Computers & Geosciences 37 (2011) 1602–1615

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

A SKOS-based multilingual thesaurus of geological time scale for interoperability of online geological maps

Xiaogang Ma [a,*], Emmanuel John M. Carranza [a], Chonglong Wu [b,c], Freek D. van der Meer [a], Gang Liu [b]

[a] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Hengelosestraat 99, 7514 AE Enschede, The Netherlands
[b] School of Computer Science, China University of Geosciences, Wuhan 430074, Hubei, China
[c] Three Gorges Research Center for Geo-hazard (TGRC), Ministry of Education, China University of Geosciences, Wuhan 430074, Hubei, China

ARTICLE INFO

Article history:
Received 20 September 2010
Received in revised form
7 February 2011
Accepted 19 February 2011
Available online 17 March 2011

Keywords:
Geodata interoperability
Semantic Web
Ontology
Web Map Service (WMS)
Information retrieval
Term recognition

ABSTRACT

The usefulness of online geological maps is hindered by linguistic barriers. Multilingual geoscience thesauri alleviate linguistic barriers of geological maps. However, the benefits of multilingual geoscience thesauri for online geological maps are less studied. In this regard, we developed a multilingual thesaurus of geological time scale (GTS) to alleviate linguistic barriers of GTS records among online geological maps. We extended the Simple Knowledge Organization System (SKOS) model to represent the ordinal hierarchical structure of GTS terms. We collected GTS terms in seven languages and encoded them into a thesaurus by using the extended SKOS model. We implemented methods of characteristic-oriented term retrieval in JavaScript programs for accessing Web Map Services (WMS), recognizing GTS terms, and making translations. With the developed thesaurus and programs, we set up a pilot system to test recognitions and translations of GTS terms in online geological maps. Results of this pilot system proved the accuracy of the developed thesaurus and the functionality of the developed programs. Therefore, with proper deployments, SKOS-based multilingual geoscience thesauri can be functional for alleviating linguistic barriers among online geological maps and, thus, improving their interoperability.

1. Introduction

Linguistic barrier is a long-term challenge for the interoperability of geodata and retrieval of geoinformation (Asch and Jackson, 2006; Gravesteijn and Rassam, 1990; Laxton et al., 2010; Lloyd, 1973). Users of geological maps have been facing that challenge since this type of geodata has evolved. With increasing internationalization and globalization of geological scientific and technological works (e.g., de Mulder et al., 2006; Jackson, 2007), overcoming that challenge has become an important issue in sharing of geodata and/or geoinformation. Most geological maps are produced by governmental organizations; thus, they are encoded in official languages of their producers. If users cannot read the languages of a geological map, then it is hard for them either to understand the meaning of that map or to use that map efficiently. Recently, some digital geological maps have been published in bilingual formats (e.g., the 1:200,000 Geological Map of Japan published in Japanese and English (GSJ-AIST, 2009)) and multilingual formats (e.g., the 1:5,500,000 Geological Map of South America published in Spanish, Portuguese, and English (CGMW et al., 2003)) to alleviate linguistic barriers for international users. However, the number of languages used in these maps is still limited and several other geological maps remain in monolingual formats. Consequently, the interoperability of most geological maps is precluded or hindered.

Since the past decades, researchers coordinated by the CGI-IUGS[1] and its predecessors have been attempting to alleviate linguistic barriers of geological maps by developing multilingual geoscience thesauri. Earlier outputs of their works include the published 1st and 2nd editions of Multilingual Thesaurus of Geosciences (or MTG) (Gravesteijn et al., 1995; Rassam et al., 1988). The 2nd edition includes 5823 terms in English (as the basic reference), French, German, Italian, Russian, and Spanish.[2] Another recently published output is the Asian Multilingual Thesaurus of Geosciences (or AMTG) (CCOP and CIFEG, 2006), which includes 5867 terms in English (as the basic reference), Khmer, Chinese, French, Indonesian, Japanese, Korean, Lao, Malaysian, Thai, and Vietnamese.[3] These thesauri help users understand and use geological maps in foreign languages. However, the MTG contains some geoscience terms that are "inconsistent,

* Corresponding author. Tel.: +31 53 4874227; fax: +31 53 4874336.
E-mail addresses: xiaogang@itc.nl, xgmachina@gmail.com (X. Ma).

[1] Commission for the Management and Application of Geoscience Information of the International Union of Geological Sciences. http://www.cgi-iugs.org [accessed February 07, 2011].

[2] The online version of MTG also includes Finnish and Swedish. http://en.gtk.fi/Geoinfo/Library/multhes.html [accessed February 07, 2011].

[3] http://www.ccop.or.th/download/pub/AMTG_2006.pdf [accessed February 07, 2011].

0098-3004/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cageo.2011.02.011

incomplete and inaccurate" (Asch and Jackson, 2006), and so its applications are limited. The newer AMTG also contains some "inconsistent, incomplete and inaccurate" terms, and its applications have not been fully demonstrated yet. Thus, it is vital to further study how to make more useful multilingual geoscience thesauri.

Rapidly evolving web technologies pave the way for development of platforms for sharing geological maps to the international community, and for developing and applying multilingual geoscience thesauri. The OGC web service standards (e.g., WMS,[4] WFS,[5] and WCS[6]) enable the flow of geodata more open and faster through the World Wide Web (Peng and Tsou, 2003). By using these web services, organizations or individuals can publish geological maps online easily. For example, through the OneGeology project (Jackson, 2007; Jackson and Wyborn, 2008), 116 countries have agreed to share geological maps by the middle of 2010, and 50 of them have already provided WMS or WFS of their national or regional geological maps.[7] Meanwhile, extensive studies related to the W3C-proposed Semantic Web[8] have been addressing the essentiality of ontologies for formal and common representations of subject domain knowledge (e.g., Antoniou et al., 2005; Davies et al., 2003; Garcia-Sanchez et al., 2009). Developments of geoscience thesauri, as basic elements for building geoscience ontologies and representing geoscience knowledge, have increasingly become one of the foci of studies in the context of the Semantic Web (e.g., Buccella et al., 2009; Deliiska, 2007; Raskin and Pan, 2005; Smits and Friis-Christensen, 2007).

Researchers in the Geoscience Concept Definitions Task Group[9] of the CGI-IUGS are currently working on multilingual geoscience thesauri with the Simple Knowledge Organization System (SKOS),[10] a standard recommended by W3C. This effort is consistent with works of MTG and AMTG, and it aims to make significant improvements to the online applications of geoscience thesauri. The MTG and AMTG classify geoscience terms by subject domains, but terms classified into each subject domain are arranged alphabetically without definitions. The current work of the Geoscience Concept Definitions Task Group is compatible with the Semantic Web and has great potential in applications with online geological maps, such as those in the OneGeology project. Although impressive progress has been made by now, the work on SKOS-based geoscience thesauri by the Geoscience Concept Definitions Task Group of the CGI-IUGS is still ongoing, and methods for developing SKOS-based geoscience thesauri still require further practical testing and discussion. Meanwhile, online services and/or applications based on SKOS-based geoscience thesauri are still rare.

The study presented in this paper aims to develop a SKOS-based multilingual thesaurus of geological time scale (MLTGTS) for alleviating linguistic barriers of geological time scale (GTS) records among online geological maps. The contributions of this study are threefold. First, to extend the SKOS model to build a more semantically expressive structure for the subject domain of GTS: This would motivate building thesauri of other subject domains in geosciences. Second, to maintain a MLTGTS and use it with developed JavaScript programs to recognize and translate GTS terms in online geological maps: The approach of characteristic-oriented term retrieval implemented in the JavaScript programs is effective for recognizing GTS terms from records in geological maps. Third, to package the first and second contributions into a novel methodology for improving the interoperability of online geological maps in the context of the Semantic Web: With functions for online recognition and translation of geoscience terms, massive monolingual geological maps can be published online directly and users can access and use them, although they cannot read their original languages.

The remainder of this paper describes the study in detail. Section 2 describes the collection and encoding of multilingual GTS terms into the MLTGTS with an extended SKOS model. Section 3 describes a workflow implemented in JavaScript programs for recognizing and translating GTS terms in online geological maps. Section 4 presents a pilot system using the developed MLTGTS and JavaScript programs, and evaluates the results of recognizing and translating GTS terms in online geological maps with some examples. Section 5 compares this study with similar studies, summarizes lessons learned, and proposes directions for future studies. Finally, Section 6 highlights the findings of this study.

2. SKOS-based multilingual thesaurus of geological time scale

2.1. Addressing the insufficiency of SKOS in the context of the Semantic Web

By using ontologies in the Semantic Web, meanings of concepts and relationships between concepts are made accessible as the material in which certain concepts appear (Berners-Lee et al., 2001). This paradigm is also supported by recent studies related to the Geospatial Semantic Web[11] (Bishr, 2006; Yue et al., 2009; Zhang et al., 2010). Ontologies in computer science are valuable functions because they are derived from shared conceptualizations of domain knowledge (Gruber, 1995). Thesauri are regarded as a necessary foundation for building ontologies in computer science (Gruber, 1995; Guarino, 1997, 1998). Professional (e.g., geoscience) terms in a thesaurus may refer to the same real-world features in a subject domain, as an ontology does. However, unlike a precise conceptualization (i.e., detailed semantics) in an ontology, a thesaurus is simpler in definitions of meanings and relationships of terms (i.e., concise semantics) and, thus, it leads to a simple organizational structure (Gilchrist, 2003). For example, the MTG and AMTG arrange geoscience terms alphabetically, and each term is tagged with a label indicating its subject domain in geosciences.

To promote functions for indexing and navigating resources on the Web, it would be useful to encode thesauri in Web-compatible formats. Similar to OWL's[12] role in editing ontologies, the SKOS can be used for encoding thesauri in the context of the Semantic Web. SKOS is a common data model based on the RDF,[13] which in turn is a standard recommended by W3C. In the SKOS model (Table 1),[14] there are predefined object properties for defining relationships between concepts and datatype properties for defining differentiating attributes (or qualities) of concepts. These properties (e.g., "skos:broader," "skos:narrower," "skos:related," "skos:prefLabel," "skos:altLabel") let users set up hierarchical and associative relationships between terms within a thesaurus, and assign essential attributes (e.g., multilingual labels) to each term (e.g., Pastor-Sanchez et al., 2009). For example, Fig. 1 shows a GTS concept "Lower_Triassic" defined with the pure SKOS model.

[4] Web Map Service. http://www.opengeospatial.org/standards/wms [accessed February 07, 2011].

[5] Web Feature Service. http://www.opengeospatial.org/standards/wfs [accessed February 07, 2011].

[6] Web Coverage Service. http://www.opengeospatial.org/standards/wcs [accessed February 07, 2011].

[7] http://www.onegeology.org/participants/app/1gCountries.cfc?method=viewCountryStatus [accessed February 07, 2011].

[8] http://www.w3.org/standards/semanticweb [accessed February 07, 2011].

[9] https://www.seegrid.csiro.au/twiki/bin/view/CGIModel/ConceptDefinitionsTG [accessed February 07, 2011].

[10] http://www.w3.org/2004/02/skos [accessed February 07, 2011].

