

Implementing TBX-Min

Adding support for TBX-Min was intended to be a straightforward process, just as TBX-Min is a straightforward and simple Glossary format. TBX-Min is an XML format, so it is easy to process the data contained within it. The structure of TBX-Min looks like the following (some elements are optional):

```
<TBX dialect = "TBX-Min">
```

```
  <header>
```

```
    <id> Dictionary/Glossary ID </id>
```

```
    <description> Description of Glossary </description>
```

```
    <dateCreated> Timestamp formatted as ISO 8601, e.g. 2014-03-02 </dateCreated>
```

```
    <languages source="source language" target="target language" />
```

```
    <creator> Creator of dictionary </creator>
```

```
    <license> License </license>
```

```
  </header>
```

```
  <body>
```

```
    <entry id="Entry ID">
```

```
      <subjectField> Subject of Dictionary </subjectField>
```

```
        <langGroup xml:lang="source/target">
```

repeat as necessary for each language group

```
          <termGroup>
```

repeat as necessary for each language group

```
            <term> Term </term>
```

```
              <termStatus> should only use the following:  
                admitted/preferred/notRecommended/obsolete  
              </termStatus>
```

```
            <partOfSpeech> should only use the following:
```

```
              noun/properNoun/verb/adjective/adverb/other
```

```
        </partOfSpeech>
        <customer>    Customer for term    </customer>
        <note> Comments and Notes on terms </note>
    </termGroup>
</langGroup>
</entry>
</body>
</TBX>
```

Understanding this structure makes writing import/export extensions straightforward as well. This can be done by following the pseudo-code maps below. If the extension is written using **Perl**, we have developed a module called **TBX::Min** which is available on **CPAN** which can be used to both import and export with ease. More details for **TBX::Min** can be found at <http://search.cpan.org/~byutrg/TBX-Min-0.06/lib/TBX/Min.pm>. The process will be similar even without the module for any who prefer not to write the extension using Perl.

Importing TBX-Min

The following is a pseudo-code that maps out the process of importing TBX-Min:

READ input file (a TBX-Min file)

CHECK root element for "TBX-Min" (to ensure validity)

READ "Header" element

GET the following information from the header:

dictionary ID

source language

target language

timestamp

creator

license

directionality

subject

description

READ "Entry" elements (These contain Concept Entries)

FOR Concept Entries

GET Entry ID

GET Subject Field (e.g. culinary, academic, technical, etc.)

READ Language Groups

FOR Language Groups

READ Term Groups

FOR Term Groups

GET the following term related values:

term

term status

part of speech

customer

note

END FOR

END FOR

END FOR

FORMAT values according to desired usage (e.g. conversion to alternate format, glossary viewer, etc.)

END

[Exporting to TBX-Min](#)

The following is a pseudo-code that maps out the process of exporting TBX-Min:

```
CREATE xml file

MAKE xml root element '<TBX dialect = "TBX-Min">'

  MAKE '<header>' element

    SET '<id>' element with Dictionary ID

    SET '<description>' element with Description

    SET '<dateCreated>' element with Timestamp (ISO 8601, e.g. 2014-03-02)

    SET '<languages source="source language" target="target language"/>'
      where source language and target language are in ISO 639 format
      (e.g. German = 'de', English = 'en', etc.)

    SET '<creator>' element with Creator

    SET '<license>' element with License

  END header element

  MAKE '<body>' element

    FOR concept entries

      MAKE/SET '<entry id="Entry_ID">' element where "Entry_ID" is the ID value

      SET '<subjectField>' element with Subject Field

      FOR language groups

        MAKE '<langGroup xml:lang="source/target">' element,
          where source/target are in ISO 639 format and depend
          on which language group you are on.

        FOR term groups

          MAKE '<termGroup>' element

            SET the following elements:

              <term>

              <termStatus>
```

<partOfSpeech>

<customer>

<note>

END FOR (term groups)

END FOR (language groups)

END FOR (concept entries)

END TBX-MIN xml file ("</TBX>")