Appendix: DMLex Examples (Draft)

OASIS LEXIDMA

This document gives examples which show how to use DMLex to model lexicographic resources. The examples are shown in three formalisms: NVH, XML and JSON.

Each example is shown in NVH first. NVH (Name-Value Hierarchy)\(^1\) is a concise serialization language designed for lexicographic data. NVH encodes data as a hierarchical list of names, values and children, which corresponds exactly to DMLex’s own data model. We use NVH here in order to demonstrate the object model at an abstract level.

After that, each example is shown in XML and JSON, two popular serialization languages. The XML and JSON encoding shown here follows DMLex’s own implementation guidance for XML and JSON.

Example 0: A basic entry

This is a basic, beginner-level example of how to use DMLex to represent a simple monolingual lexicographic resource consisting of one entry with two senses. It demonstrates some of the basic features of DMLex Core: how to subdivide a entry into senses, how attach various data such as definition, part-of-speech labels to entries and senses, and how to add labels to various objects such as senses and examples.

NVH

\begin{verbatim}
lexicographicResource: my-dictionary
entry: abandon-verb
  headword: abandon
  partOfSpeech: verb
  sense: abandon-verb-1
    definition: to suddenly leave a place or a person
    example: I'm sorry I abandoned you like that.
    example: Abandon ship!
    label: idiom
  sense: abandon-verb-2
\end{verbatim}

\(^1\)https://www.namevaluehierarchy.org/
**label:** mostly-passive

**definition:** to stop supporting an idea

**example:** That theory has been abandoned.

XML

```xml
<lexicographicResource id="my-dictionary">
  <entry id="abandon-verb">
    <headword>abandon</headword>
    <partOfSpeech value="verb"/>
    <sense id="abandon-verb-1">
      <definition>to suddenly leave a place or a person</definition>
      <example>
        <text>I'm sorry I abandoned you like that.</text>
      </example>
      <example>
        <text>Abandon ship!</text>
      </example>
    </sense>
    <sense id="abandon-verb-2">
      <label value="mostly-passive"/>
      <definition>to stop supporting an idea</definition>
      <example>
        <text>That theory has been abandoned.</text>
      </example>
    </sense>
  </entry>
</lexicographicResource>
```

JSON

```json
{
  "id": "my-dictionary",
  "entry": {
    "id": "abandon-verb",
    "headword": "abandon",
    "partOfSpeech": ["verb"],
    "senses": [{
      "id": "abandon-verb-1",
      "definitions": [{
        "text": "to suddenly leave a place or a person"
      }],
      "examples": [{
        "text": "I'm sorry I abandoned you like that."
      }]
    }]
  }
}
"text": "Abandon ship!",
"labels": ["idiom"]
}]
}, {
  "id": "abandon-verb-2",
  "labels": ["mostly-passive"],
  "definitions": ["to stop supporting an idea"],
  "examples": [ {
    "text": "That theory has been abandoned."
  }]
}]
}

Example 1: How to use inflectedForm

This is an entry from a hypothetical Irish dictionary for the headword “folúshglantóir” (“vacuum cleaner”) which gives its two inflected forms, the singular genitive and the plural.

NVH

entry: folúshglantóir-n
  headword: folúshglantóir
  partOfSpeech: n-masc
  inflectedForm: folúshglantóra
    inflectedTag: sg-gen
  inflectedForm: folúshglantóirí
    inflectedTag: pl
  sense: ...

XML

<entry id="folúshglantóir-n">
  <headword>folúshglantóir</headword>
  <partOfSpeech value="n-masc"/>
  <inflectedForm inflectedTag="sg-gen">folúshglantóra</inflectedForm>
  <inflectedForm inflectedTag="pl">folúshglantóirí</inflectedForm>
  <sense>...</sense>
</entry>

JSON

3
Example 2: Pronunciation given as transcription

NVH

code: aardvark-noun
    headword: aardvark
    pronunciation: 
      transcription: a:rdva:rk
    sense: ...

XML: TBD
JSON: TBD

Example 3: Pronunciation given as a sound file

NVH

code: aardvark-noun
    headword: aardvark
    pronunciation: 
      soundFile: aardvark.mp3
    sense: ...

XML: TBD
JSON: TBD
Example 4: Pronunciation given both ways

NVH

**entry**: aardvark-noun
  **headword**: aardvark
  **pronunciation**:
    **transcription**: a:rdva:rk
    **soundFile**: aardvark.mp3
  **sense**: ...

XML: TBD
JSON: TBD

Example 5: How to use tag

This is an entry from a hypothetical Irish dictionary for the headword “folúsghlantóir” (“vacuum cleaner”). The meaning of the various tags used in this entry is explained in the tag objects.

NVH

**entry**: folúsghlantóir-n
  **headword**: folúsghlantóir
  **partOfSpeech**: n-masc
  **inflectedForm**: folúsghlantóra
    **inflectedTag**: sg-gen
  **inflectedForm**: folúsghlantóirí
    **inflectedTag**: pl
  **sense**: ...

**tag**: n-masc
  **description**: noun, masculine
  **target**: partOfSpeech

**tag**: n-fem
  **description**: noun, feminine
  **target**: partOfSpeech

**tag**: sg-gen
  **description**: singular genitive
  **target**: inflectedTag
  **partOfSpeechConstraint**: n-masc
  **partOfSpeechConstraint**: n-fem
Example 6: Mapping tag to external inventories

This shows how to map the value of a tag such as n-masc and n-fem to items in an external inventory such as LexInfo.