[11] http://www.opengeospatial.org/projects/initiatives/gswie [accessed February 07, 2011].

[12] Web Ontology Language. http://www.w3.org/TR/owl-guide [accessed February 07, 2011].

[13] Resource Description Framework. http://www.w3.org/RDF [accessed February 07, 2011].

[14] SKOS Reference. W3C Recommendation 18 August 2009. http://www.w3.org/TR/2009/REC-skos-reference-20090818 [accessed February 07, 2011].

Table 1
Object and datatype properties in the SKOS model.

Object property	Meaning	Datatype property	Meaning
skos:broadMatch	Has broader match	skos:altLabel	Alternative label
skos:broader	Has broader	skos:changeNote	Change note
skos:broaderTransitive	Has broader transitive	skos:definition	Definition
skos:closeMatch	Has close match	skos:editorialNote	Editorial note
skos:exactMatch	Has exact match	skos:example	Example
skos:hasTopConcept	Has top concept	skos:hiddenLabel	Hidden label
skos:inScheme	Is in scheme	skos:historyNote	History note
skos:mappingRelation	Is in mapping relation with	skos:notation	Notation
skos:member	Has member	skos:note	Note
skos:memberList	Has member list	skos:prefLabel	Preferred label
skos:narrowMatch	Has narrower match	skos:scopeNote	Scope note
skos:narrower	Has narrower		
skos:narrowerTransitive	Has narrower transitive		
skos:related	Has related		
skos:relatedMatch	Has related match		
skos:semanticRelation	Is in semantic relation with		
skos:topConceptOf	Is top concept in scheme		

a

```
<skos:Concept rdf:ID="Lower_Triassic">
  <skos:definition>Lower/Early Triassic Series/Epoch; 251.0±0.4—~245.9 Ma</skos:definition>
  <skos:prefLabel xml:lang="cn">下三叠统 </skos:prefLabel>
  <skos:prefLabel xml:lang="de">Untertrias</skos:prefLabel>
  <skos:prefLabel xml:lang="en">Lower Triassic</skos:prefLabel>
  <skos:prefLabel xml:lang="es">Triásico Inferior</skos:prefLabel>
  <skos:prefLabel xml:lang="fr">Trias Inférieur</skos:prefLabel>
  <skos:prefLabel xml:lang="jp">下部三畳系 </skos:prefLabel>
  <skos:prefLabel xml:lang="nl">Onder Trias</skos:prefLabel>
  <skos:altLabel xml:lang="en">Early Triassic</skos:altLabel>
  <skos:altLabel xml:lang="jp"> 三畳紀前期 </skos:altLabel>
  <skos:altLabel xml:lang="jp">前期三畳紀</skos:altLabel>
  <skos:altLabel xml:lang="jp"> 三畳系下部 </skos:altLabel>
  <skos:altLabel xml:lang="cn">早三叠世 </skos:altLabel>
  <skos:altLabel xml:lang="nl">Vroeg Trias</skos:altLabel>
  <skos:altLabel xml:lang="es">Triasico Inferior</skos:altLabel>
  <skos:broader rdf:resource="#Triassic"/>
  <skos:narrower rdf:resource="#Olenekian"/>
  <skos:narrower rdf:resource="#Induan"/>
  <skos:related rdf:resource="#Middle_Triassic"/>
  <skos:related rdf:resource="#Lopingian"/>
  <skos:inScheme rdf:resource="urn:ICS:International_Stratigraphic_Chart:2010"/>
</skos:Concept>
```

b

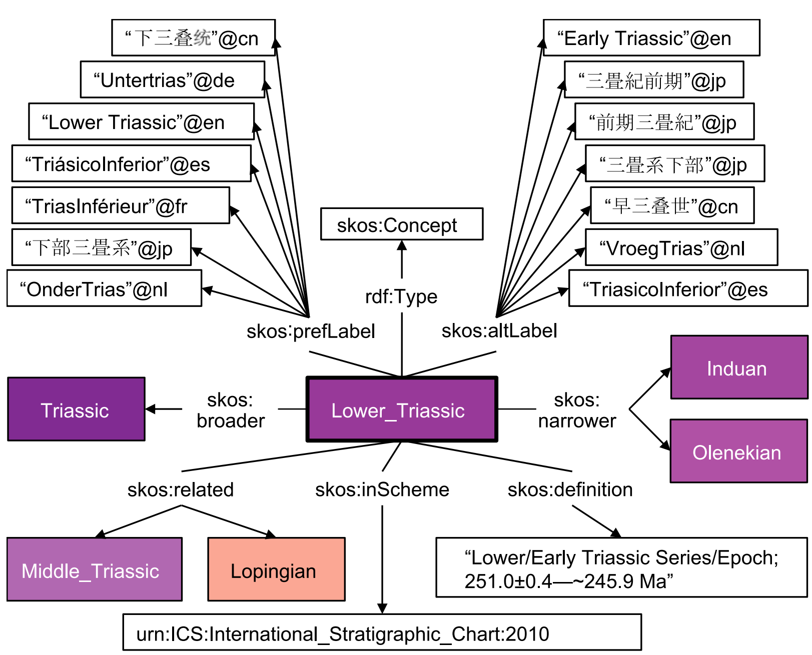

Fig. 1. Definition of "Lower_Triassic" as a GTS concept with object and datatype properties of the SKOS model: (a) source code and (b) graphic view of (a).

Despite those features, the SKOS model is insufficient for encoding certain semantics in a thesaurus (cf. Tennis and Sutton, 2008). Because GTS is not just a hierarchical structure but an ordinal hierarchical scheme divided by time boundaries (Cox and Richard, 2005), a pure SKOS model cannot properly represent this core feature of GTS. For example, "Lopingian" rocks are older than "Lower_Triassic" rocks, which are, in turn, older than "Middle_-Triassic" rocks. These ordinal relationships cannot be represented properly with the "skos:related" property (Fig. 1). To address this problem, we used the SKOS model in a RDF approach. That is, we extended the SKOS model by adding some user-defined properties and some predefined properties in RDF (cf. Jupp et al., 2008; Pan and Horrocks, 2007; Rector et al., 2004). After collecting multilingual GTS terms, we encoded them with this extended SKOS model.

2.2. Addressing semantics and syntax/lexicon in multilingual GTS terms

Like studies on interoperability of multisource geodata and/or geoinformation (Bishr, 1998; Brodaric and Gahegan, 2006; Ludäscher et al., 2006), collecting multilingual terms of a subject domain in geosciences also involves semantic and syntactic/lexical issues. Semantics deals with meanings of terms whereas syntax/lexicon deals with words and structures of expressions in each language.

The first challenge is addressing semantics. If the meanings of several terms in different languages are the same, then they are semantically matched and they can be registered as entries in a multilingual thesaurus. For a certain geoscience concept, if there are no semantically matched terms in different languages (i.e., one cannot find multilingual terms describing exactly the same thing or falling exactly into the same interrelationships), then it is difficult to register a full entry in a multilingual thesaurus for this concept. The meanings of geoscience concepts, in general, are defined by international commissions of different subject domains in geosciences. In our study of the MLTGTS, global boundaries of geological time are defined by the ICS[15] and the International Stratigraphic Chart compiled by ICS is globally accepted and used by the international geoscience community (cf. Ogg, 2009; U.S. Geological Survey Geologic Names Committee, 2010; Walker and Geissman, 2009). These formed our stable basis for collecting semantically matched GTS terms in different languages.

The second challenge is addressing syntax/lexicon. For semantically matched terms in different languages for the same concept, there may be several synonyms describing the same concept in every language. It is not wrong to use synonyms in one language to record geodata as long as users can read them and understand their meanings in that language. However, for a multilingual geoscience thesaurus, many synonyms in different languages should be collected as much as possible so that they are all recognized when that thesaurus is used by a computer. For example, "Cainozoic" is a synonym of "Cenozoic" in English; "Paleogeno" is a synonym of "Paleógeno" in Spanish; "Quartaer" is a synonym of "Quartär" in German; and "ジュラ紀前期" is a synonym of "前期ジュラ紀" in Japanese. Several of such synonyms in different languages were collected for the MLTGTS discussed here.

Related to the semantic and syntactic/lexical issues addressed in collecting multilingual GTS terms, there are two approaches commonly applied to match multilingual terms (Miles et al., 2001): (1) interlingual mapping or (2) multilingual labeling. The first approach can be used to address the lack of semantically matched multilingual terms. For instance, consider at least two independent monolingual thesauri covering the same or similar subject domains but with different hierarchical and associative relationships. Mappings between terms in each pair of thesauri can be performed via the first approach, but such mappings are time-consuming and, sometimes, even impossible. In contrast, the second approach deals with terms in different languages with the same conceptual structure (i.e., terms that have already been semantically matched). Thus, the second approach can be used to arrange multilingual terms in a thesaurus.

We applied the multilingual labeling approach in developing the MLTGTS because boundaries in the GTS proposed by ICS are accepted globally as a common conceptual schema in this subject domain. We collected standard GTS terms in seven languages (i.e., English, Dutch, German, Spanish, French, Chinese, and Japanese) by referring to the MTG and AMTG. However, some GTS terms are (a) not available in MTG and AMTG (e.g., terms at levels of "Series/Epoch" and "Stage/Age" in "Permian" and "Silurian"), (b) out of date (e.g., the Chinese term "晚第三纪" and Japanese term "新第三紀" of "Neogene" in AMTG). In addition, some GTS terms in the AMTG are mismatched (e.g., the Chinese geochronologic term "早泥盆世" ("Early Devonian Epoch") is mismatched with the Japanese chronostratigraphic term "下部デボン系" ("Lower Devonian Series") in the entry with English term "Lower Devonian" as the basic reference). In this regard and to make the collection of multilingual GTS terms complete and up-to-date, we searched websites of geological institutions of different countries and a multilingual GTS thesaurus[16] recently edited by the Geoscience Concept Definitions Task Group of the CGI-IUGS.

Moreover, we considered the two nomenclature systems for GTS terms—one for chronostratigraphy (i.e., Eonothem, Erathem, System, Series, and Stage) and the other for geochronology (i.e., Eon, Era, Period, Epoch, and Age). In some western languages (e.g., English or Spanish) wherein the basic terms are the same, chronostratigraphic and geochronologic terms are often indistinct in actual applications, but in some other languages (e.g., Chinese or Japanese) GTS terms include units by which chronostratigraphic terms are distinguished from geochronologic terms. For example, in Chinese, the chronostratigraphic term "泥盆统" ("Devonian Series") corresponds to but is distinct from the geochronologic term "泥盆世" ("Devonian Epoch"). Another concern from the twofold nomenclature relates to GTS terms containing "Upper/Late" and "Lower/Early" at the level of "Series/Epoch." GTS terms containing "Upper" and "Lower" were originally proposed for chronostratigraphy, whereas terms containing "Late" and "Early" are for geochronology (Haile, 1987; U.S. Geological Survey Geologic Names Committee, 2010). In many actual works, "Upper" is equated to "Late" and "Lower" to "Early" and are used interchangeably, causing confusions for other workers. To improve the semantic precision of the developed MLTGTS, we regard terms containing "Upper" or "Lower" (e.g., "Upper Cretaceous," "Lower Triassic") as chronostratigraphic terms and, correspondingly, terms containing "Late" or "Early" (e.g., "Late Cretaceous," "Early Triassic") as geochronologic terms.

