

Report on use of LMF and TMF for representing raw terms

May 2008

Monica Monachini Claudia Soria
CNR ILC



Knowledge Yielding Ontologies for Transition-based Organization

ICT 211423

1 Purpose

The purpose of this report is to investigate suitability of two available standards, i.e. TMF and LMF for the representation of raw terms.

2 Approach

In order to explore suitability of the two selected representation formats, we have used samples of raw terms lists provided by Irion and Synthema.

For illustrative purposes, we have selected the term “climate change”, in English and Italian, and represented it in the two formats. For an example of the Irion and Synthema formats, see sections 8 and 9.

As we need to make sure that all relevant information is represented, we would encourage that all partners try to implement a TMF and LMF-compliant representation of terms based on their resources so that they can get a feeling of the representational potential and limitations of both formats.

3 Introduction of TMF

TMF stands for Terminological Markup Framework (ISO 16642 2001), an international standard designed in the framework of the ISO initiatives to support the creation and use of computer applications for terminological data and exchange of such data between different applications. TMF can be described as a meta-model consisting of two levels of abstraction.

1) The most abstract level is the *meta-model*, (abstract and conceptual data model level) which supports analysis and design of terminological data at a very general level.

The terminological data model comprises the following structural nodes:

- **TDC**, the Terminological Data Collection, the top level container for all the information contained in a terminology system. This is used as a container for other containers;
- **GIS** and **CI**, respectively, Global Information Section and Complementary Information, used to contain external administrative data (the title of the file, the institution originating the file, address, copyright ...) or reference to contextual links to related text corpora;
- **TE**, the terminological entry, i.e. the term assigned to a concept. It can contain one or more language sections, depending on whether the termbase is mono- or multi-lingual.
- **LS**, Language Section, containing all the Term Sections for a terminological entry
- **TS**, the Term Section, where information about terms is held. It usually contains a single term used to designate a concept as well as any other information (e.g. definitions, contexts ...)
- **TCS**, the Term Component Section, where information about the components of a term are described. For some languages it could be useful to include information about individual words used to construct a multiword term.

2) The second level is the *data model level*, which enables the designer of a terminological data collection to make particular choices according to his/her particular needs. A specific implementation of the meta-model for terminology markup expressed in XML is called a Terminological Markup Language, **TML**.

The abstract meta-model with the various structural nodes can be instantiated in XML (cf. the DTD available on the Twiki site) by means of a the generic element <struct> which can be recursively expressed. Each structural node can be identified by means of a type attribute whose

possible values can be the identifiers of the levels in the meta-model, TDC, GIS, TE, CI, LS, TS, TCS.

A complete description of TMF can be found in the document available on the Kyoto Twiki site: http://www2.let.vu.nl/twiki/pub/Kyoto/WP02:SystemDesign/TMF_ISO16642_160802.pdf.

4 TMF for raw terms

Translation into TMF is a quite straightforward task. Each term is represented by means of a `<struct>` element specified for the attribute “type” as TE (Terminological Entry).

```
<struct type="TE" id="t_001">
```

An additional element `<feat>` can be added at this stage to encode the particular domain to which the term belongs:

```
<feat type="domain">environment</feat>
```

Another `<struct>` element delimitates the language-specific sections (type value: LS). In our case, thus, we have two `<struct>` elements of the type “LS”, one for English and the other for Italian.

```
<struct type="LS">
  <feat type="language">en</feat>
```

Inside each language-specific section, we have the term proper. This is encoded by means of a `<struct>` of type “TS” (Term Section). Thus, for English:

```
<struct type="TS" id="t_66">
  <feat type="term">climate change</feat>
```

For Italian:

```
<struct type="TS">
  <feat type="term">cambiamento climatico</feat>
```

