

# An LMF-based Web Service for Accessing WordNet-type Semantic Lexicons

Bora Savas<sup>§</sup>, Yoshihiko Hayashi<sup>§</sup>, Monica Monachini<sup>¶</sup>, Claudia Soria<sup>¶</sup>, Nicoletta Calzolari<sup>¶</sup>

<sup>§</sup>Graduate School of Language and Culture, Osaka University  
1-8 Machikaneyama, Toyonaka, 5600043, Japan  
bsavas@gs.lang.osaka-u.ac.jp, hayashi@lang.osaka-u.ac.jp

<sup>¶</sup>Istituto di Linguistica Computazionale "A. Zampolli", Consiglio Nazionale delle Ricerche  
Via Moruzzi 1 – 56126 Pisa, Italy  
{monica.monachini, claudia.soria, nicoletta.calzolari}@ilc.cnr.it

## Abstract

This paper describes a Web service for accessing WordNet-type semantic lexicons. The central idea behind the service design is: given a query, the primary functionality of lexicon access is to present a *partial lexicon* by extracting the relevant part of the target lexicon. Based on this idea, we implemented the system as a *RESTful Web service* whose input query is specified by the access URI and whose output is presented in a standardized XML data format. LMF, an ISO standard for modeling lexicons, plays the most prominent role: the access URI pattern basically reflects the lexicon structure as defined by LMF; the access results are rendered based on Wordnet-LMF, which is a version of LMF XML-serialization. The Web service currently provides accesses to Princeton WordNet, Japanese WordNet, as well as the EDR Electronic Dictionary as a trial. To accommodate the EDR dictionary within the same framework, we modeled it also as a WordNet-type semantic lexicon. This paper thus argues possible alternatives to model innately bilingual/multilingual lexicons like EDR with LMF, and proposes possible revisions to Wordnet-LMF.

## 1. Introduction

Although not every language resource is suitable for networked query-based access, many types of language resources are successfully wrapped as Web services, opening a new dimension for dissemination and utilization of language resources and technologies. Among them, lexical resource such as a semantic lexicon is an excellent target for the *Web servicization*, because the issues associated with language resources, including intellectual property right issue, can be solved/remedied by this solution.

Given this trend as the background, this paper describes a Web service for accessing WordNet-type semantic lexicons. The central idea behind the service design is: given a query, the primary functionality of lexicon access is to present a *partial lexicon* by extracting the relevant part of the target lexicon. Based on this idea, we implemented the system as a *RESTful Web service* (Richardson and Ruby, 2007) whose input (query) is specified by the access URI and whose output is represented by a standardized XML data format.

LMF (Lexical Markup Framework) (Francopoulo et al., 2008), an ISO standard (ISO 24613, 2008) for modeling lexicons, plays the most prominent role: the access URI pattern basically reflects the lexicon structure as defined by LMF; the access results are rendered based on Wordnet-LMF (Soria et al., 2009) XML schema, which is a version of LMF XML-serialization.

The Web service currently provides access to Princeton WordNet 3.0 (Fellbaum, 1998) as well as Japanese WordNet 0.9 (Bond et al., 2008). This paper also describes our trial to encode the EDR Electronic Dictionary (Yokoi, 1995; EDR, 2007) also as a WordNet-type semantic lexicon. This enables us to implement the access service by the same framework. This trial, on the other hand, revealed an issue in modeling innately bilingual/multilingual lexicons

like EDR with LMF. This paper thus argues possible alternatives to the modeling, and proposes possible revisions to Wordnet-LMF.

## 2. Web Servicizing Language Resources

### 2.1. Reasons to Web Servicize Language Resources

The notion of service in the world of software has been becoming more important as illuminated by the terms such as *SOA* (Service Oriented Architecture) or *SaaS* (Software as a Service). In parallel with this, service-oriented language infrastructures to push forward the notion of "*LRaaS*" (Language Resources as a Service) have come to the front. With a carefully designed and adequately operated service-oriented language infrastructure:

- More non-expert can have accesses to language resources (LR) and language technologies (LT) through usable access interfaces (APIs);
- Complicated intellectual property right issues can be (partly) remedied by the access control policies and mechanisms maintained by the service infrastructure;
- A virtual language resource can be realized as a language service through useful combination of the existing language services, thanks to Web service workflow technology.