2.3. Extending SKOS model to capture GTS structure

We used "skos:prefLabel" and "skos:altLabel" of the SKOS model to capture multilingual GTS terms in the developed MLTGTS. To capture interrelationships between GTS terms and add more semantic expressions, we extended the SKOS model by adding several other object and datatype properties in the developed MLTGTS. We collected English chronostratigraphic terms from the 2009 ICS International Stratigraphic Chart (2009

[15] International Commission on Stratigraphy. http://www.stratigraphy.org [accessed February 07, 2011].

[16] https://www.seegrid.csiro.au/subversion/CGI_CDTGVocabulary/tags/SKOSVocabularies/ICS_TimeScale2008.rdf [accessed February 07, 2011].

a

```
1  <skos:Concept rdf:ID="Lower_Triassic">
2    <skos:definition>Lower/Early Triassic Series/Epoch; 251.0±0.4—~245.9 Ma</skos:definition>
3    <skos:prefLabel xml:lang="cn"> 下三叠统 </skos:prefLabel>
4    <skos:prefLabel xml:lang="de">Untertrias</skos:prefLabel>
5    <skos:prefLabel xml:lang="en">Lower Triassic</skos:prefLabel>
6    <skos:prefLabel xml:lang="es">Triásico Inferior</skos:prefLabel>
7    <skos:prefLabel xml:lang="fr">Trias Inférieur</skos:prefLabel>
8    <skos:prefLabel xml:lang="jp">下部三畳系 </skos:prefLabel>
9    <skos:prefLabel xml:lang="nl">Onder Trias</skos:prefLabel>
10   <skos:altLabel xml:lang="en">Early Triassic</skos:altLabel>
11   <skos:altLabel xml:lang="jp"> 三畳紀前期 </skos:altLabel>
12   <skos:altLabel xml:lang="jp">前期三畳紀</skos:altLabel>
13   <skos:altLabel xml:lang="jp"> 三畳系下部 </skos:altLabel>
14   <skos:altLabel xml:lang="cn">早三叠世</skos:altLabel>
15   <skos:altLabel xml:lang="nl">Vroeg Trias</skos:altLabel>
16   <skos:altLabel xml:lang="es">Triasico Inferior</skos:altLabel>
17   <skos:broader rdf:resource="#Triassic"/>
18   <skos:narrower rdf:resource="#Olenekian"/>
19   <skos:narrower rdf:resource="#Induan"/>
20   <rdf:type rdf:resource="#Series"/>
21   <gts:lowerThan rdf:resource="#Middle_Triassic"/>
22   <gts:upperThan rdf:resource="#Lopingian"/>
23   <gts:upperBoundaryTime>~245.9 Ma</gts:upperBoundaryTime>
24   <gts:lowerBoundaryTime>251.0±0.4 Ma</gts:lowerBoundaryTime>
25   <gts:basalGsspInfo>
26   https://engineering.purdue.edu/stratigraphy/gssp/detail.php?periodid=76·top_parentid=35
27   </gts:basalGsspInfo>
28   <skos:inScheme rdf:resource="urn:ICS:International_Stratigraphic_Chart:2010"/>
29 </skos:Concept>
```

b

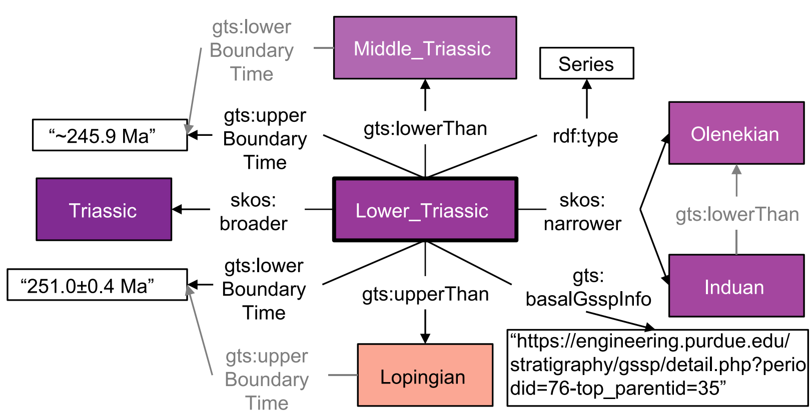

Fig. 2. Definition of "Lower_Triassic" as a GTS concept with an extended SKOS model. Several pure SKOS properties in (a) (i.e., "skos:Concept," "skos:definition," "skos:prefLabel," "skos:altLabel," "skos:inScheme") are omitted in (b) to show the difference between (b) and Fig. 1b. Text and symbols in gray color are not shown in (a), but are included in definitions of other GTS terms in the developed MLTGTS. (a) Source code and (b) graphic view of a part of (a).

ICS chart) (Ogg, 2009), and used them as basic references of GTS concepts in the MLTGTS. For example, "Lower_Triassic" is encoded with "skos:Concept" in line 1 in Fig. 2a. In line 2, the definition of "Lower_Triassic" is encoded with "skos:definition." In lines 3–9, chronostratigraphic terms in seven languages are encoded as preferred labels of "Lower_Triassic" with "skos:prefLabel." In lines 10, 12, 14, and 15, geochronologic terms in other languages are encoded as alternative labels of "Lower_Triassic" with "skos:altLabel." In lines 11, 13, and 16, synonyms of both chronostratigraphic and geochronologic terms of the same SKOS concept are also encoded with "skos:altLabel."

Unlike the definition of "Lower_Triassic" with the pure SKOS model (Fig. 1a), several external properties are used in the definition of that concept in the extended SKOS model (Fig. 2a) to represent the ordinal hierarchical structure of GTS. In line 20 in Fig. 2a, "Lower_Triassic" is defined as an instance of "Series" (a subclass of chronostratigraphic units) by using a RDF property "rdf:type." In lines 21 and 22, two properties "gts:lowerThan" and "gts:upperThan" are used to represent that "Lower_Triassic" rocks are stratigraphically lower than "Middle_Triassic" rocks and stratigraphically upper than "Lopingian" rocks, respectively. In lines 23 and 24, properties "gts:upperBoundaryTime" and "gts:lowerBoundaryTime" are used to record, respectively, the upper and lower time boundaries of "Lower_Triassic," which are derived from the 2009 ICS chart. In lines 25–27, "gts:basalGsspInfo" is used to record a web address pointing to the information of the basal Global Boundary Stratotype Section and Point (GSSP)[17] of "Lower_Triassic." Fig. 2b shows the difference of Fig. 2a from Fig. 1a.

2.4. Summary of building the SKOS-based MLTGTS

The process consists of the following steps. First, a subject domain (i.e., GTS) of a thesaurus is chosen. Second, multilingual GTS terms are collected. To improve the interoperability of the resulting thesaurus, we collected most terms from the MTG and AMTG because they have been extensively reviewed and are

[17] GSSP information is maintained by the Subcommission for Stratigraphic Information of the International Commission of Stratigraphy (https://engineering.purdue.edu/Stratigraphy/index.html) [accessed February 07, 2011].

widely accepted. Third, multilingual labeling (Miles et al., 2001) is used for organizing multilingual GTS terms because of the global common conceptual structure of GTS coordinated by the ICS; and English chronostratigraphic terms from the 2009 ICS chart are used as the basic references because this chart is accepted and used in the geoscience community globally. Fourth, an extended SKOS model is used to encode multilingual GTS terms and to represent the ordinal hierarchical structure of GTS. Finally, the MLTGTS is refined by adding more synonyms of multilingual GTS terms obtained from actual geodata. The workload for the first version of the MLTGTS is about 150 man-hours, with participation of geologists from different language backgrounds. However, the refinement of the MLTGTS is a continuous process, because there are certainly other synonyms of GTS terms existing in actual geodata in different languages. The size of the current MLTGTS is about 225 kB and it is still increasing slowly.

3. Recognizing and translating GTS terms retrieved from WMS

The primary functions of the developed MLTGTS are to recognize and translate GTS terms in GTS records retrieved from geological maps on WMS servers. These can be achieved in four steps (Fig. 3). First, the GTS record of an area is retrieved from an online geological map provided by a WMS server. Second, GTS terms in that record and their languages are recognized. Third, a recognized GTS term is chosen and information about it is searched in the MLTGTS, and then displayed on the user interface. Finally, other languages supported by the MLTGTS can be chosen on the user interface, such that the MLTGTS is researched and the displayed term and related information are translated into the chosen language. In this workflow, the SKOS-based MLTGTS is a RDF document and the user interface is a webpage encoded with HTML (HyperText Markup Language).

We used JavaScript and Ajax (Asynchronous JavaScript and XML) techniques (Garrett, 2005) for accessing WMS and parsing the MLTGTS through the user interface, such that we could realize with one programming language most of the designed functions in the workflow (Fig. 3). We first developed functions for retrieving GTS records from WMS servers (i.e., step 1). Because geological maps are produced by different countries and are in different languages, original GTS records are often presented in diverse styles. Thus, we developed a function to correct spelling errors (Kukich, 1992; Lam-Adesina and Jones, 2006) and to reformat texts of retrieved GTS records for compatibility with terms in the MLTGTS. Because, in step 2, a GTS term under operation has already been recognized in the MLTGTS, steps 3 and 4 simply involve finding the term again in the MLTGTS, and then retrieving information about that term in the chosen language. Compared to steps 1, 3, and 4, recognizing GTS terms in a GTS record (i.e., step 2) costs most of the time in the workflow.

To recognize certain terms from given texts, there are various methods of information retrieval (e.g., Baeza-Yates and Ribeiro-Neto, 2011; Gey et al., 2005; Lazarinis et al., 2009) that can be adopted and/or adapted. Here, our target is comparing a preprocessed GTS record with standard GTS terms in the MLTGTS and then recognizing all GTS terms in that record. To achieve this target, we developed an algorithm (Fig. 4) based on Boolean information retrieval (Koubarakis et al., 2006; Losee, 1997; Radecki, 1983). In this algorithm, we imposed a letter-by-letter equality (LBLE) condition for finding standard GTS terms in a GTS record and, accordingly, this involved an iterative process of letter-by-letter comparison. We used the LBLE condition because, in general, a GTS record is a phrase or a short sentence and, thus, letter-by-letter comparison between a GTS record and standard terms in the MLTGTS does not entail huge workloads for a computer. An ideal result of Boolean information retrieval is that the GTS record contains only one GTS term and it is recognized either as a "skos:prefLabel" term or a "skos:altLabel" term in the MLTGTS (Fig. 4).

However, multilingual GTS terms have their own spelling features, which defy a pure LBLE search. For example, "Trias" (in German) and "Triassic" (in English) are both "skos:prefLabel"s for the concept "Triassic" in the MLTGTS. If a GTS record is "Triassic to Jurassic" (in English), then, by using a pure LBLE search, "Trias" will also be recognized as a GTS term in that GTS record, which is not the case however. The terms "Lower/Early," "Middle," and "Upper/Late" diversify GTS records and pose obstacles in a pure LBLE search. For example, a GTS record "Lower Triassic to Middle Triassic" may be written as "Lower to Middle Triassic." A pure LBLE search can recognize "Middle Triassic" from the abridged record, whereas the GTS term "Lower" (i.e., "Lower Triassic") is ignored. Worse still, a pure LBLE search will also recognize "Triassic" and "Trias" as GTS terms in that GTS record.