Inside each term section, all proprietary information is then represented by means of attribute-value pairs under the form of `<feat>` elements.

```
<feat type="pos">9</feat>
<feat type="preferredForm">climate change</feat>
<feat type="cum_leb_frequency">0</feat>
<feat type="value">0</feat>
<feat type="forms">climate change</feat>
<feat type="parent_name" target="">change</feat>
<feat type="parent_form">change</feat>
<feat type="parent_id">0</feat>
<feat type="documents">2</feat>
<feat type="frequency">2</feat>
<feat type="salience">1.0407</feat>
<feat type="disjoints"></feat>
<feat type="connectivity">13</feat>
<feat type="profileMatch">0.6981</feat>
```

The element `<brack>` in TMF allows to express compound information units whose meaning is interrelated. We have used it, for instance, to represent Irion’s information about “profile”:

```

<brack>
  <feat type="profile"></feat>
  <brack>
    <feat type="MW">Geography</feat>
    <feat type="MW_SCORE">0.88</feat>
  </brack>
  <brack>
    <feat type="MW">Finance</feat>
    <feat type="MW_SCORE">0.83</feat>
  </brack>
  <brack>
    <feat type="MW">Metereology</feat>
    <feat type="MW_SCORE">0.82</feat>
  </brack>
  <brack>
    <feat type="MW">Society</feat>
    <feat type="MW_SCORE">0.75</feat>
  </brack>
</brack>

```

Multiword terms are represented in TMF through an additional <struct> element, the Term Component Section or TCS, where each component can be further specified and linguistically described:

```

<struct type="TCS">
  <feat type="grammaticalinfo">N</feat>
  <feat type="grammaticalinfo">Ms</feat>
  <feat type="component">climate</feat>
  <feat type="rank">1</feat>
  <feat type="role">modifier</feat>
</struct>
<struct type="TCS">
  <feat type="grammaticalinfo">N</feat>
  <feat type="grammaticalinfo">Ms</feat>
  <feat type="component">change</feat>
  <feat type="rank">2</feat>
</struct>

```

5 Introduction of LMF

A complete description of LMF can be found in the document available on the Kyoto Twiki site: <http://www2.let.vu.nl/twiki/bin/view/Kyoto/WP02:SystemDesign> .

6 LMF for raw terms

For this translation exercise, we have used the Multiword Expression Patterns Extension (see LMF specifications, page 73). This extension allows the representation of multiwords and their internal structure.

6.1 Representation of Terms

Single terms

Each term in an LMF lexicon is represented via a `<LexicalEntry>` object. The attribute `id` codifies its original offset:

```
<LexicalEntry id="LE_climate">
  <Lemma>
    <FormRepresentation>
      <feat att="preferredForm" val="climate"></feat>
    </FormRepresentation>
  </Lemma>
</LexicalEntry>
```

6.1.1 *<Form Representation>*

The `<FormRepresentation>` element allows to represent the preferred orthographic form of a term in a way that is more constrained as compared to TMF and/or the proprietary formats, where this information is encoded via features.

Compound Terms

6.1.2 *<ListOfComponents>* and *<Component>*

Multiword term expressions are encoded as `<LexicalEntry>` objects, but they are further specified by the child element `<ListOfComponents>`, which represents the aggregative aspect of a MWE. It has the specific function to bracket the individual components of a compound term. The object `<Component>` specifies each MWE single lexical component and for each of them represents the reference to an independent lexical entry. For instance, the MW expression “climate_change” appears as follows:

```
<LexicalEntry id="LE_climate_change" mwePattern="NN">
  <Lemma>
    <FormRepresentation>
      <feat att="preferredForm" val="climate change"></feat>
    </FormRepresentation>
  </Lemma>
  <ListOfComponents>
    <Component entry="LE_climate"></Component>
```

```

    <Component entry="LE_change"></Component>
  </ListOfComponents>
</LexicalEntry>

```

where LE_climate and LE_change are the identifiers of the independent lexical entries in the lexicon.

