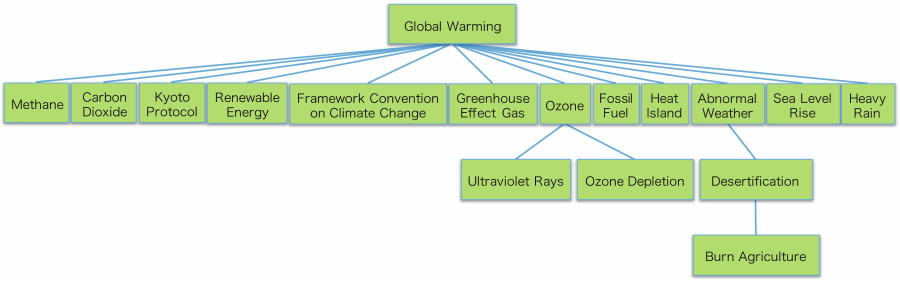

Fig. 7. Example of learning scenario generation

In grouping the related topics about “Global Warming”, we introduce three viewpoints which are “Cause”, “Effect”, and “Countermeasure”. Following these viewpoints, we classified the related topics into three groups and simulate the automatic scenario generation. Figure 8 shows the learning scenario to be generated. We do not currently implement the updated method. In future, we need to refine the learning scenario automatic generation method.

On the other hand, there is another approach to generating a structured scenario with the categories used in Wikipedia, which classifies the Wikipedia articles by fields.

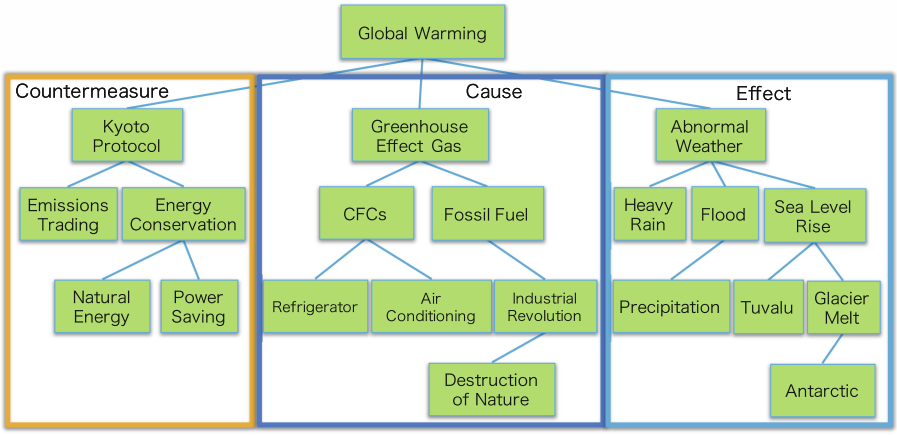

Fig. 8. Example of learning scenario by grouping related topics

In this approach, the sub-topics to be investigated are selected from the categories. Figure 9 shows an example of a query to the SPARQL endpoint. This asks what category belongs to "global warming" indicated by "category-ja:" Fig. 10 shows the result, which includes the categories belonging to "global warming".

Default Data Set Name (Graph IRI)
http://ja.dbpedia.org

SPARQL Query

Query Text

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX category-ja: <http://ja.dbpedia.org/resource/Category:>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>