To address those challenges and to complement the LBLE search, we developed a group of other methods forming what we call a characteristic-oriented term retrieval algorithm (Fig. 4), because they focus on the characteristics of GTS terms and GTS records. For example, one of the methods extends abridged terms in the GTS record into full terms before using the LBLE search. For instance, if a GTS record is "Lower to Upper Triassic," it is extended to "Lower Triassic to Upper Triassic" prior to LBLE search. Another method deletes fake terms in the recognized term list. For instance, if a GTS record is "Lower Triassic to Lower Jurassic," a pure LBLE search will recognize "Lower Triassic," "Lower Jurassic," "Triassic," "Jurassic," "Trias," and "Jura" as GTS terms in that record. The latter four terms will be recognized as fake terms and will be deleted by the algorithm shown in Fig. 4, leaving only "Lower Triassic" and "Lower Jurassic," as desired, in the result. This characteristic-oriented term retrieval algorithm substantially improved the accuracy of GTS term recognition.

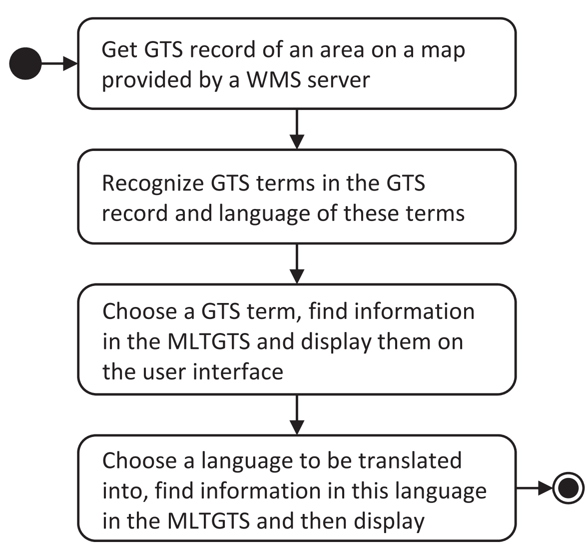

Fig. 3. Four-step workflow for recognizing and translating GTS terms in GTS records retrieved from geological maps on WMS servers.

4. Pilot system, results, and evaluation

We set up a pilot system to test the accuracy of the MLTGTS and the functionality of the JavaScript programs for recognizing and translating GTS terms in GTS records of online geological maps. A purpose of this pilot system, being consistent with the aforementioned workflow, is retrieving background knowledge of GTS terms from the MLTGTS and showing them in a way that is

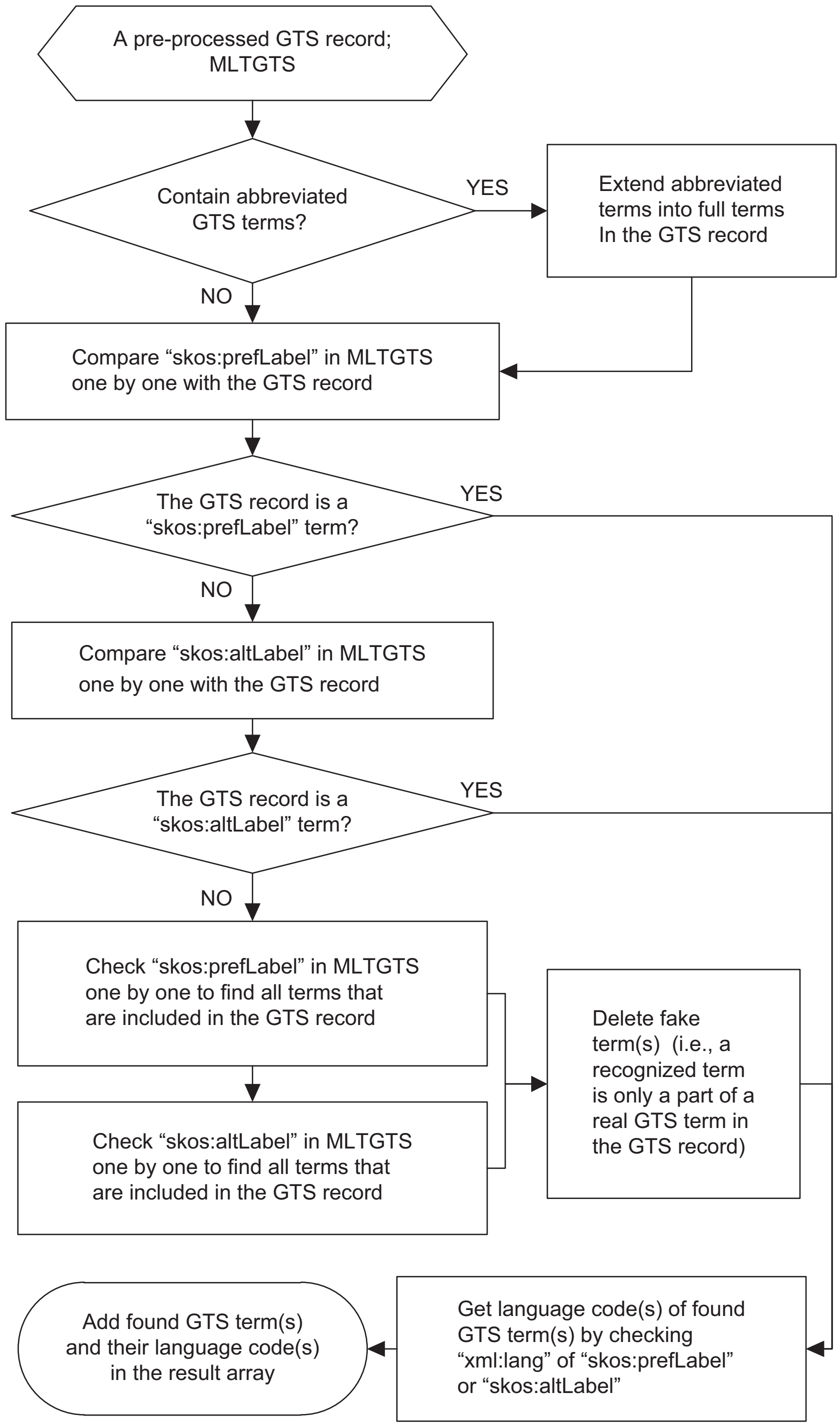

Fig. 4. Algorithm for recognizing GTS terms and their languages in a GTS record by using the developed MLTGTS.

easy for both geologists and nongeologists to access. Datasets in the pilot system include the following geological maps (as ARCGIS .shp files) stored in a self-built WMS server:

- 1:200,000 Geological Map of Japan (GSJ-AIST, 2009), for testing recognition and translation of Japanese GTS terms;
- 1:5,500,000 Geological Map of South America (CGMW et al., 2003), for testing recognition and translation of Spanish GTS terms;
- 1:600,000 Superficial Rock Age Map of The Netherlands (Schokker, 2010), for testing recognition and translation of Dutch GTS terms;
- 1:625,000 Bedrock Age Map of United Kingdom (BGS, 2005), for testing recognition and translation of British English GTS terms;
- 1:250,000 Geologic Map of New York (Dicken et al., 2008), for testing recognition and translation of American English GTS terms.

From remote WMS servers maintained by geological surveys of several countries, we also used the following geological map layers. In the following list, layers from the first four WMS servers are used for testing recognition and translation of English GTS terms, and layers in the last WMS server are used for testing recognition and translation of Dutch and German GTS terms.

- WMS: CCOP Combined Bedrock and Superficial Geology and Age[18];
  Layer: EASIA CCOP 1:2,000,000 Combined Bedrock and Superficial Geology and Age.
- WMS: GSJ Combined Bedrock and Superficial Geology and Age[19];
  Layer: JPN GSJ 1:1,000,000 Combined Bedrock and Superficial Geology and Age.
- WMS: BGS GSN Bedrock geology[20];
  Layer: NAM GSN 1:1,000,000 Bedrock Age.
- WMS: BGS Bedrock and Superficial geology[21];
  Layer: GBR BGS 1:625,000 Bedrock Age.
- WMS: DinoMap Geological maps[22];
  Layers: Geological map of NL 600k (German legend); Geological map of NRW 100k (original map).

We used the following open-source or free software programs for developing the pilot system:

- Protégé 4.0.2[23] and SKOSEd-1.0-alpha (build04),[24] for editing the MLTGTS;
- Notepad++ 5.7,[25] for revising the MLTGTS, and for editing JavaScript programs and the HTML file of the user interface;
- Firefox 3.6.8[26] and Firebug 1.5.4,[27] for debugging JavaScript programs and browsing the HTML file of the user interface;
- GeoServer 2.0.1,[28] for setting up a WMS server;
- uDig 1.1,[29] for editing SLD (Style Layer Descriptor) files of geological maps stored in GeoServer 2.0.1;
- OpenLayers 2.9.1[30]: for retrieving spatial and attribute data from a WMS server.

By operating the pilot system with geological maps on the self-built and remote WMS servers, we could translate recognized GTS terms in English, Spanish, Dutch, German, or Japanese and relevant information about them into any one of the seven languages supported by the MLTGTS. Fig. 5a shows the user interface of the pilot system with the layer "EASIA CCOP 1:2,000,000 Combined Bedrock and Superficial Geology and Age" retrieved from a remote WMS server. In this example, an area in the map was clicked first. However, because the cross-domain data access is limited by JavaScript currently,[31] we used a pop-up window to show the retrieved GTS record (bottom part of Fig. 5a). Then, the record was copied manually to a text box at the bottom left part of the main user interface, and the button "Translate" was clicked. Two GTS terms, "Paleogene" and "Cretaceous," were recognized and listed at the top of the main user interface. Meanwhile, the language code (i.e., "en") of the two terms was recognized and recorded, but not shown on the user interface. Then, either recognized term can be clicked and information of this term is shown in its original language. Fig. 5a shows the result after the term "Paleogene" was clicked. After that, any one of the seven flag buttons at the bottom right part of the user interface can be clicked, and information shown on the user interface is translated into the corresponding language. Fig. 5b shows the result after the Dutch flag button was clicked.

Applying the MLTGTS and JavaScript programs in the pilot system is not a "one station stop" work. Instead, evaluations and revisions on them are iterative. By operating the pilot system with GTS records retrieved from actual geological maps in different languages, we can find insufficiencies of the MLTGTS and the JavaScript programs. We can then evaluate the results, and revise the MLTGTS and/or the JavaScript programs to improve their accuracy and functionality. For example, we recently took the Geological Map of Kumamoto from the 1:200,000 Geological Map of Japan and operated the recognition and translation of Japanese GTS records. This map contains 2822 original GTS records, which can be condensed into 32 differentiated GTS records (i.e., the 2822 original GTS records are just repetitions of these 32 differentiated records). With an earlier version of the MLTGTS, the pilot system successfully translated GTS terms in 21 of those 32 differentiated GTS records but failed for the other 11 records. The compositions of the successfully translated 21 GTS records can be classified into four types, whereas the compositions of the unsuccessfully translated 11 GTS records can be classified into three types. Table 2 shows a list of examples of all these seven types. Nine of the 11 unsuccessfully translated GTS records contain valid GTS terms that are not included in the 2009 ICS chart. These unsuccessful pilot results allowed us to revise the MLTGTS by encoding those valid GTS terms that are not included in the 2009 ICS chart. With the revised MLTGTS, the pilot system successfully translated GTS terms in 30 (types 1–6) of the 32 differentiated GTS records but still failed for the other two records (type 7) that contain no GTS terms. We further tested the pilot system (with the revised MLTGTS) to recognize and translate GTS terms in GTS records of other 1:200,000 Geological Maps of Japan, and we obtained satisfactory results (Table 3).