6.1.3 MWEPattern

The attribute “mwePattern” element is a pointer to an object which allows to express the internal syntactic structure of a MWE. This is defined as an independent object outside the lexical entry and is intended to be shared across different entries that behave similarly:

```

<MWEPattern id="NN">
  <feat att="comment" val="NOUN_NOUN"></feat>
  <MWENode>
    <feat att="syntacticConstituent" val="NP"></feat>
    <MWEEdge>
      <feat att="function" val="modifier"></feat>
    </MWEEdge>
    <MWELex>
      <feat att="rank" val="1"></feat>
    </MWELex>
  </MWENode>
</MWEPattern>

```

The MWEPattern model offers a device to represent details about the internal structure of multiword expressions and describe the functions of the individual components. For instance, the pattern exemplified above specifies that the first component syntactically behaves as a modifier.

6.1.3.1 Comments

In comparison with the proprietary formats examined, LMF modularity presents the advantage to define once a number of typical syntactic patterns for multiwords and reuse them without the need to specify them for each terminological instance that have the same internal constitution.

6.1.4 <Related Form>

In the Irion’s proprietary format, information about typical left and right raw term collocates is provided and encoded by means of a combined set of attribute-value features. In the definition of a KYOTO format for raw terms, we decided to represent this information by exploiting the <RelatedForm> object. In LMF, it was originally intended as a class of the morphological extension able to represent any word forms or morphs that can be related to a lexical entry in a variety of ways, e.g. derived forms, roots etc. We tried to adapt this object to suit the representation of raw term related forms.

The attribute `targets` specifies the <RelatedForm> in the form of an identifier pointing to the related form represented in turn as a lexical entry in the lexicon. Then, the attribute-value <feat> object contains information about the position of the collocate, i.e. if it is typically found on the right- or left hand-side of the term.

Again, this format, even if more costly, has the advantage to offer a more constrained and powerful representation of typical left and right raw terms collocate.

```

<LexicalEntry id="LE_climate_change" mwePattern="NN">
  <Lemma> </Lemma>

```

```

<RelatedForm targets="LE_sayan">
  <feat att="relation" val="LEFT"></feat>
</RelatedForm>
<RelatedForm targets="LE_location">
  <feat att="relation" val="RIGHT"></feat>
</RelatedForm>
</LexicalEntry>

```

6.1.5 <feat>

All other proprietary information, similarly to TMF, is then represented by means of a combination of attribute-value pairs of the <feat> object.

```

<LexicalEntry id="LE_climate_change" mwePattern="NN">
  <Lemma>
    <feat att="pos" val="9"></feat>
    <feat att="cum_leb_frequency" val="0"></feat>
    <feat att="value" val="0"></feat>
    <feat att="profile" val="Geography"></feat>
    <feat att="MWscore" val="0.88"></feat>
    <feat att="profile" val="Finance"></feat>
    <feat att="MWscore" val="0.83"></feat>
  </Lemma>

```

7 Considerations on LMF and TMF

Both LMF and TMF allow to specify the internal structure of MWEs through their lexical components. LMF, in addition, offers the possibility to represent the syntactic pattern, i.e. the internal syntactic structure of a MWE as well as the position and role associated with each component.

A syntactic pattern defined outside the lexical entry can be shared among lexical entries that behave in the same way and there is no need to redefine and describe it for each entry.

By “adapting” the morphological object relatedForm, LMF can univoquely express the collocates of a MWE that are further specified as for their position wrt the MWE.