```xml
<tag>
  <description>noun, masculine</description>
  <target>partOfSpeech</target>
  <sameAs>http://www.lexinfo.net/ontology/3.0/lexinfo#noun</sameAs>
  <sameAs>http://www.lexinfo.net/ontology/3.0/lexinfo#masculine</sameAs>
</tag>
```

```xml
<tag>
  <description>noun, feminine</description>
  <target>partOfSpeech</target>
  <sameAs>http://www.lexinfo.net/ontology/3.0/lexinfo#noun</sameAs>
  <sameAs>http://www.lexinfo.net/ontology/3.0/lexinfo#feminine</sameAs>
</tag>
```

XML: TBD
JSON: TBD

Example 7: Defining a bilingual lexicographic resource

This defines a lexicographic resource where the source language is German and the translation language is English and the English translations are going to come with pronunciation transcriptions in English IPA.

```xml
<lexicographicResource>
  <description>My German-English Dictionary</description>
  <language>de</language>
  <translationLanguage>en</translationLanguage>
</lexicographicResource>
```

XML: TBD
JSON: TBD
Example 8: How to use headwordTranslation

This is an entry from a hypothetical English-German dictionary for English-speaking learners of German.

NVH

entry: doctor-n
  headword: doctor
  sense: doctor-n-1
    indicator: medical doctor
    headwordTranslation: Arzt
    part0fSpeech: n-masc
    headwordTranslation: Ärztin
    part0fSpeech: n-fem
  sense: doctor-n-2
    indicator: academic title
    headwordTranslation: Doktor
    part0fSpeech: n-masc
    headwordTranslation: Doktorin
    part0fSpeech: n-fem
    label: rare

Example 9: How to use headwordExplanation

NVH

entry: treppenwitz
  headword: Treppenwitz
  part0fSpeech: n-masc
  sense: treppenwitz-1
    headwordExplanation: belated realisation of what one could have said
    headwordTranslation: staircase wit
Example 10: Defining a multilingual lexicographic resource

This defines a lexicographic resource where the source language is Irish and the translation languages are English, German and Czech.

```
NVH

lexicographicResource: irish-multilingual
  description: My Irish-Multilingual Dictionary
  language: ga
  translationLanguage: en
  translationLanguage: de
  translationLanguage: cs
```

Example 11: How to use headwordTranslation in a multilingual lexicographic resource

This is an entry from a hypothetical Irish-multilingual dictionary.

```
NVH

entry: fómhar-n
  headword: fómhar
  partOfSpeech: n-masc
  inflectedForm: fómhair
    inflectedTag: genitive-case
  sense: fómhar-n-1
    headwordTranslation: autumn
      language: en
    headwordTranslation: fall
      language: en
    headwordTranslation: Herbst
      language: de
    headwordTranslation: podzim
      language: cs
```
Example 12: Modelling parts and wholes

We have three entries with one sense each: “glasses”, “microscope” and “lens”. We want to represent the fact that “lens” is a meronym of both “glasses” and “microscope”, and simultaneously that “glasses” and “microscope” are both holonyms of “lens”.

**NL**