In the pilot system, we also performed a case study of translating GTS terms in geological maps of areas across, say, borders between two countries. That is because geological mapping of border areas is an increasingly discussed topic in recent years (Asch, 2001; OneGeology-Europe Consortium, 2010; Podemski, 2005; Satkunas and Graniczny, 1997). Geological units are naturally independent of administrative borders, but geological maps in border areas usually exhibit certain inconsistencies, including GTS nomenclatures (Satkunas et al., 2004). There are vast challenges, including the linguistic barriers discussed above, in harmonizing geological maps in border areas (e.g., Asch, 2001, 2005; Delgado et al., 2001). The purpose of this case study was to check whether SKOS-based multilingual geoscience thesauri can address the challenge of linguistic barriers in harmonizing geological maps in border areas.

For this case study, we retrieved two geological maps – geological map of NL 600k (German legend) and geological map of NRW 100k (original map) – from the WMS server "DinoMap Geological maps" (see footnote 22). They cover the Dutch–German border areas between the provinces Overijssel, Gelderland, and Limburg of the Netherlands and the state North Rhine–Westphalia (NRW) of Germany. In the first map, 11 differentiated Dutch GTS records were found along the Dutch–German border areas; whereas in the second map, 14 differentiated German GTS records were found. Compositions of GTS records in the two maps can be classified into three and four types, respectively (Table 4). With the earlier version of the MLTGTS, the pilot system

[18] http://geodata1.geogrid.org/mapserv/CCOP_Combined_Bedrock_and_Superficial_Geology_and_Age/wms? [accessed February 07, 2011].
[19] http://geodata1.geogrid.org/mapserv/GSJ_Combined_Bedrock_and_Superficial_Geology_and_Age/wms? [accessed February 07, 2011].
[20] http://ogc.bgs.ac.uk/cgi-bin/BGS_GSN_Bedrock_Geology/wms? [accessed February 07, 2011].
[21] http://ogc.bgs.ac.uk/cgi-bin/BGS_Bedrock_and_Superficial_Geology/wms? [accessed February 07, 2011].
[22] http://www.dinoservices.nl/wms/dinomap/M07M0034? [accessed February 07, 2011].
[23] http://protege.stanford.edu [accessed February 07, 2011].
[24] http://code.google.com/p/skoseditor [accessed February 07, 2011].
[25] http://notepad-plus-plus.org [accessed February 07, 2011].
[26] http://www.mozilla-europe.org/en/firefox [accessed February 07, 2011].
[27] http://www.getfirebug.com [accessed February 07, 2011].
[28] http://www.geoserver.org [accessed February 07, 2011].
[29] http://udig.refractions.net [accessed February 07, 2011].
[30] http://www.openlayers.org [accessed February 07, 2011].
[31] Technologies of cross-domain data access is now being worked on by the W3C community. http://www.w3.org/TR/cors [accessed February 07, 2011].

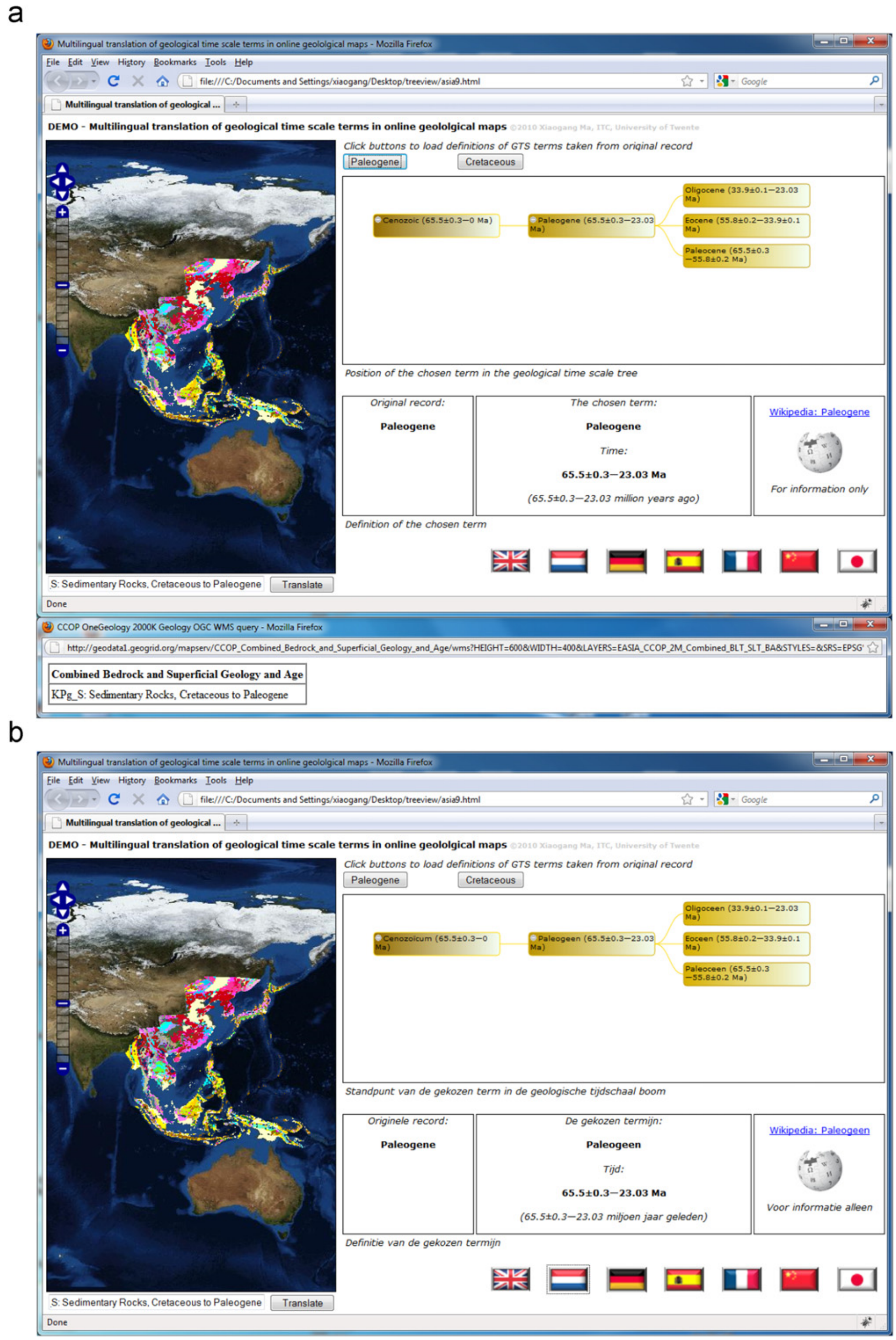

Fig. 5. Running JavaScript programs and MLTGTS through a web browser to translate GTS terms in GTS records retrieved from a WMS server. (a) GTS terms retrieved from a WMS server are recognized and listed on the user interface. Users can choose a GTS term from the list and then its location in the GTS tree structure; its definition and a link to its corresponding Wikipedia page are shown. (b) Users can choose a preferred language by clicking a flag button and then the GTS tree structure; the definition of the GTS term and additional information on the user interface are translated into the chosen language. Geological map reproduced with the permission of the OneGeology secretariat and registered participants. All Rights Reserved. (a) Information of "Paleogene" shown in English and (b) information of "Paleogene" shown in Dutch.

successfully translated GTS terms in 6 (types 1.1 and 1.2) of the 11 differentiated GTS records in the first map and in 11 (types 2.1–2.3) of the 14 differentiated GTS records in the second map. In the first map, we found that the five unsuccessfully translated GTS records (type 1.3) include two terms "Weichseliën" (Weichselian) and "Saaliën" (Saalian). These terms are not included in the 2009 ICS chart, but are used as Stage terms in regional subdivisions of the Pleistocene Series in North West Europe.[32] In the second map, the three unsuccessfully translated GTS records (type 2.4) include two terms "Weichselium" (Weichselian) and

[32] Stratigraphical charts for the Quaternary. http://www.quaternary.stratigraphy.org.uk/charts [accessed February 07, 2011].

Table 2
GTS records in the Geological Map of Kumamoto that are successfully and unsuccessfully translated by the pilot system of the MLTGTS.

Original record	In English	Record type	Pilot system translation
白亜紀	Cretaceous	1. Containing only one GTS term	Successful
シルル紀-デボン紀	Silurian to Devonian	2. Containing two GTS terms, without "Upper," "Late," "Middle," "Lower," or "Early"	Successful
ペルム紀-前期白亜紀	Permian to Early Cretaceous	3. Containing two GTS terms, one or two with "Upper," "Late," "Middle," "Lower," or "Early"	Successful
前-後期ジュラ紀	Early to Late Jurassic	4. Containing two GTS terms, one in abbreviated format	Successful
中期始新世	Middle Eocene	5. Containing only one GTS term, not found in 2009 ICS chart	Unsuccessful
中期始新世-前期漸新世	Middle Eocene to Early Oligocene	6. Containing two GTS terms, not found in 2009 ICS chart	Unsuccessful
時代未詳	Unknown age	7. Containing no GTS terms	Unsuccessful

Table 3
Results of recognizing and translating GTS terms in GTS records of some 1:200,000 Geological Maps of Japan.

Map name	GTS records	Differentiated GTS records	Successful translations	Failed translations
Geological Map of Hyogo	3747	27	25	2
Geological Map of Ibaraki	2419	21	19	2
Geological Map of Miyazaki	4188	31	29	2
Geological Map of Niigata	4906	28	26	2
Geological Map of Osaka	1442	16	15	1

Table 4
GTS records of Dutch–German border areas in the geological map of NL 600k (German legend) and the geological map of NRW 100k (original map) that are successfully and unsuccessfully translated by the pilot system of the MLTGTS.