SELECT *
WHERE {
        category-ja:地球温暖化 ^skos:broader ?thing1.
        ?thing1 rdfs:label ?thing2.
}
```

Fig. 9. Example of query to acquire the lower category

Figure 11 shows a learning scenario generated with the category names extracted from the Wikipedia. In this example, the category name of "global warming" is given as the initial topic. From the "global warming" category, the category names such as "fossil fuel", "greenhouse gas", etc. are extracted. As we have confirmed, there are currently some cases where important topics such as "Kyoto Protocol" in learning "global warming" are missing in this method. In order to resolve such problems, it is necessary to use the property "dbp-owl:wikiPageWikiLink" as discussed in Fig. 5.

thing1	thing2
http://ja.dbpedia.org/resource/Category:化石燃料	"化石燃料"@ja
http://ja.dbpedia.org/resource/Category:温室効果ガス	"温室効果ガス"@ja
http://ja.dbpedia.org/resource/Category:気候変動に関する政策	"気候変動に関する政策"@ja
http://ja.dbpedia.org/resource/Category:気候変動関連組織	"気候変動関連組織"@ja
http://ja.dbpedia.org/resource/Category:気候変動の影響	"気候変動の影響"@ja
http://ja.dbpedia.org/resource/Category:反地球温暖化論	"反地球温暖化論"@ja
http://ja.dbpedia.org/resource/Category:気候変動の原因	"気候変動の原因"@ja
http://ja.dbpedia.org/resource/Category:気候変動防止	"気候変動防止"@ja
http://ja.dbpedia.org/resource/Category:気候変動防止活動	"気候変動防止活動"@ja

Fig. 10. Example of category output results of SPARQL query

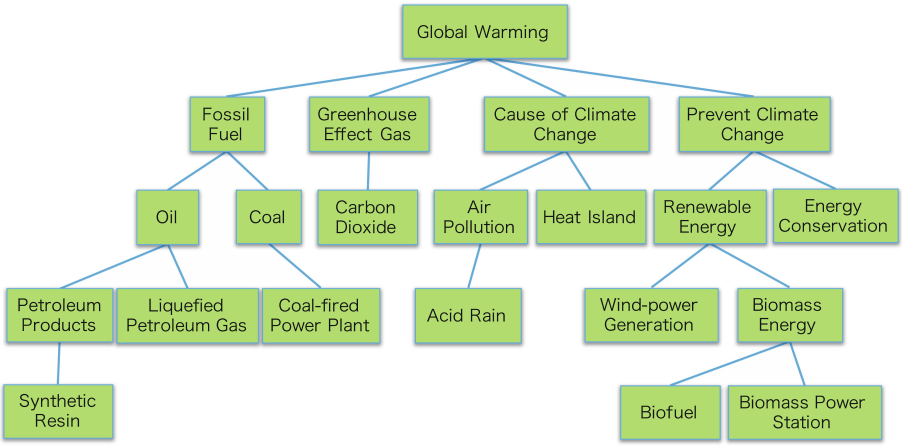

Fig. 11. Example of learning scenario by categorizing

4 iLSB-Tablet

In this work, we have developed interactive Learning Scenario Builder on iPad (iLSB-tablet for short), which is a Web-based investigative learning support system [12]. Figure 2 shows the user interface of iLSB-tablet.

This system follows the model of Web-based investigative learning as shown in Fig. 1, and provides learners with the functions that are search engine, keyword repository for storing keywords representing the contents learned about the topics, and learning scenario builder. In addition, the system presents a list of attributes representing the relationships between topics, which functions as a scaffold for the learners to pick the important topic keywords up from the Web pages or the keyword repository

[12]. This allows them to decompose the topic into the sub-topics to generate the meaningful scenario.

In the future, we will plan to develop a function for comparing the learning scenario generated by the learners with the scenario automatically generated by the system. Presenting the difference between the scenarios promotes the investigative learning process. The comparison function could also provide some foundation for adaptation of investigative learning on the Micro-Web (Fig. 12).

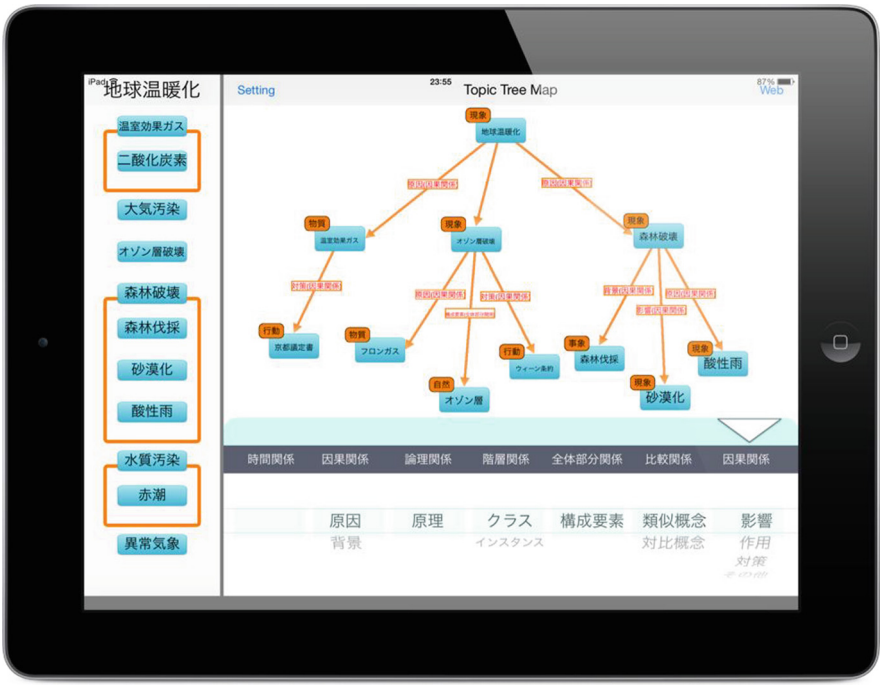

Fig. 12. User interface of iLSB-tablet

5 Conclusion

In this paper, we have proposed the Micro-Web and learning scenario generation with LOD provided by DBpedia Japanese. As future work, we need to implement iLSB-tablet with automatic scenario generation and with the comparison function for adaptive investigative learning. We will also evaluate the proposed method in detail and refine it base on the results.

Acknowledgment. This work is in part supported by JSPS KAKENHI Grand Number 26282047.

References

1. Kashihara, A., Hasegawa, S.: A model of meta-learning for web-based navigational learning. Int. J. Adv. Technol. Learn. **2**(4), 198–206 (2005)
2. Land, S.M.: Cognitive requirements for learning open-ended learning environments. Educ. Technol. Res. Dev. **48**(3), 61–78 (2000)
3. Ministry of Education, Culture, Sports, Science and Technology (MEXT),“ the Period for Integrated Studies” (in Japanese)
4. Kashihara, A., Akiyama, N.: Learner-created scenario for investigative learning with web resources. In: Lane, H., Yacef, K., Mostow, J., Pavlik, P. (eds.) AIED 2013. LNCS, vol. 7926, pp. 700–703. Springer, Heidelberg (2013)
5. Matsubara, Y.: Simulation Virtual reality system, Knowledge base, The Institute of Electronics, Information and Communication Engineers, group S3 part 11 “Education support system”. In: Education/Learning Support System, Chap. 4, pp. 17−22 (2011). http://www.ieice-hbkb.org/files/S3/S3gun_11hen_04.pdf
6. Freebase. https://www.freebase.com
7. DBpedia Japanese. http://ja.dbpedia.org
8. National Institute of Informatics. http://www.nii.ac.jp
9. For RDF query language SPARQL. http://www.asahi-net.or.jp/～ax2s-kmtn/internet/rdf/rdf-sparql-query.html
10. About: Link from a Wikipage to another Wikipage – DBpedia. http://dbpedia.org/ontology/wikiPageWikiLink
11. SPARQL Endpoint: Virtuoso SPARQL Query Editor. http://ja.dbpedia.org/sparql
12. Kinoshita, K., Kashihara, A.: Scaffolding topic decomposition in investigative learning with web resources. In: Proceedings of ICCE 2014, Nara, Japan (2014)