8 Irion proprietary format – climate change

```
<class>
  <name><![CDATA[climate change]]></name>
    <id>66</id>
    <pos>9</pos>
  <preferred_form><![CDATA[climate change]]></preferred_form>
    <cum_leb_frequency>0</cum_leb_frequency>
    <value>0</value>
    <forms><![CDATA[climate change]]></forms>
  <parent_name><![CDATA[change]]></parent_name>
  <parent_form><![CDATA[change]]></parent_form>
    <parent_id>0</parent_id>
    <documents>2</documents>
    <frequency>2</frequency>
    <salience>1.0407</salience>
    <disjoints>
    </disjoints>
  <connectivity>13</connectivity>
    <modifiers>
      <modifier>climate</modifier>
      <modifier>climate</modifier>
    </modifiers>
  <profileMatch>0.6981</profileMatch>
  <profile>
    <MW>Geography</MW>
    <MW_SCORE>0.88</MW_SCORE>
    <MW>Finance</MW>
    <MW_SCORE>0.83</MW_SCORE>
    <MW>Metereology</MW>
    <MW_SCORE>0.82</MW_SCORE>
    <MW>Society</MW>
    <MW_SCORE>0.75</MW_SCORE>
  </profile>
  <termSource><![CDATA[]]></termSource>
  <cumfrequency_parent>5</cumfrequency_parent>
  <cumdocuments_parent>5</cumdocuments_parent>
  <siblings>3</siblings>
  <description><![CDATA[]]></description>
  <preferred_forms>
  </preferred_forms>
  <related_forms>
    <target>
      <relation>LEFT</relation>
    </target>
  <target_form><![CDATA[sayan]]></target_form>
    <concept_id></concept_id>
    <source></source>
    <targetHead></targetHead>
    <searchScore>0</searchScore>
    <editDistance>0</editDistance>
    <targetValue>0</targetValue>
    <leb_frequency>0</leb_frequency>
  </target>
  <target>
    <relation>RIGHT</relation>
  </target>
  <target_form><![CDATA[location]]></target_form>
    <concept_id></concept_id>
    <source></source>
    <targetHead></targetHead>
```



```
                <searchScore>0</searchScore>
                <editDistance>0</editDistance>
                <targetValue>0</targetValue>
                <leb_frequency>0</leb_frequency>
            </target>
        </related_forms>
        <features>
    </features>
</class>
```