Map	Original record	In English	Record type	Pilot system translation
NL 600k	Boven Krijt	Upper Cretaceous	1.1. Containing only one GTS term	Successful
NL 600k	Tertiair	Tertiary	1.2. Containing only one GTS term, not found in 2009 ICS chart but found in MLTGTS	Successful
NL 600k	Weichseliën -Saaliën	Saalian to Weichselian	1.3. Containing "Weichseliën" and/or "Saaliën"	Unsuccessful
NRW 100k	Pliozän	Pliocene	2.1. Containing only one GTS term	Successful
NRW 100k	Miozän bis Oligozän	Miocene to Oligocene	2.2. Containing two GTS terms	Successful
NRW 100k	Unterpleistozän	Lower Pleistocene	2.3. Containing only one GTS term, not found in 2009 ICS chart but found in MLTGTS	Successful
NRW 100k	Weichselium	Weichselian	2.4. Containing "Weichselium" and/or "Saalium"	Unsuccessful

"Saalium" (Saalian), which are not included in the 2009 ICS chart. By referring to the chart (v. 2010) provided by the Subcommission on Quaternary Stratigraphy of ICS (see footnote 32), we added these terms and their multilingual versions into the MLTGTS. With the updated MLTGTS, the pilot systems successfully translated all GTS terms in the 25 differentiated GTS records from both maps. By operating the recognition and translation, original Dutch or German GTS terms can be translated into any one of the seven language supported by the MLTGTS. This would provide convenience to geologists working on geological maps of Dutch–German border areas. Results of this case study show the benefits of using SKOS-based multilingual geoscience thesauri in using geological maps of border areas, although it addresses only a part (i.e., linguistic barriers) of the challenges in harmonizing multisource geological maps.

5. Discussion

With proper extensions, the SKOS is functional for encoding multilingual geoscience thesauri into a format that is compatible with the Semantic Web, and SKOS-based multilingual geoscience thesauri are efficient for translating online geoscience records into any language that is supported by the thesauri. We only developed a multilingual thesaurus of GTS in this study, but intuitively we can see that other SKOS-based multilingual thesauri of different subject domains in geosciences can be developed. By using these multilingual geoscience thesauri, geological maps in their native languages can be published online directly, while users can translate the maps and browse the data in their preferred languages. In this way, linguistic barriers between online geological maps can be reduced and, thus, their interoperability can be improved. Results of the pilot system demonstrate the accuracy of the MLTGTS and the functionality of the JavaScript programs to recognize and translate GTS terms in multilingual geological maps. Meanwhile, background information of GTS terms retrieved from the MLTGTS are also displayed on the user interface in the chosen language. Because the multilingual terms and their definitions and relationships in the MLTGTS were collected from credible sources, users can get precise explanations of GTS terms from the MLTGTS. By reading the translated GTS terms and background information in preferred languages, users can access the information represented by the GTS records in an easier way.

It has been extensively discussed that the SKOS model has several advantages compared to other models for encoding thesauri in the Web context (Miles and Pérez-Agüera, 2007; Pastor-Sanchez et al., 2009; van Assem et al., 2006). The SKOS model is based on RDF, making SKOS-based thesauri compatible with other standards and technologies of the Semantic Web. In recent years, the SKOS model has been applied to build or rebuild thesauri in various fields, such as agriculture (Soergel et al., 2004), stratigraphy (Fils et al., 2009), authority files (Voss, 2009), and the aforementioned geoscience thesauri edited by the Geoscience Concept Definitions Task Group of the CGI-IUGS. There are also significant studies on building thesauri from diverse resources (e.g., Broughton, 2006; Fang et al., 2008; Hepp, 2006; Tsuruoka et al., 2008), addressing semantic and syntactic issues from various aspects. In this study, the semantics of the GTS has been defined by the 2009 ICS chart, which is accepted as a global standard. Instead of redefining semantics of the GTS, we adopted the existing standard (cf. Bibby, 2006; McGuinness, 2003) and then adapted it with an extended SKOS model. Although we collected multilingual GTS terms from different resources, we could add these GTS terms into the MLTGTS easily, because the meanings (i.e., semantics) of these terms are defined by their time boundaries and, thus, their locations in the ordinal hierarchical structure GTS are clear.

For GTS thesauri, Cox and Richard (2005) discussed in detail components in the GTS and drew conceptual schemas for them by using Unified Modeling Language (UML). They also transformed the UML schemas into XML formats so that they can be used on the Web. The resulting schemas represent units, boundaries, and GSSPs in the GTS and their relationships. Thus, those schemas not only represent GTS terms but also show how they were derived. Compared to the extended SKOS model used for the MLTGTS, the conceptual schemas of Cox and Richard (2005) are more thorough. The SKOS-based thesaurus of GTS (see footnote 16) edited by the Geoscience Concept Definitions Task Group of CGI-IUGS adapted the work of Cox and Richard (2005) by simplifying the components to fit the SKOS model. The current version of the CGI-IUGS GTS thesaurus also refers to the ICS chart and covers GTS terms in English, French, Italian, and Slovakian. However, unlike the MLTGTS discussed here, the CGI-IUGS GTS thesaurus does not distinguish between chronostratigraphic and geochronologic terms. For some GTS concepts, the CGI-IUGS GTS thesaurus combines chronostratigraphic terms with geochronologic units in their definitions, which potentially causes confusion. For example, "Upper Cretaceous" in the CGI-IUGS GTS thesaurus is defined as "Upper Cretaceous Epoch," whereas it should be "Upper Cretaceous Series" or "Late Cretaceous Epoch." Such issues were discussed within the CGI-IUGS community in a recent workshop[33] and more international cooperation was proposed.

Another example of GTS thesaurus is the multilingual geological age thesaurus developed in the OneGeology-Europe project (1G-E)[34] recently (Asch, 2010). The 1G-E hosts a web portal[35] providing multilingual (i.e., 18 European languages) access to contents of semantically and technically interoperable 1:1000,000 scale geological maps for the whole of Europe (Laxton et al., 2010). Such functions are bolstered by SKOS-based multilingual thesauri of lithology, age (geochronology), genesis, and structures and faults developed by the 1G-E Work Package 3 (1G-E WP3) (Asch, 2010). There are several differences between the SKOS-based thesauri of 1G-E and the MLTGTS discussed here. The MLTGTS in our work adopts the 2009 ICS chart and uses chronostratigraphic units. The geological age thesaurus of 1G-E uses geochronologic units, and subdivides the periods of Precambrian for Europe and adds 27 new terms accordingly (Asch et al., 2010). Another difference is that the MLTGTS in our study includes two Asian languages, while the geological age thesaurus of 1G-E focuses on European languages. Besides these differences, our goal of using MLTGTS to translate GTS records of online geological maps is similar to that of the 1G-E project, although from the current literature of 1G-E we know little about whether or not its web portal applies a workflow of retrieval, recognition, translation, and display that is similar to what we developed in our study (Fig. 3).

The benefits of embedding ontologies in Spatial Data Infrastructures (SDI) have also been discussed significantly in recent years (Georgiadou, 2006; Lacasta et al., 2007; Ludäscher et al., 2003; Sinha et al., 2007). By using ontologies in a SDI, heterogeneous geodata sources can be mapped to common models; meanings of inconsistent concepts can be harmonized and the semantic interoperability of geodata can be improved. Geoscience thesauri, as "simple ontologies," are also functional for improving the interoperability of geodata (Ma et al., 2007; Ma et al., 2010). Because SKOS-based thesauri are compatible with standards and technologies of the Semantic Web, using them can potentially lead to more features in a SDI. The results of the pilot system in our study and the 1G-E web portal already show some of these features. Recently, the AuScope[36] project has built services using SKOS vocabularies for querying geodata (Woodcock et al., 2010). The AuScope vocabularies record synonyms of geoscience terms by using the label "skos:altLabel." Even when users input alternative names of geoscience terms for querying, the vocabulary services can find certain concepts and then retrieve desired geodata. We also used the label "skos:altLabel" in our study for recording synonyms of GTS terms. However, compared to the vocabulary services/applications of AuScope, the application of MLTGTS in our study is not for querying geodata but for recognizing and translating GTS terms and showing background knowledge about them.

Several lessons are learned from this study. First, SKOS provides a concise model for representing hierarchical structures, but it may be insufficient or inappropriate for structures of certain subject domains in geosciences and this may require an extension to the SKOS model in practice. Thus, because GTS is not a pure hierarchical structure but an ordinal hierarchical structure divided by time boundaries, we extended the SKOS model by adding several other properties, as described earlier, so that that the extended model can represent the ordinal hierarchical structure of GTS properly. Second, SKOS is good for encoding multilingual geoscience thesauri, but matching multilingual geoscience terms and building interrelationships still need geoscience knowledge and cooperation of experts from different language backgrounds. Thus, we referred to the 2009 ICS chart for GTS terms in English and to the MTG and AMTG for multilingual GTS terms because they are results of international cooperation that are accepted globally. However, because some GTS terms are mismatched or missed in the MTG and AMTG, we had to refer to various other sources to collect credible multilingual GTS terms. Third, many synonyms in different languages should be collected as much as possible in geoscience thesauri. Although international standards or agreements on professional terms of a certain subject domain exist, synonyms are still used in current geoscience works. For example, in British English there are three GTS terms "Cainozoic," "Palaeozoic," and "Archaean," which correspond to "Cenozoic," "Paleozoic," and "Archean," respectively. Such synonyms were encoded in the MLTGTS so that if they are encountered in practice, they can be recognized by using the MLTGTS. Finally, we cannot use "new" standards to explain "old" data, denoting that if a concept's meaning is changed in the thesaurus, we cannot use it to explain a record using the previous meaning of that concept. For example, in the 2009 ICS chart, the

[33] IUGS-CGI and OneGeology-Europe Geoscience Language Workshop (IGSL 2010). http://www.bgr.bund.de/cln_116/nn_1951520/EN/Themen/GG__geol__Info/IGSL2010 [accessed February 07, 2011].

[34] http://www.onegeology-europe.org/home [accessed February 07, 2011].

[35] http://onegeology-europe.brgm.fr/geoportal/viewer.jsp [accessed February 07, 2011].

[36] http://www.auscope.org [accessed February 07, 2011].

basal boundary of Quaternary is different from that in previous versions of ICS charts. The MLTGTS refers to the 2009 ICS chart for the most recently defined meaning of Quaternary. However, if a record "Quaternary" in a map refers to the 2008 ICS chart, then we cannot use the definition of "Quaternary" in the MLTGTS to explain the meaning of that record (cf. Mascarelli, 2009). This reminds us that thesauri used by a geodata source could be attached along with the geodata, or at least, a record of used thesauri could be noted in the metadata of a geodata source.

Because SKOS-based multilingual geoscience thesauri are still an emerging topic in the field of geosciences, many future works can be proposed. One possible work is collecting more synonyms for GTS terms not only in the seven languages considered in this study but also in other languages to enrich the MLTGTS. Another work is incorporating results and lessons of this study with other efforts for developing multilingual geoscience thesauri, such as that of the Geoscience Concept Definitions Task Group of the CGI-IUGS. In a broader perspective, SKOS-based multilingual geoscience thesauri can be maintained by international task groups in the CGI-IUGS and published online. Meanwhile, they can be accessed and used by many different organizations and individuals globally for various applications (cf. Schäffer et al., 2010). New technologies for parsing SKOS-based thesauri can also be studied further. JavaScript programs are efficient for parsing the MLTGTS in this study because it is small; for parsing a large SKOS-based thesaurus or a group of large thesauri, those programs require further testing. Some other technologies for parsing thesaurus (e.g., SPARQL[37]) can be tested in the future. Transforming the SKOS-based MLTGTS into an OWL-based ontology of GTS is an open topic, because although SKOS and OWL are compatible in physical formats, an OWL-based ontology of GTS is capable of adding more semantic descriptions for concepts and relationships. The work of OWL-based geological time ontology[38] in the SWEET project (Raskin and Pan, 2005) can be referred to in this future study.