### 2.2. Language Service Infrastructure and Web APIs

The Language Grid (Ishida, 2006; Murakami et al., 2010) is a multilingual language service infrastructure on the Internet whose primary goal is to provide solutions enabling the above mentioned environment. Envisaged majority of the users are non LR/LT experts who are involved in the

activities of intercultural collaboration. It currently accommodates more than sixty Web services<sup>1</sup>.

These services are classified into one of the around twenty service types, and utilized through accordingly defined APIs<sup>2</sup>. In other words, a language resource has to be *wrapped* as a Web service of the appropriate type in order to be accommodated by the infrastructure. The APIs defined by the Language Grid are carefully designed so that non-LR/LT expert users are able to use the language services relatively easily. The provided APIs however are not linguistically fine-grained, nor compliant with LR-related international standards. Thus a possible direction for the next generation APIs or APIs in another layer may be to accommodate more fine-grained linguistic data that are represented by adopting relevant LR-related international standards.

### 3. WordNet-type Semantic Lexicon

#### 3.1. What is the WordNet-type Semantic Lexicon?

We mean by *WordNet-type semantic lexicon* a lexical resource whose fundamental structure is same as the Princeton WordNet (Fellbaum, 1998) (hereafter PWN). That is, a lexicon of the type consists of: a set of synset nodes; a set of links, each connecting a synset node with another one under some lexical/conceptual relation. A synset denotes a lexicalized concept and the associated synset node gathers synonymous word forms, each representing one sense carried by a word form. More precisely, a word sense in PWN is defined by the triple {word-form, part-of-speech, sense-number}, and functions as a pointer to the associated synset.

A number of lexical resources sharing this information structure have been developed for many languages, including the Japanese WordNet (Bond et al., 2008) (hereafter WN-ja), and these lexical resources are expected to be integrated via Global WordNet Grid<sup>3</sup>. Note that, in some literatures, WordNet-type semantic lexicon is described as “wordnet” and we adopt this convention in this paper.

#### 3.2. Modeling EDR Dictionary as a WordNet-type Semantic Lexicon

The *EDR Electronic Dictionary* (Yokoi, 1995; EDR, 2007) is not a single dictionary; rather it is actually a dictionary system consisting of sub-dictionaries, including monolingual dictionaries (Japanese and English), bilingual dictionaries (J-to-E and E-to-J), and a concept dictionary, along with co-occurrence dictionaries and corpora. The EDR dictionary (hereafter EDR) is the result of a nine-year national project in Japan (from 1986 through 1994) whose aim is to establish a lexical knowledge infrastructure that is useful for intelligent information processing, including Japanese-to/from-English machine translation systems.

The core logical structure of EDR can be depicted as shown in Fig. 1.

<sup>1</sup><http://langrid.org/operation/service-manager/language-services>

<sup>2</sup><http://langrid.nict.go.jp/langrid-developers-wiki-en/#f6d501a8>

<sup>3</sup><http://www.globalwordnet.org/gwa/gwa.grid.htm>

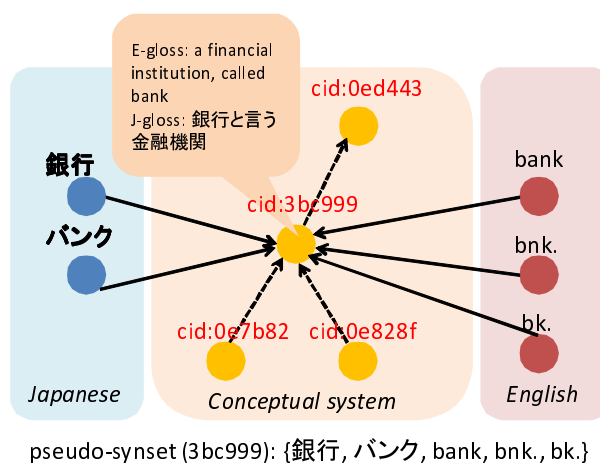


Figure 1: Core logical structure of EDR.

