

TBX-Min to TBX-Basic Conversion Mapping

The conversion to TBX-Basic from TBX-Min is mostly lossless. The loss in information is in the header, which has more categories in TBX-Min than in TBX-Basic. This document will describe how information in a TBX-Min document is represented in a corresponding TBX-Basic document. You are invited to comment on the conversion strategy and suggest changes, point out errors, etc.

Converter Tool

You may use the TBX-Min to TBX-Basic converter online at this site: <http://tbxconvert.gevterm.net/tbx-min/>. Please feel free to test it and report any problems without the generated documents.

Example Documents

Below is an example of a TBX-Min document and its TBX-Basic equivalent.

TBX-Min:

```
<?xml version='1.0' encoding="UTF-8"?>
<TBX dialect="TBX-Min">
  <header>
    <id>TBX sample</id>
    <languages source="de" target="en"/>
    <description>A short sample file demonstrating TBX-
Min</description>
    <dateCreated>2013-11-12T00:00:00</dateCreated>
    <creator>Klaus-Dirk Schmidt</creator>
    <directionality>bidirectional</directionality>
    <license>CC BY license can be freely copied and
modified</license>
  </header>
  <body>
    <termEntry id="C002">
      <subjectField>animals</subjectField>
      <langSet xml:lang="de">
        <tig>
          <term>hund</term>
        </tig>
      </langSet>
      <langSet xml:lang="en">
        <tig>
          <term>dog</term>
        </tig>
      </langSet>
    </termEntry>
  </body>
</TBX>
```

```

    </termEntry>
  </body>
</TBX>

```

TBX-Basic:

```

<martif type="TBX-Basic" xml:lang="de">
  <martifHeader>
    <fileDesc>
      <titleStmt>
        <title>TBX sample</title>
      </titleStmt>
      <sourceDesc>
        <p>license: CC BY license can be freely copied
and modified</p>
        <p>directionality: bidirectional</p>
        <p>description: A short sample file demonstrating
TBX-Min</p>
        <p>creator: Klaus-Dirk Schmidt</p>
      </sourceDesc>
    </fileDesc>
    <encodingDesc>
      <p type="XCSURI">TBXBasicXCSV02.xcs</p>
    </encodingDesc>
  </martifHeader>
  <text>
    <body>
      <termEntry id="C002">
        <descrip type="subjectField">whatever</descrip>
        <langSet xml:lang="de">
          <tig>
            <term>hund</term>
          </tig>
        </langSet>
        <langSet xml:lang="en">
          <tig>
            <term>dog</term>
          </tig>
        </langSet>
      </termEntry>
    </body>
  </text>
</martif>

```

Lossy Header Conversion

The information lost is in the header. Notice how in the TBX-Min document there are special tags for date created, creator, license, directionality and description. TBX-Basic, as far as I know, does not provide separate tags for each of these (you are welcome to correct me). Therefore, in the output all of these categories are placed in the TBX-Basic `<sourceDesc>` element, pre-pended with the name of the original tag and a colon (":"). Furthermore, the `<languages>` element is not converted at all, since the languages contained in a TBX-Basic file are indicated simply through the `xml:lang` attributes or the associated XCS file.

Conversion Mapping

The table below lists all of the elements and attributes in the TBX-Min specification and how they are converted to TBX-Basic.

TBX-Min	TBX-Basic
tbx	martif
@dialect	@type
Header	
header	martifHeader
id	fileDesc/titleStatement/title
description	fileDesc/sourceDesc (append name)
creator	fileDesc/sourceDesc (append name)
directionality	fileDesc/sourceDesc (append name)
license	fileDesc/sourceDesc (append name)
Body	
body	text/body
termEntry	termEntry
subjectField	descrip @type="subjectField"
langSet	langSet
tig	tig
term	term
NoteGrp/note/noteValue	note
status	termNote @type="administrativeStatus"
customer	admin @type="customerSubset"
partOfSpeech	termNote @type="partOfSpeech"

The picklist values for part of speech are the same between TBX-Min and TBX-Basic. The status picklist values are close; the ones used in TBX-Min are the recommended display values from the TBX-Basic standard (preferred, obsolete, etc.) and map to the corresponding XCS values in TBX-Basic.