6. Conclusions

Fast evolving Web-based technologies provide not only platforms for building online geodata services but also opportunities for alleviating linguistic barriers to geodata use. Among various proposed technologies, the SKOS model is advantageous as a starting point for encoding and applying multilingual geoscience thesauri in the context of the Semantic Web, and it can be extended in conjunction with other approaches to express concepts and relationships of a subject domain properly. In this study, we encoded a multilingual thesaurus of geological time scale with an extended SKOS model and, coupled with the thesaurus, we implemented methods of characteristic-oriented term retrieval in JavaScript programs for recognizing and translating geological time-scale terms in online geological maps. The developed thesaurus and associated programs were used in a pilot system to recognize and translate geological time-scale terms in actual geological maps. Results of the pilot system proved the accuracy of the developed multilingual thesaurus of geological time scale and the functionality of the JavaScript programs. Our study shows that SKOS-based multilingual geoscience thesauri can be functional for alleviating linguistic barriers between online geological maps and, thus, improving their interoperability. However, background knowledge of a subject domain is essential when SKOS is used for building a multilingual geoscience thesaurus of that domain. In addition, it may be necessary to extend the SKOS model in order to obtain satisfactory semantic expressions in certain subject domains in geosciences.

[37] http://www.w3.org/TR/rdf-sparql-query [accessed February 07, 2011].

[38] http://sweet.jpl.nasa.gov/2.0/timeGeologic.owl [accessed February 07, 2011].

Acknowledgments

This work was partly supported by the key project of Chinese Ministry of Education (No. 109107). We express our appreciation to researchers of the CGI-IUGS community for their helpful comments and suggestions. Especially, we thank Dr. Steven Richard, Dr. Guillaume Duclaux, Mr. Jan Jellema, Prof. Peter Fox, and Dr. Marcus Ebner for their suggestions on encoding SKOS-based geoscience thesauri, and Dr. Kristine Asch for inviting us to attend the IGSL2010 workshop. We thank Prof. Stanley Finney, Dr. Frits Agterberg, Dr. Koji Wakita, and Dr. Harald Poelchau for their discussions on nomenclatures of geological time-scale terms in different languages; Dr. Jeroen Schokker for providing a part of the pilot dataset; and Mr. Barend Köbben for discussing techniques of Ajax. We also thank Mr. Emilio Cortegoso Lobato for his open source Graphic JavaScript Tree, which enriched features of the pilot system developed in this study. Finally, we thank Dr. Mark van Assem, Dr. Simon Cox, and Dr. Boyan Brodaric for their insightful and constructive comments on our paper.

References

Antoniou, G., Franconi, E., van Harmelen, R., 2005. Introduction to semantic web ontology languages. In: Eisinger, N., Maluszynski, J. (Eds.), Reasoning Web; Lecture Notes in Computer Science, vol. 3564. Springer-Verlag, Berlin, pp. 1–21.

Asch, K., 2001. The IGME 5000: a means to pan-European geological standards (or just another local and transient convention)? In: Proceedings of 2001 Annual Conference of the International Association for Mathematical Geology, Cancún, Mexico, 7 pp., 〈http://www.kgs.ku.edu/Conferences/IAMG/Sessions/I/Papers/asch_o.pdf〉 [Accessed February 07, 2011].

Asch, K., 2005. The new digital geological map of Europe and standardisation: consistency as the last refuge of the unimaginative?. In: Ostaficzuk, S. (Ed.), The Current Role of Geological Mapping in Geosciences, pp. 1–8.

Asch, K., 2010. The challenge of building a vocabulary for the geology of Europe. IUGS-CGI and OneGeology-Europe International Geoscience Language Workshop (IGSL 2010), 25–26 August 2010, Berlin, Germany (18 slides), 〈http://www.bgr.bund.de/cln_144/nn_1960030/EN/Themen/GG__geol__Info/IGSL2010/Downloads/Asch__IGSL__2,templateId=raw,property=publicationFile.pdf/Asch_IGSL_2.pdf〉 [Accessed February 07, 2011].

Asch, K., Jackson, I., 2006. Commission for the Management & Application of Geoscience Information (CGI). Episodes 29 (3), 231–233.

Asch, K., Laxton, J., Bavec, M., Bergman, S., Perez Cerdan, F., Declercq, P.Y., Janjou, D., Kacer, S., Klicker, M., Nironen, M., Pantaloni, M., Schubert, C., 2010. Explanatory notes for the vocabulary to describe spatial geological data in Europe at a 1: 1 million scale—for the eContentPlus project OneGeology-Europe, 86 pp., 〈http://www.bgr.bund.de/cln_144/nn_1960030/EN/Themen/GG__geol__Info/IGSL2010/Downloads/Asch__IGSL__2,templateId=raw,property=publicationFile.pdf/Asch_IGSL_2.pdf〉 [Accessed February 07, 2011].

Baeza-Yates, R., Ribeiro-Neto, B., 2011. Modern Information Retrieval 2nd, ed. Addison-Wesley, Harlow, England, 870 pp.

Berners-Lee, T., Hendler, J., Lassila, O., 2001. The Semantic Web. Scientific American 284 (5), 34–43.

Bibby, L., 2006. Establishing vocabularies for the exchange of geological map data? How to herd stray cats. ASEG Extended Abstracts (online publication of Australian Society of Exploration Geophysicists) 2006 (1). doi:10.1071/ASEG2006ab1015.

Bishr, Y., 1998. Overcoming the semantic and other barriers to GIS interoperability. International Journal of Geographical Information Science 12 (4), 299–314.

Bishr, Y., 2006. Geospatial Semantic Web. In: Rana, S., Sharma, J. (Eds.), Frontiers of Geographic Information Technology. Springer, Berlin, pp. 139–154.

BGS, 2005. Britain Beneath Our Feet, British Geological Survey (BGS) (on CD-ROM).

Brodaric, B., Gahegan, M., 2006. Representing geoscientific knowledge in cyberinfrastructure: challenges, approaches and implementations. In: Sinha, A.K. (Ed.), Geoinformatics: Data to Knowledge; Geological Society of America Special Paper 397. Geological Society of America, Boulder, CO, USA, pp. 1–20.

Broughton, V., 2006. The need for a faceted methods of information retrieval. Aslib Proceedings 58 (1–2), 49–72.

Buccella, A., Cechich, A., Fillottrani, P., 2009. Ontology-driven geographic information integration: a survey of current approaches. Computers & Geosciences 35 (4), 710–723.

CCOP, CIFEG (Eds.), 2006. Asian Multilingual Thesaurus of Geosciences. Coordinating Committee for Geoscience Programmes in East and Southeast Asia (CCOP) & Centre International pour la Formation et les Echanges en Géosciences

(CIFEG), 563 pp., 〈http://www.ccop.or.th/download/pub/AMTG_2006.pdf〉 [Accessed February 07, 2011].

CGMW, CPRM, DNPM, 2003. 1:5,500,000 Geological Map of South America. Commission for the Geological Map of the World (CGMW), Geological Survey of Brazil (CPRM), National Department of Mineral Production, Brazil (DNPM) (on CD-ROM).

Cox, S.J.D., Richard, S.M., 2005. A formal model for the geologic time scale and global stratotype section and point, compatible with geospatial information transfer standards. Geosphere 1 (3), 119–137.

Davies, J., Fensel, D., van Harmelen, F., 2003. Towards the Semantic Web: Ontology-Driven Knowledge Management. John Wiley & Sons, Chichester, England, 328 pp.

Delgado, L., Ortíz, G., Hernández, J.R., Cerdán, F.P., 2001. Building of trans-national geologic map databases—The Iberian Pyrite Belt case. In: Proceedings of 2001 Annual Conference of the International Association for Mathematical Geology, Cancún, Mexico, 10 pp., 〈http://www.kgs.ku.edu/Conferences/IAMG/Sessions/I/Papers/Delgado/delgado.pdf〉 [Accessed February 07, 2011].

Deliiska, B., 2007. Thesaurus and domain ontology of geoinformatics. Transactions in GIS 11 (4), 637–651.

de Mulder, E.F.J., Nield, T., Derbyshire, E., 2006. The international year of planet Earth (2007–2009): earth sciences for society. Episodes 29, 82–86.

Dicken, C.L., Nicholson, S.W., Horton, J.D., Kinney, S.A., Gunther, G., Foose, M.P., Mueller, J.A.L., 2008. Preliminary Integrated Geologic Map Databases for the United States: Delaware, Maryland, New York, Pennsylvania, and Virginia (Version 1.1), 〈http://pubs.usgs.gov/of/2005/1325〉 [Accessed February 07, 2011].

Fang, Y.C., Huang, H.C., Chen, H.H., Juan, H.F., 2008. TCMGeneDIT: a database for associated traditional Chinese medicine, gene and disease information using text mining. BMC Complementary and Alternative Medicine, 8. doi:10.1186/1472-6882-8-58.

Fils, D., Cervato, C., Reed, J., Diver, P., Tang, X., Bohling, G., Greer, D., 2009. CHRONOS architecture: experiences with an open-source services-oriented architecture for geoinformatics. Computers & Geosciences 35 (4), 774–782.

Garcia-Sanchez, F., Valencia-Garcia, R., Martinez-Bejar, R., Fernandez-Breis, J.T., 2009. An ontology, intelligent agent-based framework for the provision of semantic web services. Expert Systems with Applications 36 (2), 3167–3187.

Garrett, J.J., 2005. Ajax: a new approach to web applications, 〈http://www.adaptivepath.com/ideas/essays/archives/000385.php〉 [Accessed February 07, 2011].

Georgiadou, P.Y., 2006. SDI ontology and implications for research in the developing world: editorial. International Journal of Spatial Data Infrastructures Research 1 (1), 51–64.

Gey, F.C., Kando, N., Peters, C., 2005. Cross-language information retrieval: the way ahead. Information Processing & Management 41 (3), 415–431.

Gilchrist, A., 2003. Thesauri, taxonomies and ontologies—an etymological note. Journal of Documentation 59 (1), 7–18.

Gravesteijn, J., Kortman, C., Potenza, R., Rassam, G.N. (Eds.), 1995. Multilingual Thesaurus of Geosciences 2nd ed. Information Today, Inc., Medford, NJ, USA.

Gravesteijn, J., Rassam, G.N., 1990. International cooperation in geoscience information: a global challenge. Episodes 13 (3), 178–183.

Gruber, T.R., 1995. Toward principles for the design of ontologies used for knowledge sharing. International Journal of Human–Computer Studies 43 (5–6), 907–928.

GSJ-AIST, 2009. Digital Geoscience Map Series G-16: Seamless Geological Map of Japan at a scale of 1:200,000. Geological Survey of Japan (GSJ)—National Institute of Advanced Industrial Science and Technology (AIST) (on CD-ROM).

Guarino, N., 1997. Understanding, building and using ontologies. International Journal of Human–Computer Studies 46 (2–3), 293–310.

Guarino, N., 1998. Formal ontology and information systems. In: Guarino, N. (Ed.), Formal Ontology in Information Systems. IOS Press, Amsterdam, the Netherlands, pp. 3–15.