In EDR, each entry in every sub-dictionary is associated with a concept identifier (CID) which represents a fine-grained language-independent (or Japanese/English bilingual) concept. A CID can be referred by multiple word entries whose meanings are thought to be equivalent. For example in Fig. 1, the Japanese words (“銀行”, and “バンク”) and the English words (“bank”, “bnk.”, “bk.”) have the same CID (3bc999), showing that they all denote a same concept (the financial institution sense). This enables us to form a pseudo-synset for a concept node. Note here that the pseudo-synset can be bilingual, given a possible situation where a CID is shared by both the Japanese and English entries. As shown in Fig. 1, a concept node can also have glosses both in Japanese and English.

The concept nodes make up a kind of taxonomy or ontological structure (conceptual system in the figure) in which a node is connected to another by some conceptual/semantic relation. In short, the overall logical structure is quite similar to the PWN structure; hence EDR can be modeled as a WordNet-type semantic lexicon. This provides us an opportunity to realize an access service for EDR by exact the same framework for PWN/WN-ja.

#### 3.3. Wordnet-LMF: a Modeling Framework for WordNet-type Lexicons

Wordnet-LMF (Soria et al., 2009), developed by the EU KYOTO project<sup>4</sup> (Vossen et al., 2008), is a dialect of LMF (Lexical Markup Framework) (Francopoulo et al., 2008), which is an ISO international standard (ISO 24613, 2008) to model a broad range of lexical resources. Wordnet-LMF was especially designed to facilitate interchange of lexico-semantic information encoded in wordnets for multiple languages. The WN-ja (Bond et al., 2008) is a remarkable instance of the wordnet that demonstrates the applicability of Wordnet-LMF to encode a large-scale lexicon in a language other than English.

As suggested by the specification of LMF, it is basically expected that the multilingual associations among lexical entries across lexical resources are modeled by using the

<sup>4</sup><http://www.kyoto-project.eu/>

*Sense Axis* introduced by the multilingual notation extension, because LMF assumes that lexicons modeled with the framework are essentially monolingual. Wordnet-LMF also fully makes use of the *Sense Axis* to interrelate lexicons in different languages. It serves as a pivot to connect *Senses/Synsets* in different languages (hence, in different lexicons) that share an equivalent meaning. In fact, a synset in WN-ja is associated with the corresponding synset in PWN via a *Sense Axis* node.

Figure 2<sup>5</sup> overviews the core structure of LMF that is relevant for describing Wordnet-LMF. The top-most element is *Lexical Resource* accompanied by *Global Information* which essentially stores profile-like information. A lexical resource consists of one or more *Lexicons* and zero or more *Sense Axes*. A lexicon consists of one or more *Lexical Entries* and zero or more *Synsets*. Observing Saussure’s tradition, a lexical entry consists of form (Lemma) and *Sense*. A *Sense* node functions as a pointer to the corresponding *Synset* node. An instance of *Synset Relation* represents a conceptual relationship between synsets. The core information structure plays a crucial role in designing the Web service API.

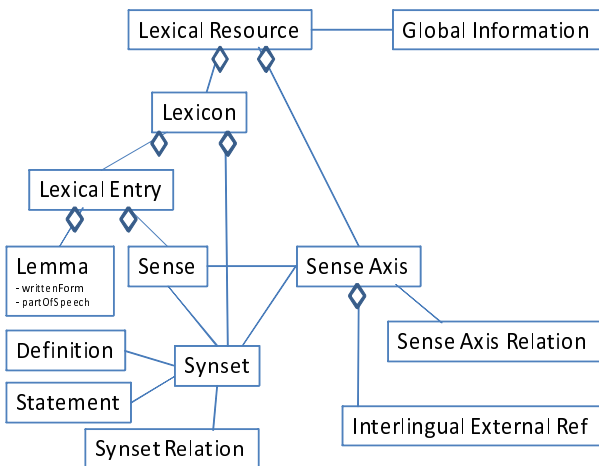


Figure 2: LMF core structure relevant to Wordnet-LMF.