9 Synthema proprietary format – cambiamento climatico

```
<sentence>
  <handle>1470</handle>
  <content>
    Prevenire o mitigare l'impatto dei cambiamenti climatici è un altro modo per
    proteggere i raccolti, così come eliminare l'uso di sostanze chimiche tossiche che
    degradano gli ecosistemi.
  </content>
  <morpho-syn>

1:prevenire{v(inf),0,9},2:o{o(coord),10,1},3:mitigare{v(inf),objnoth,12,8},4:lo{d(m,sg,def),21,2
},5:impatto{n(cn,m,sg),23,7},6:di{p(art,m,pl),31,3},7:cambiamento{n(cn,m,pl),38,11},8:climatico{
g(m,pl),50,9},9:essere{v(mod),60,1},10:uno{d(m,sg),62,2},11:altro{g(m,sg),65,5},12:modo{n(cn,m,s
g),71,4},13:per{p(sempl),76,3},14:proteggere{v(inf),80,10},15:lo{d(m,pl,def),91,1},16:raccolto{n
(cn,m,pl),96,8},18:così{g(m,sg),106,4},19:come{subconj,111,4},20:eliminare{v(inf),116,9},21:lo
{d(m,sg,def),126,2},22:uso{n(cn,m,sg),128,3},23:di{p(sempl),132,2},24:sostanza{n(cn,f,pl),135,8}
,25:chimico{g(f,pl),144,8},26:tossico{g(f,pl),156,8},27:che{n(pron(rel)),165,3},28:degradare{v(f
in,pres,pers3,pl,aff),169,9},29:lo{d(m,pl,def),179,3},30:ecosistema{n(cn,m,pl),183,10},0:cambiam
ento climatico{n(@ng),,38,21},0:sostanza chimica{n(@ng),,135,17},0:chimica
tossica{n(@ng),,144,20},0:sostanza chimiche{n(@nN),,135,17},0:sostanza chimica
tossica{n(@ngg),,135,29},0:impatto dei cambiamenti{n(@nPN),,23,26},0:uso di
sostanze{n(@nPN),,128,15},0:impatto dei cambiamenti climatici{n(@nPNG),,23,36},0:uso di sostanze
chimiche{n(@nPNG),,128,24},
  </morpho-syn>
  <functional>
    subj(n){1,9}, obj(n){5,1}, ndet{4,5}, nprep(di){7,5}, nadj{8,7}, subj(n){3,9},
    obj(n){5,3}, ndet{10,12}, nadj{11,12}, vprep(per){14,9}, obj(n){16,14}, ndet{15,16}, abscl{22,9},
    absadj{18,22}, vsubconj(come){20,18}, ndet{21,22}, nprep(di){24,22}, nadj{25,24}, nadj{26,24},
    subj(nrel){24,28}, obj(n){30,28}, ndet{29,30}
  </functional>
  <lemma>
    <C frequency="1" features="|inf|" start="0" len="9">/V/prevenire</C>
    <C frequency="1" features="|inf|objnoth|" start="12" len="8">/V/mitigare</C>
    <C frequency="1" features="|cn|m|sg|" start="23" len="7">/N/impatto</C>
    <C frequency="1" features="noun phrase" start="23" len="26">/N/impatto dei cambiamenti</C>
    <C frequency="1" features="noun phrase" start="23" len="36">/N/impatto dei cambiamenti
    climatici</C>
    <C frequency="1" features="noun phrase" start="38" len="21">/N/cambiamento climatico</C>
    <C frequency="1" features="|cn|m|pl|" start="38" len="11">/N/cambiamento</C>
    <C frequency="1" features="|m|pl|" start="50" len="9">/G/climatico</C>
    <C frequency="1" features="|mod|" start="60" len="1">/V/essere</C>
    <C frequency="1" features="|m|sg|" start="65" len="5">/G/altro</C>
    <C frequency="1" features="|cn|m|sg|" start="71" len="4">/N/modo</C>
    <C frequency="1" features="|inf|" start="80" len="10">/V/proteggere</C>
    <C frequency="1" features="|cn|m|pl|" start="96" len="8">/N/raccolto</C>
    <C frequency="1" features="|m|sg|" start="106" len="4">/G/così</C>
    <C frequency="1" features="|inf|" start="116" len="9">/V/eliminare</C>
    <C frequency="1" features="noun phrase" start="128" len="15">/N/uso di sostanze</C>
    <C frequency="1" features="|cn|m|sg|" start="128" len="3">/N/uso</C>
    <C frequency="1" features="noun phrase" start="128" len="24">/N/uso di sostanze
    chimiche</C>
    <C frequency="1" features="|cn|f|pl|" start="135" len="8">/N/sostanza</C>
    <C frequency="1" features="noun phrase" start="135" len="29">/N/sostanza chimica
    tossica</C>
    <C frequency="1" features="noun phrase" start="135" len="17">/N/sostanza chimiche</C>
    <C frequency="1" features="noun phrase" start="135" len="17">/N/sostanza chimica</C>
    <C frequency="1" features="noun phrase" start="144" len="20">/N/chimica tossica</C>
    <C frequency="1" features="|f|pl|" start="144" len="8">/G/chimico</C>
    <C frequency="1" features="|f|pl|" start="156" len="8">/G/tossico</C>
    <C frequency="1" features="|fin|pres|pers3|pl|aff|" start="169" len="9">/V/degradare</C>
    <C frequency="1" features="|cn|m|pl|" start="183" len="10">/N/ecosistema</C>
  </lemma>
</sentence>
```

