

Documentation of LMF file merger Web Service¹

Author: Muntsa Padró. Barcelona, 2012

Contact: muntsa.padro@upf.edu

1 Overview

Given two LMF files, this webservice merges them into a single LMF file. It works for LMF files encoding the information in the same way, i.e. same labels, values and structure. This will work, for example, for merging different lexica learnt under PANACEA platform. If the LMF files contain equivalent information encoded in different ways, a mapping into a common format should be previously performed.

This webservice can be used, for example, to merge the predictions of different classes learned with PANACEA (see the different available classifiers¹) or predictions of the same class for different corpora.

WARNING: this version of the webservice only works for LMF files without references. This is, LMF files containing only “<LexicalEntry>” elements. I.e. it works for morphological dictionaries, noun classification, etc but not for SCF lexica, for example. SCF lexica define first the frames associated to each verb and then the frames itself. The link between the two is made using the IDs. This version of the webservice is not able to follow such links, so it won’t perform the merging of elements outside “<LexicalEntry>” correctly. This will be developed in the future.

2 Inputs, outputs and formats

2.1 Inputs

- *input_1*, *input_2*: The two LMF files to be merged. They must encode equivalent information in the same way. For this version of the webservice, they should have only <LexicalEntry> elements (no subcats, for example).

Example:

Input1:

¹ This documented is licensed under a Creative Commons Attribution 3.0 Spain License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0/es/>.

```
<Lexicon>
  <LexicalEntry id="le_1">
    <feat att="partOfSpeech" val="noun"/>
    <Lemma>
      <feat att="writtenForm" val="boy"/>
    </Lemma>
    <Sense>
      <feat att="corpusLabel" val="labour"/>
      <feat att="hum" val="yes"/>
    </Sense>
  </LexicalEntry>
</Lexicon>
```

Input2:

```
<Lexicon>
  <LexicalEntry id="le_1">
    <feat att="partOfSpeech" val="noun"/>
    <Lemma>
      <feat att="writtenForm" val="boy"/>
    </Lemma>
    <Sense>
      <feat att="corpusLabel" val="labour"/>
      <feat att="event" val="no"/>
    </Sense>
  </LexicalEntry>
  <LexicalEntry id="le_2">
    <feat att="partOfSpeech" val="noun"/>
    <Lemma>
      <feat att="writtenForm" val="party"/>
    </Lemma>
    <Sense>
      <feat att="corpusLabel" val="labour"/>
      <feat att="event" val="yes"/>
    </Sense>
  </LexicalEntry>
</Lexicon>
```

2.2 Outputs

- *merged_LMF*: LMF file with information of the two input files merged.

Example:

```
<Lexicon>
  <LexicalEntry id="le_1">
    <feat att="partOfSpeech" val="noun"/>
    <Lemma>
      <feat att="writtenForm" val="boy"/>
    </Lemma>
```

```
<Sense>
  <feat att="corpusLabel" val="labour"/>
  <feat att="hum" val="yes"/>
  <feat att="event" val="no"/>
</Sense>
</LexicalEntry>
<LexicalEntry id="le_2">
  <feat att="partOfSpeech" val="noun"/>
  <Lemma>
    <feat att="writtenForm" val="party"/>
  </Lemma>
  <Sense>
    <feat att="corpusLabel" val="labour"/>
    <feat att="event" val="yes"/>
  </Sense>
</LexicalEntry>
</Lexicon>
```

Related Web Service:

LMF file merger Web Service: <http://lod.iula.upf.edu/resources/245>

ⁱ Available classifiers:

- Abstract nouns classifier Web Service: <http://lod.iula.upf.edu/resources/264>
- Artifact nouns classifier Web Service: <http://lod.iula.upf.edu/resources/265>
- Eventive nouns classifier Web Service: <http://lod.iula.upf.edu/resources/227>
- Human nouns classifier Web Service: <http://lod.iula.upf.edu/resources/243>
- Location nouns classifier Web Service: <http://lod.iula.upf.edu/resources/244>
- Matter nouns classifier Web Service: <http://lod.iula.upf.edu/resources/266>
- Process nouns classifier Web Service: <http://lod.iula.upf.edu/resources/267>
- Semiotic nouns classifier Web Service: <http://lod.iula.upf.edu/resources/268>
- Social nouns classifier Web Service: <http://lod.iula.upf.edu/resources/269>