## 4. LMF-based Web Service for WordNet-type Semantic Lexicons

### 4.1. Design Principle

*REST* (Representational State Transfer) (Richardson and Ruby, 2007) is a style of software architecture to realize Web services. With this architecture, inputs (requests) and outputs (responses) of a Web service are built around the transfer of *representations of resources*. Here, a resource is coherent and meaningful object that should be designated by URI (Uniform Resource Identifier); the resources should be represented in the format that is shared by the possible clients in advance.

<sup>5</sup>A diamond shape in the diagram indicates “aggregation” as in the formal UML class diagram. Cardinalities are however omitted.

Since the primary functionality of the lexicon access, given a query, is to present a partial lexicon extracted from the target lexicon, REST’s resource-centric approach is highly applicable. That is:

- Access queries are *structured* in the sense that they assume particular information structure in the target lexicon, thus can be mapped to URIs that reflect the information structure;
- The resulted partial lexicon should be represented in a pre-shared standard data format, providing an opportunity to apply LR-related international standards.

It is LMF that not only provides the standardized data format for representing access results but also presents a solid structural specification for the target lexicon type.

### 4.2. Access URI Pattern and Examples

We designed the URI pattern for invoking the Web service primarily based on the lexicon structure specified by Wordnet-LMF, and secondarily by considering possible usage patterns. Figure 3 shows the basic structure of the proposed URI pattern, which consists of the *core\_part* and the *optional\_part*.

URI pattern: `http://[server]/core_part/optional_part`  
*core\_part*: `[lexicon]/[target]/[search_by]/[qterm]`  
*optional\_part*: `?[query_parameter]=[value]`

Figure 3: Structure of the proposed URI pattern.

The organization of the *core\_part* is summarized as shown in Table 1 and Table 2. The `[qterm]` part designates a query term string.

path in the URI	possible value	LMF
<code>[lexicon]</code>	pwn, wn-ja, edr, ...	<Lexicon>
<code>[target]</code>	word synset	<LexicalEntry> <Synset>

Table 1: Organization of the *core\_part*.

<code>[target]</code> part	<code>[search_by]</code> part	LMF
word	form	writtenForm
	pos	partOfSpeech
	form-pos	writtenForm, partOfSpeech
synset	id	id
	definition	<Definition>
	statement	<Statement>

Table 2: Organization of the *search\_by* part.

Figure 4 shows the result with the access URI `/wn-ja/word/form/銀行`, showing that it has three senses. Notice here that according to the partial lexicon concept, the top element in the resulted XML

```

<LexicalResource>
  <GlobalInformation label="/scope_web/wn-ja/word/form/銀行/">
  -<Lexicon languageCoding="ISO 639-3" label="Japanese Wordnet 3.0" language="jpn" owner="NICT" version="0.9">
    -<LexicalEntry id="w211859">
      <Lemma writtenForm="銀行" partOfSpeech="n"/>
      <Sense index="1" id="w211859_02787772-n" synset="jpn-09-02787772-n"/>
      <Sense index="2" id="w211859_13368318-n" synset="jpn-09-13368318-n"/>
      <Sense index="3" id="w211859_08420278-n" synset="jpn-09-08420278-n"/>
    </LexicalEntry>
  </Lexicon>
</LexicalResource>

```

Figure 4: Result of /wn-ja/word/form/銀行.

document is `<LexicalResource>` whose immediate daughters are `<GlobalInformation>` (decorated with the access URI) and `<Lexicon>`.