10 TMF Representation of climate change, cambiamento climatico

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE tmf SYSTEM "tmf.dtd">
<tmf>
  <struct type="TE" id="t_001">
    <!-- in this example we assume that the various language dependent terms are wrapped
under the same TE section.
-->
    <!-- alternatively, and ID can be placed inside the language-specific section -->
    <!-- all values of the "type" attribute in the example are taken as they are from the original formats.
In case of adoption TMF as encoding format, they should be mapped against the ISO Data Category
Registry 12620 -->
    <feat type="domain">environment</feat>
    <struct type="LS">
      <feat type="language">en</feat>
      <!-- English example based on data provided by Irion -->
      <!-- In this case, ID value corresponds to the one from the Irion tool -->
      <struct type="TS" id="t_66">
        <feat type="term">climate change</feat>
        <feat type="pos">9</feat>
        <feat type="preferredForm">climate change</feat>
        <feat type="cum_leb_frequency">0</feat>
        <feat type="value">0</feat>
        <feat type="forms">climate change</feat>
        <!-- wouldn't it be nice if reference to parent name could be via
IDREF? The attribute "target" allows it -->
        <feat type="parent_name" target="">change</feat>
        <feat type="parent_form">change</feat>
        <feat type="parent_id">0</feat>
        <feat type="documents">2</feat>
        <feat type="frequency">2</feat>
        <feat type="saliency">1.0407</feat>
        <feat type="disjoints"></feat>
        <feat type="connectivity">13</feat>
        <feat type="profileMatch">0.6981</feat>
      <brack>
        <feat type="profile"></feat>
      <brack>
        <feat type="MW">Geography</feat>
        <feat type="MW_SCORE">0.88</feat>
      </brack>
      <brack>
        <feat type="MW">Finance</feat>
        <feat type="MW_SCORE">0.83</feat>
      </brack>
      <brack>
        <feat type="MW">Metereology</feat>
        <feat type="MW_SCORE">0.82</feat>
      </brack>
    </struct>
  </struct>
</tmf>
```

```

        <brack>
            <feat type="MW">Society</feat>
            <feat type="MW_SCORE">0.75</feat>
        </brack>
    </brack>
    <feat type="termSource"> </feat>
    <feat type="cumfrequency_parent">5</feat>
    <feat type="cumdocuments_parent">5</feat>
    <feat type="siblings">3</feat>
    <feat type="description"><![CDATA[]]></feat>
    <feat type="preferredForms"></feat>
    <brack>
        <feat type="related_forms"></feat>
    <brack>
        <feat type="target"></feat>
        <brack>
            <feat type="relation">LEFT</feat>
            <feat
type="target_form"><![CDATA[sayan]]></feat>
            <feat type="concept_id"></feat>
            <feat type="source"></feat>
            <feat type="targetHead"></feat>
            <feat type="searchScore">0</feat>
            <feat type="editDistance">0</feat>
            <feat type="targetValue">0</feat>
            <feat type="leb_frequency">0</feat>
        </brack>
    </brack>
    <brack>
        <feat type="target"></feat>
        <brack>
            <feat type="relation">RIGHT</feat>
            <feat
type="target_form"><![CDATA[location]]></feat>
            <feat type="concept_id"></feat>
            <feat type="source"></feat>
            <feat type="targetHead"></feat>
            <feat type="searchScore">0</feat>
            <feat type="editDistance">0</feat>
            <feat type="targetValue">0</feat>
            <feat type="leb_frequency">0</feat>
        </brack>
    </brack>
    <feat type="context">...</feat>
    <struct type="TCS">
        <feat type="grammaticalinfo">N</feat>
        <feat type="grammaticalinfo">Ms</feat>
        <feat type="component">climate</feat>
        <feat type="rank">1</feat>

```

<!-- the following feat shifts on the term component section an information that was placed differently in the

format-->

<feat type="role">modifier</feat>

</struct>

<struct type="TCS">

<feat type="grammaticalinfo">N</feat>

<feat type="grammaticalinfo">Ms</feat>

<feat type="component">change</feat>

<feat type="rank">2</feat>

</struct>

</struct>

</struct>

<struct type="LS">

<feat type="language">it</feat>

<!--Italian example based on data provided by Synthema -->

<struct type="TS">

<feat type="term">cambiamento climatico</feat>

<feat type="frequency">3</feat>

<feat type="grammaticalinfo">N</feat>

<feat type="context">obiettivo della Convenzione delle Nazioni

Unite sui Cambiamenti Climatici</feat>

<feat type="context">La contabilità e il monitoraggio delle risorse

sono essenziali per combattere i cambiamenti climatici, </feat>

<feat type="context">Prevenire o mitigare l'impatto dei

cambiamenti climatici è un altro modo per proteggere i raccolti</feat>

<struct type="TCS">

<feat type="grammaticalinfo">N</feat>

<feat type="grammaticalinfo">Ms</feat>

<feat type="component">cambiamento</feat>

<feat type="rank">1</feat>

</struct>

<struct type="TCS">

<feat type="grammaticalinfo">A</feat>

<feat type="grammaticalinfo">Ms</feat>

<feat type="component">climatico</feat>

<feat type="rank">2</feat>

</struct>

</struct>

</struct>

<struct type="TE">

<feat type="id">1</feat>

<feat type="domain">environment</feat>

<struct type="LS">

<feat type="language">en</feat>

<struct type="TS">

<feat type="term">biodiversity</feat>

<feat type="grammaticalinfo">N</feat>

<feat type="context">...</feat>

</struct>