Haile, N.S., 1987. Time and age in geology: the use of upper/lower, late/early in stratigraphic nomenclature. Marine and Petroleum Geology 4 (3), 255–257.

Hepp, M., 2006. Products and services ontologies: a methodology for deriving OWL ontologies from industrial categorization standards. International Journal on Semantic Web and Information Systems 2 (1), 72–99.

Jackson, I., 2007. OneGeology—making geological map data for the earth accessible. Episodes 30 (1), 60–61.

Jackson, I., Wyborn, L., 2008. One planet: OneGeology? The Google Earth revolution and the geological data deficit. Environmental Geology 53 (6), 1377–1380.

Jupp, S., Stevens, R., Bechhofer, S., Yesilada, Y., Kostkova, P., 2008. Knowledge representation for web navigation. In: Proceedings of the Semantic Web Applications and Tools for the Life Sciences (SWAT4LS 2008), e-Science Institute, Edinburgh, Scotland, 15 pp., 〈http://www.cs.man.ac.uk/~sjupp/publications/2008/swat4ls-Jupp.pdf〉 [Accessed February 07, 2011].

Koubarakis, M., Skiadopoulos, S., Tryfonopoulos, C., 2006. Logic and computational complexity for Boolean information retrieval. IEEE Transactions on Knowledge and Data Engineering 18 (12), 1659–1666.

Kukich, K., 1992. Techniques for automatically correcting words in text. Computing Surveys 24 (4), 377–439.

Lacasta, J., Nogueras-Iso, J., Béjar, R., Muro-Medrano, P.R., Zarazaga-Soria, F.J., 2007. A web ontology service to facilitate interoperability within a spatial data infrastructure: applicability to discovery. Data & Knowledge Engineering 63 (3), 947–971.

Lam-Adesina, A.M., Jones, G.J.F., 2006. Examining and improving the effectiveness of relevance feedback for retrieval of scanned text documents. Information Processing & Management 42 (3), 633–649.

Laxton, J., Serrano, J.-J., Tellez-Arenas, A., 2010. Geological applications using geospatial standards—an example from OneGeology-Europe and GeoSciML. International Journal of Digital Earth 3 (Suppl. 1), 31–49.

Lazarinis, F., Vilares, J., Tait, J., Efthimiadis, E., 2009. Current research issues and trends in non-English Web searching. Information Retrieval 12 (3), 230–250.

Lloyd, J.J., 1973. Concept relations in a multilingual thesaurus. Geology 1 (4), 152.

Losee, R.M., 1997. Comparing Boolean and probabilistic information retrieval systems across queries and disciplines. Journal of the American Society for Information Science 48 (2), 143–156.

Ludäscher, B., Lin, K., Bowers, S., Jaeger-Frank, E., Brodaric, B., Baru, C., 2006. Managing scientific data: from data integration to scientific workflows. In: Sinha, A.K. (Ed.), Geoinformatics: Data to Knowledge; Geological Society of America Special Paper 397. Geological Society of America, Boulder, CO, USA, pp. 109–130.

Ludäscher, B., Lin, K., Brodaric, B., Baru, C., 2003. GEON: toward a cyberinfrastructure for the geosciences—a prototype for geological map interoperability via domain ontologies. In: Proceedings of the Workshop on Digital Mapping Techniques'03, Millersville, Pennsylvania, USA, pp. 223–229.

Ma, X., Wang, X., Wu, C., Ju, F., 2007. Metadata hierarchy in integrated geoscientific database for regional mineral prospecting. Geo-spatial Information Science 10 (3), 223–227.

Ma, X., Wu, C., Carranza, E.J.M., Schetselaar, E.M., van der Meer, F.D., Liu, G., Wang, X., Zhang, X., 2010. Development of a controlled vocabulary for semantic interoperability of mineral exploration geodata for mining projects. Computers & Geosciences 36 (12), 1512–1522.

Mascarelli, A.L., 2009. Quaternary geologists win timescale vote. Nature 459, 624.

McGuinness, D.L., 2003. Ontologies come of age. In: Fensel, D., Hendler, J., Lieberman, H., Wahlster, W. (Eds.), Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential. MIT Press, Cambridge, MA, USA, pp. 171–196.

Miles, A., Matthews, B., Wilson, M., 2001. RDF Encoding of Multilingual Thesauri, 〈http://www.w3c.rl.ac.uk/SWAD/deliverables/8.3.html〉 [Accessed February 07, 2011].

Miles, A., Pérez-Agüera, J.R., 2007. SKOS: simple knowledge organisation for the web. Cataloging & Classification Quarterly 43 (3), 69–83.

Ogg, J.G., 2009. International Stratigraphic Chart, International Commission on Stratigraphy, 1 p., 〈http://www.stratigraphy.org/upload/ISChart2009.pdf〉 [Accessed February 07, 2011].

OneGeology-Europe Consortium, 2010. One Europe One Geology: Applying Geoscience for Society, OneGeology-Europe Consortium, 84 pp.

Pan, J.Z., Horrocks, I., 2007. RDFS(FA): connecting RDF(S) and OWL DL. IEEE Transactions on Knowledge and Data Engineering 19 (2), 192–206.

Pastor-Sanchez, J.A., Mendez, F.J.M., Rodriguez-Munoz, J.V., 2009. Advantages of thesaurus representation using the Simple Knowledge Organization System (SKOS) compared with proposed alternatives. Information Research-an International Electronic Journal 14 (4) paper 422, 〈http://informationr.net/ir/14-4/paper422.html〉 [Accessed February 07, 2011].

Peng, Z.-R., Tsou, M.H., 2003. Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks. John Wiley & Sons, Inc., Hoboken, NJ, USA, 679 pp.

Podemski, M., 2005. The Polish Geological Institute activities in the cross-border geological mapping. Przeglad Geologiczny 53 (10/2), 894–898.

Radecki, T., 1983. Generalized Boolean methods of information-retrieval. International Journal of Man–Machine Studies 18 (5), 407–439.

Raskin, R.G., Pan, M.J., 2005. Knowledge representation in the semantic web for earth and environmental terminology (SWEET). Computers & Geosciences 31 (9), 1119–1125.

Rassam, G.N., Gravesteijn, J., Potenza, R. (Eds.), 1988. Multilingual Thesaurus of Geosciences. Pergamon Press, New York, USA.

Rector, A., Drummond, N., Horridge, M., Rogers, J., Knublauch, H., Stevens, R., Wang, H., Wroe, C., 2004. OWL pizzas: practical experience of teaching OWL-DL: common errors and common patterns. In: Motta, E., Shadbolt, N., Stutt, A., Gibbins, N. (Eds.), Engineering Knowledge in the Age of the Semantic Web, Proceedings; Lecture Notes in Computer Science, vol. 3257. Springer-Verlag, Berlin, pp. 63–81.

Satkunas, J., Graniczny, M., 1997. Lithuanian Polish cross-border geoenvironmental mapping and its relevancy to spatial development. In: Marinos, P.G., Koukis, G.C., Tsiambaos, G.C., Stournaras, G.C. (Eds.), Engineering Geology and the Environment. A.A. Balkema, Rotterdam, Netherlands, pp. 1483–1486.

Satkunas, J., Graniczny, M., Lazauskiene, J., 2004. Geoscientific data, sustainable development and international borders. In: Proceedings of the Thematic Session on "Sustainable Development of Geo-resources and Geo-environment," 42nd Coordinating Committee for Geoscience Programmes in East and Southeast Asia (CCOP) Annual Session, Beijing, China, pp. 35–41.

Schäffer, B., Baranski, B., Foerster, T., 2010. Towards Spatial data infrastructures in the clouds. In: Painho, M., Santos, M.Y., Pundt, H. (Eds.), Geospatial Thinking. Springer-Verlag, Berlin, pp. 399–418.

Schokker, J., 2010. 1:600,000 Geologische overzichtskaart van Nederland, TNO Bouw en Ondergrond (Geological Survey of the Netherlands) (on CD-ROM).

Sinha, A.K., McGuinness, D.L., Fox, P., Raskin, R., Condie, K., Stern, R., Hanan, B., Seber, D., 2007. Towards a reference plate tectonics and volcano ontology for semantic scientific data integration. In: Brady, S.R., Sinha, A.K., Gundersen, L.C. (Eds.), Proceedings of the Geoinformatics 2007—Data to Knowledge, Reston, VA, USA, pp. 43–46.

Smits, P.C., Friis-Christensen, A., 2007. Resource discovery in a European spatial data infrastructure. IEEE Transactions on Knowledge and Data Engineering 19 (1), 85–95.
Soergel, D., Lauser, B., Liang, A., Fisseha, F., Keizer, J., Katz, S., 2004. Reengineering thesauri for new applications: the AGROVOC example. Journal of Digital Information 4 (4) Online Journal, 〈http://journals.tdl.org/jodi/rt/printer Friendly/112/111〉 [Accessed February 07, 2011].
Tennis, J.T., Sutton, S.A., 2008. Extending the simple knowledge organization system for concept management in vocabulary development applications. Journal of the American Society for Information Science and Technology 59 (1), 25–37.
Tsuruoka, Y., Tsujii, J., Ananiadou, S., 2008. FACTA: a text search engine for finding associated biomedical concepts. Bioinformatics 24 (21), 2559–2560.
U.S. Geological Survey Geologic Names Committee, 2010. Divisions of geologic time-major chronostratigraphic and geochronologic units: U.S. Geological Survey Fact Sheet 2010-3059, 2 pp., 〈http://pubs.usgs.gov/fs/2010/3059/pdf/FS10-3059.pdf〉 [Accessed February 07, 2011].
van Assem, M., Malaisé, V., Miles, A., Schreiber, G., 2006. A method to convert thesauri to SKOS. In: Sure, Y., Domingue, J. (Eds.), The Semantic Web: Research and Applications; Lecture Notes in Computer Science, vol. 4011. Springer-Verlag, Berlin, pp. 95–109.
Voss, J., 2009. Encoding changing country codes for the Semantic Web with ISO 3166 and SKOS. In: Sicilia, M.-A., Lytras, M.D. (Eds.), Metadata and Semantics. Springer-US, New York, pp. 211–221.
Walker, J.D., Geissman, J.W. (compilers), 2009. Geologic Time Scale: Geological Society of America, 1 pp., 〈http://www.geosociety.org/science/timescale/timescl.pdf〉 [Accessed February 07, 2011].
Woodcock, R., Simons, B., Duclaux, G., Cox, S., 2010. AuScope's use of standards to deliver earth resource data. Geophysical Research Abstracts 12, EGU2010-1556, 〈http://meetingorganizer.copernicus.org/EGU2010/EGU2010-1556-2011.pdf〉 [Accessed February 07, 2011].
Yue, P., Di, L., Yang, W., Yu, G., Zhao, P., Gong, J., 2009. Semantic Web Services-based process planning for earth science applications. International Journal of Geographical Information Science 23 (9), 1139–1163.
Zhang, C.R., Zhao, T., Li, W.D., 2010. The framework of a geospatial semantic web-based spatial decision support system for Digital Earth. International Journal of Digital Earth 3 (2), 111–134.