Another example that would be of interest is full-text search on the gloss definition of a synset. For example, the URI `/edr/synset/definition/the capital of France/` gives us a set of synsets whose gloss text contains the string “the capital of France” as shown in Fig. 5. Notice from this example that not only the `<Synset>` elements but the associated `<LexicalEntry>` elements are included in the XML document to fulfill the structural requirement of the `<Lexicon>` element imposed by LMF. This figure also serves as a nice example of the access to EDR, where the definition of a concept is given both in English and Japanese: in Fig. 5 the Japanese `<LexicalEntry>` (`writtenForm="パリ"`), along with two other variants, and the English `<LexicalEntry>` (`writtenForm="Paris"`) simultaneously refer to a `<Synset>` (`id=edr_cph-104edc-x`), whose definition is given in the two languages.

As described above, the `core_part` in the URI pattern reflects the lexicon structure specified by LMF, while the optional `part` that utilizes the query parameter construct is introduced to support other possible usage patterns. Here we show two synset search examples: (1) by specifying a {word-form, part-of-speech, sense-number} triple, and (2) by restricting a conceptual relation.

Figure 6 shows the former example, where a PWN synset is queried by specifying the triple {`form="bank"`, `pos=noun`, `sense-number=2`}. Notice here that we introduced the query parameter (`?synset_by_index`) along with the `/word/form-pos/` construct that specifies a combination of `writtenForm` and `partOfSpeech` attributes.

Figure 7 shows the latter example, where a set of PWN synsets is queried by specifying a base synset (`id=eng-09-08420278-n`; it is the result synset shown in Fig. 6) and a conceptual relation (hyponym). Thanks to the rendering with Firefox browser, only two of eleven `<Synset>` elements are unfolded and shown.

### 4.3. A Note on the Implementation

Each target lexical resource has been stored in a relational database in advance, and the user-specified URI is mapped into the corresponding SQL statement. We employed

Python-based Django<sup>6</sup> as the back-end mechanism. It is extremely effective in implementing this type of Web service, because it provides a built-in mechanism for URI-to-resource (database entry) mapping (URL dispatcher) along with a template mechanism for generating XML documents that observe an XML schema.

## 5. Modeling Multilingual Semantic Lexicons with LMF: A Revision Proposal

As already shown, EDR is an innately bilingual lexicon, insisting that the `<Synset>` configuration can be bilingual. In fact, a substantial number of concept nodes in EDR have both Japanese and English definitions. To encode such a bilingual/multilingual synset configuration, at least two solutions can be possible:

1. allowing a `<Synset>` element to have multiple `<Definition>` elements, each of which is given in an explicitly specified language ;
2. utilizing the `Sense Axis` device provided by LMF.

The former is the solution proposed in this paper, where `lang` attribute is being introduced to `<Definition>` element in order to explicitly specify the language of a definition text. Figure 5 exemplifies the case.

On the other hand, with the latter solution, the “the capital of France” entry in EDR, as already shown in Fig. 5, can be alternatively organized as illustrated in Fig. 8.

As depicted in the figure, the original EDR concept node (`id=edr_cph-104edc-x`) is separated into the Japanese Synset node (`id=edr_cph-104edc-J`) and the English Synset node (`id=edr_cph-104edc-E`). These two nodes are then interconnected by a `Sense Axis` node. This organization may be particularly effective in case where other synsets in other languages would be further incorporated. With this organization, however, the `Sense Axis` node has to be linked to the EDR conceptual system via `Interlingual External Ref`. That is, the EDR conceptual system, now considered as an external ontology rather than a part of the dictionary system, has to be modeled and represented using some framework other than LMF.