```
</struct>
<struct type="LS">
  <feat type="language">it</feat>
  <struct type="TS">
    <feat type="term">biodiversità</feat>
    <feat type="context">La nostra Associazione sta lavorando anche
con i governi che cercano di arginare la perdita di biodiversità</feat>
    <feat type="context">L'altro indice di questo rapporto, l'indice del
pianeta vivente, mostra una perdita rapida e continua di biodiversità</feat>
    <feat type="context">Per mantenere la biodiversità è indispensabile
che una parte della capacità produttiva della biosfera sia messa a disposizione per la
sopravvivenza delle altre specie:</feat>
    <feat type="grammaticalinfo">N</feat>
    <feat type="grammaticalinfo">Fs</feat>
  </struct>
</struct>
</struct>
</struct>
</tmf>
```

11 LMF Representation of climate change

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE LexicalResource SYSTEM "LMF_rev16.dtd">
<LexicalResource>
  <GlobalInformation/>
  <Lexicon>
<LexicalEntry id="LE_climate">
  <Lemma>
    <FormRepresentation>
      <feat att="preferredForm" val="climate"></feat>
    </FormRepresentation>
  </Lemma>
</LexicalEntry>
<LexicalEntry id="LE_change">
  <Lemma>
    <feat att="wordform" val="change"></feat>
  </Lemma>
</LexicalEntry>
<LexicalEntry id="LE_sayan">
  <Lemma>
    <feat att="wordform" val="sayan"></feat>
  </Lemma>
</LexicalEntry>
<LexicalEntry id="LE_location">
  <Lemma>
    <feat att="wordform" val="location"></feat>
  </Lemma>
</LexicalEntry>
<LexicalEntry id="LE_climate_change" mwePattern="NN">
  <Lemma>
    <feat att="pos" val="9"></feat>
    <feat att="cum_leb_frequency" val="0"></feat>
    <feat att="value" val="0"></feat>
  </Lemma>
  <!-- In the following, a combination of att-val pairs allows to encode what in TMF is encoded via
  the brack structures. -->
    <feat att="profile" val="Geography"></feat>
    <feat att="MWscore" val="0.88"></feat>
    <feat att="profile" val="Finance"></feat>
    <feat att="MWscore" val="0.83"></feat>
  <FormRepresentation>
    <feat att="preferredForm" val="climate change"></feat>
  </FormRepresentation>
  </Lemma>
  <ListOfComponents>
    <Component entry="LE_climate"></Component>
    <Component entry="LE_change"></Component>
  </ListOfComponents>
  <RelatedForm targets="LE_sayan">
    <feat att="relation" val="LEFT"></feat>
  </RelatedForm>
</LexicalEntry>
</Lexicon>
</LexicalResource>
```

```
        </RelatedForm>
        <RelatedForm targets="LE_location">
        <feat att="relation" val="RIGHT"></feat>
        </RelatedForm>
</LexicalEntry>
<!-- description of the MWE syntactic pattern that can be shared by different Lex Entries-->
<MWEPattern id="NN">
<feat att="comment" val="NOUN_NOUN"></feat>
  <MWENode>
    <feat att="syntacticConstituent" val="NP"></feat>
    <MWEEdge><feat att="function"
val="modifier"></feat></MWEEdge>
    <MWELex><feat att="rank" val="1"></feat></MWELex>
  </MWENode>
</MWEPattern>
</Lexicon>
</LexicalResource>
```