Comparing these two solutions, it would be still natural and simple to accept the revision of LMF as proposed to represent innately bilingual/multilingual lexicons, as far as we

<sup>6</sup><http://www.djangoproject.com/>

```

<LexicalResource>
  <GlobalInformation label="/scope_web/edr/synset/definition/the capital of France/">
  - <Lexicon languageCoding="ISO 639-3" label="EDR" language="en ja" owner="NICT" version="1.0">
    - <LexicalEntry id="edr_je_276947">
      <Lemma writtenForm="パリ" partOfSpeech="JN2" pronunciation="パリ"/>
      <Sense index="1" id="x_276947_JN2" synset="edr_cph-104edc-x"/>
    </LexicalEntry>
    + <LexicalEntry id="edr_je_276949"></LexicalEntry>
    + <LexicalEntry id="edr_je_276948"></LexicalEntry>
    - <LexicalEntry id="edr_ej_15247">
      <Lemma writtenForm="Paris" partOfSpeech="EN2" pronunciation=""/>
      <Sense index="1" id="x_15247_EN2" synset="edr_cph-104edc-x"/>
      <Sense index="2" id="x_15248_EN2" synset="edr_cph-3bd10a-x"/>
    </LexicalEntry>
    - <Synset id="edr_cph-104edc-x" baseConcept="undef">
      <Definition lang="en" gloss="the capital of France, a city called Paris"/>
      <Definition lang="ja" gloss="フランスの首都のパリ"/>
      + <SynsetRelations></SynsetRelations>
    </Synset>
  </Lexicon>
</LexicalResource>

```

Figure 5: Result of /edr/synset/definition/the capital of France/.

```

<LexicalResource>
  <GlobalInformation label="/scope_web/pwn/word/form-pos/bank-n/?synset_by_index=2"/>
  - <Lexicon languageCoding="ISO 639-3" label="English Wordnet 3.0" language="eng" owner="PRINCETON" version="3.0">
    + <LexicalEntry id="w109396"></LexicalEntry>
    - <Synset id="eng-09-08420278-n" baseConcept="3">
      <Definition gloss="a financial institution that accepts deposits and channels the money into lending activities;"/>
      <Statement example="he cashed a check at the bank'; 'that bank holds the mortgage on my home"/>
    - <SynsetRelations>
      <SynsetRelation target="eng-09-08054721-n" relType="hype"/>
      <SynsetRelation target="eng-09-08234628-n" relType="hypo"/>
      <SynsetRelation target="eng-09-08350919-n" relType="hypo"/>
      <SynsetRelation target="eng-09-08418316-n" relType="hypo"/>
      <SynsetRelation target="eng-09-08418420-n" relType="hypo"/>
      <SynsetRelation target="eng-09-08418763-n" relType="hypo"/>
      <SynsetRelation target="eng-09-08418885-n" relType="hypo"/>
      <SynsetRelation target="eng-09-08419033-n" relType="hypo"/>
      <SynsetRelation target="eng-09-08419163-n" relType="hypo"/>
      <SynsetRelation target="eng-09-08419562-n" relType="hypo"/>
      <SynsetRelation target="eng-09-08422524-n" relType="hypo"/>
      <SynsetRelation target="eng-09-08423298-n" relType="hypo"/>
      <SynsetRelation target="eng-09-08066491-n" relType="hmem"/>
    </SynsetRelations>
  </Synset>
</Lexicon>
</LexicalResource>

```

Figure 6: Result of /pwn/word/form-pos/bank-n/?synset\_by\_index=2.

want the conceptual system as a part of the dictionary system. Therefore, this paper argues that the proposed revision of LMF should be considered, given a situation where inately bilingual/multilingual lexicons like EDR are to be modeled and represented.

## 6. Related Work

Access to PWN via a REST-style Web service was firstly described in (Assem et al., 2006)<sup>7</sup>, where the authors discusses a URI format and a content representation schema in

RDF/OWL. Our proposal was in part inspired by their URI proposal: for example, to designate a synset by hyphenated triples of word-form, part-of-speech, and sense number.

Soria et al. (2006) presented an architecture where distributed wordnets are linked via the central ILI (Inter-Lingual Index), and proposed a set of three Web API methods that a local wordnet has to provide. Among them, the functionalities provided by `GetSynsetById` and `GetSynsetByLemma` are also supported by the proposed Web service, while `GetWeightedSynsetsByIli` which returns a set of synsets associated with an ILI record

<sup>7</sup>They however do not use the term "REST."

```

<LexicalResource>
  <GlobalInformation label="/scope_web/pwn/synset/id/eng-09-08420278-n/?relation=hypo"/>
  - <Lexicon languageCoding="ISO 639-3" label="English Wordnet 3.0" language="eng" owner="PRINCETON" version="3.0">
    + <LexicalEntry id="w50873"></LexicalEntry>
    + <LexicalEntry id="w84396"></LexicalEntry>
    + <LexicalEntry id="w94765"></LexicalEntry>
    + <LexicalEntry id="w109396"></LexicalEntry>
  - <Synset id="eng-09-08234628-n" baseConcept="3">
    <Definition gloss="a cooperative depository financial institution whose members can obtain loans from their combined savings"/>
    - <SynsetRelations>
      <SynsetRelation target="eng-09-08420278-n" relType="hype"/>
    </SynsetRelations>
  </Synset>
  - <Synset id="eng-09-08350919-n" baseConcept="3">
    <Definition gloss="one of 12 regional banks that monitor and act as depositories for banks in their region"/>
    - <SynsetRelations>
      <SynsetRelation target="eng-09-08420278-n" relType="hype"/>
      <SynsetRelation target="eng-09-08350470-n" relType="hmem"/>
    </SynsetRelations>
  </Synset>
  + <Synset id="eng-09-08418316-n" baseConcept="3"></Synset>
  + <Synset id="eng-09-08418420-n" baseConcept="3"></Synset>
  + <Synset id="eng-09-08418763-n" baseConcept="3"></Synset>
  + <Synset id="eng-09-08418885-n" baseConcept="3"></Synset>
  + <Synset id="eng-09-08419033-n" baseConcept="3"></Synset>
  + <Synset id="eng-09-08419163-n" baseConcept="3"></Synset>
  + <Synset id="eng-09-08419562-n" baseConcept="3"></Synset>
  + <Synset id="eng-09-08422524-n" baseConcept="3"></Synset>
  + <Synset id="eng-09-08423298-n" baseConcept="3"></Synset>
</Lexicon>
</LexicalResource>

```

Figure 7: Result of /pwd/word/synset/id/eng-09-08420278-n/?relation=hypo.

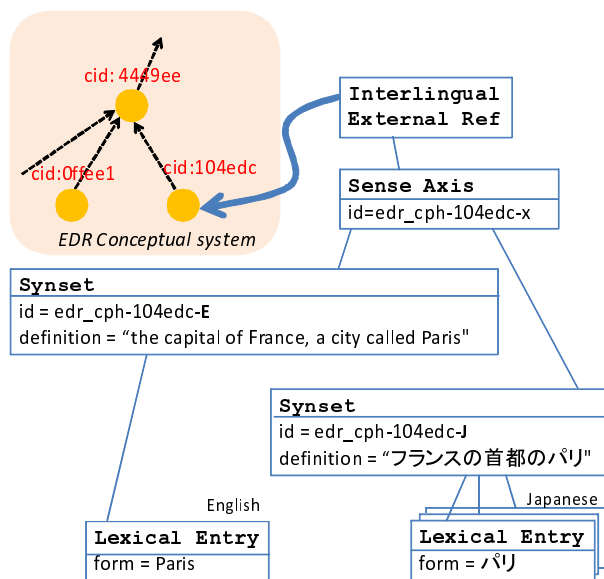


Figure 8: Alternative modeling of EDR with Interlingual External Ref.

is not. They also proposed a schema for synset representation that is employed by the Web service. The schema however is not LMF-compliant.

Recently, a language service type named *concept dictionary* has been published by the Language Grid, which

provides accesses to WordNet-type semantic lexicons, including PWN, WN-ja, and EDR. The service however is SOAP/WSDL-based, rather than REST. In addition, the data format adopted for representing the access results is not LMF-compliant.

## 7. Concluding Remarks

This paper presented a RESTful Web service for accessing WordNet-type semantic lexicons. In particular, the API was described: we discussed the design of URI structure for querying, and outlined the output data format which is based on the Wordnet-LMF schema. Possible revisions to the Wordnet-LMF schema were also proposed to accommodate innately bilingual/multilingual semantic lexicons like EDR. Other revisions to the XML schema of Wordnet-LMF include *index* attribute in the *<Sense>* element that allows the user to specify a sense number of a word form. These are however slight revisions and may not affect the generality of Wordnet-LMF as well as the original LMF. In fact, any XML document validated with the Wordnet-LMF schema is valid with our revised XML schema.

For future work, inclusion of other wordnets should be considered to demonstrate the applicability of the proposed framework. In addition, we will explore the issues associated with the notion of *on-demand composite lexicon access*. Such a Web service requires a mechanism to associate lexical entries across lexical resources on-demand/on-the-fly. To enable this, we have to develop an efficient Web service (hidden to users) to semantically align possi-

ble lexical entries. Furthermore, we will need to develop a mechanism to formally represent and store the uncovered semantic links. Fortunately LMF allows a Language Resource to gather Sense Axes, but it might force us to (re)consider the global ID issue, given a highly distributed environment.

## 8. Acknowledgments

The presented work has been supported by Strategic Information and Communications R&D Promotion Programme (SCOPE) of the Ministry of Internal Affairs and Communications of Japan.

## 9. References

- van Assem, M., Gangemi, A., and Schreiber, G. (2006). Conversion of WordNet to a Standard RDF/OWL Representation. *Proc. LREC2006*, pp. 237–242.
- Bond, F., Isahara, H., Kanzaki, K., and Uchimoto, K. (2008). Boot-strapping a WordNet Using Multiple Existing WordNets. *LREC2008*.
- EDR. (2007). The EDR Dictionary. <http://www2.nict.go.jp/r/r312/EDR/index.html>.
- Fellbaum, C. (Ed). (1998). *WordNet. An Electronic Lexical Database*. The MIT Press.
- Francopoulo, G., Bel, N., George, M., Calzolari, N., Monachini, M., Pet, M., and Soria, C. (2008). Multilingual Resources for NLP in the Lexical Markup Framework (LMF). *Language Resources and Evaluation*, Vol.43, No.1, pp. 57–70.
- Hayashi, Y., and Ishida, T. (2006). A Dictionary Model for Unifying Machine Readable Dictionaries and Computational Concept Lexicons. *Proc. LREC2006*, pp. 1–6.
- Ishida, T. (2006). Language Grid: An Infrastructure for Intercultural Collaboration. *Proc. IEEE/IPSJ Symposium on Applications and the Internet*, pp. 96–100.
- ISO 24613. (2008). Lexical Markup Framework (LMF). ISO 24613:2008.
- Murakami, Y., Lin, D., Tanaka, M., Nakaguchi, T, and Ishida, T. (2010). Language Service Management with the Language Grid. *Proc. LREC2010*.
- Richardson, L., and Ruby, S. (2007). *RESTful Web Services*. O'Reilly.
- Soria, C., Tesconi, M., Marchetti, A., Bertagna, F., Monachini, M., Huang, C.R., and Calzolari, N. (2006). Towards Agent-based Cross-lingual Interoperability of Distributed Lexical Resources. *Proc. COLING-ACL 2006 Workshop on Multilingual Language Resources and Interoperability*, pp. 17–24.
- Soria, C., Monachini, M., and Vossen, P. (2009). Wordnet-LMF: Fleshing out a Standardized Format for Wordnet Interoperability. *Proc. IWIC2009*, pp. 139–146.
- Vossen, P., Agirre, E., Calzolari, N., Fellbaum, C. Hsieh, S., Huang, C.R., Isahara, H., Kanzaki, K., Marchetti, A., Monachini, M., Neri, F., Raffaelli, R., Rigau, G., Tescon, M., and VanGent, J. (2008). KYOTO: a System for Mining, Structuring and Distributing Knowledge across Languages and Cultures. *Proc. LREC2008*.
- Yokoi, T. (1995). The EDR Electronic Dictionary. *Communications of the ACM*, Volume 38, Issue 11, pp. 42–44.