```xml
<lexicographicResource language="en">
  <entry headword="glasses">
    <sense id="glasses-1">
      <definition>an optical seeing aid</definition>
    </sense>
  </entry>
  <entry headword="microscope">
    <sense id="microscope-1">
      <definition>equipment for looking at very small things</definition>
    </sense>
  </entry>
  <entry headword="lens">
    <sense id="lens-1">
      <definition>curved glass that makes things seem bigger</definition>
    </sense>
  </entry>
  <relation type="meronymy">
    <member id="glasses-1" role="whole"/>
    <member id="lens-1" role="part"/>
  </relation>
  <relation type="meronymy">
    <member id="microscope-1" role="whole"/>
    <member id="lens-1" role="part"/>
  </relation>
</lexicographicResource>
```
member: lens-1  
  role: part

relationType: meronomy
  description: used for modelling part-whole relationships
  memberRole: whole
    description: the whole
    memberType: sense
    min: 1
    max: 1
    action: navigate
  memberRole: part
    description: the part
    memberType: sense
    min: 1
    max: 1
    action: navigate

XML: TBD
JSON: TBD

Suggested rendering of the entry “lens” for human users:

lens

  • curved glass that makes things seem bigger
  things that contain lens: glasses, microscope

Example 13: Modelling antonyms

We have two entries for the verbs “buy” and “sell” with one sense each. We want to express the fact that the senses are antonyms.

NVH

lexicographicResource:
  language: en
  entry: buy
    headword: buy
    sense: buy-1
      definition: get something by paying money for it
  entry: sell
    headword: sell
    sense: see-1
**definition:** exchange something for money

**relation:** ants
  - member: buy-1
  - member: sell-1

**relationType:** ants
  - description: antonyms
  - memberRole:
    - memberType: sense
    - min: 2
    - max: 2
    - action: navigate

XML: TBD
JSON: TBD

Suggested rendering of the entry “buy” for human users:

**buy**

- get something by paying money for it
  
  **opposite meaning:** sell

**Example 14: Modelling synonyms**

We have three German entries with one sense each, two which mean “sea” and one which means “ocean”. We want to set up a relation which brings these three sense together as near-synonyms.

**NVH**

*lexicographicResource:*
  - language: de
  - translationLanguage: en

*entry: die-see*
  - headword: See
  - partOfSpeech: n-fem
  - sense: die-see-1
    - headwordTranslation: see

*entry: das-meer*
  - headword: Meer
partOfSpeech: n-neut
sense: das-meer-1

headwordTranslation: see

entry: der-ozean
headword: Ozean
partOfSpeech: n-masc
sense: der-ozean-1
translation: ocean

relation: syns
description: words that mean sea and ocean
member: die-see-1
member: das-meer-1
member: der-ozean-1

relationType: syns
description: synonyms and near synonyms
memberRole:
  memberType: sense
  min: 2
  action: navigate

XML: TBD
JSON: TBD

Suggested rendering of the entry “See” for human users:

See feminine noun
  - sea
    same or similar meaning: Meer, Ozean

Example 15: Modelling variants

We have two entries in our lexicographic resource, one for the headword “colour” and one for the headword “color”. We want to create a relation to represent the fact that these are spelling variants.

NVH

lexicographicResource:
  language: en
  entry: colour
Example 16: Modelling subsenses

We have an entry for the noun “colour” with four senses. We want to express the fact that senses number two and three are subsenses of sense number one, and should be displayed as such to human users.

NVH
entry: colour
  headword: colour
  sense: colour-1
    definition: red, blue, yellow etc.
    example: What is your favourite colour?
  sense: colour-2
    definition: not being black and white
    example: Back then owning a colour TV meant you were rich.
  sense: colour-3
    definition: a sign of a person's race
    example: We welcome people of all creeds and colours.
  sense: colour-4
    definition: interest or excitement
    example: Examples add colour to your writing.

relation: subsensing
  member: colour-1
    role: supersense
    member: colour-2
    role: subsense
relation: subsensing
  member: colour-1
    role: supersense
    member: colour-3
    role: subsense

relationType: subsensing
  description: expresses the fact that a sense is a subsense of another sense
  scope: sameEntry
  memberRole: supersense
    memberType: sense
    min: 1
    max: 1
    action: none
  memberRole: subsense
    memberType: sense
    min: 1
    max: 1
    action: embed

XML: TBD
Suggested rendering of the entry for human users:

colour

1. red, blue, yellow etc.
   What is your favourite colour?
   a. not being black and white
      Back then owning a colour TV meant you were rich.
   b. a sign of a person’s race
      We welcome people of all creeds and colours.

2. interest or excitement
   Examples add colour to your writing.

Example 17: Modelling subentries (at subsense level)

We have an entry for the adjective “safe” with two senses, and an entry for the multi-word expression “better safe than sorry” with one sense. We want to express the fact that the multi-word entry should appear under the first sense of “safe” as a subentry.

NVH

lexicographicResource:
   language: en

entry: safe
   headword: safe
   sense: safe-1
      indicator: protected from harm
      example: It isn't safe to park here.
   sense: safe-2
      indicator: not likely to cause harm
      example: Is the ride safe for a small child?

entry: better-safe
   headword: better safe than sorry
   sense:
      definition: you should be careful even if it seems unnecessary

relation: subentrying
   membership: safe-2
   role: container
   membership: better-safe
role: subentry

relationType: subentrying
  scope: sameResource
  memberRole: container
    memberType: sense
    min: 1
    max: 1
    action: navigate
  memberRole: subentry
    memberType: entry
    min: 1
    max: 1
    action: embed

XML: TBD
JSON: TBD

Suggested rendering of the entry “safe” for human users:

safe

1. protected from harm: *It isn’t safe to park here.*
   - **better safe than sorry** you should be careful even if it seems unnecessary
2. not likely to cause harm: *Is the ride safe for a small child?*

Suggested rendering of the entry “better safe than sorry” for human users:

better safe than sorry

- you should be careful even if it seems unnecessary

see also: safe

**Example 18: Modelling subentries (at sense level)**

We have an entry for the word “bible” and another entry for the expression “the Bible”. We want to make sure that, when a human user is viewing the entry for “bible”, the entry for “the Bible” is shown as a subentry of it, as if it were its first sense.

NVH
entry: the-bible
  headword: the Bible
  Sense: the-bible-1
    definition: the book considered holy by Christians

entry: bible
  headword: bible
  sense: bible-1
  sense: bible-2
    definition: a book considered important for a subject

relation: subentrying
  member: bible-1
    role: container
  member: the-bible
    role: subentry

relationType: subentrying
  scope: sameResource
  memberRole: container
    memberType: sense
    min: 1
    max: 1
    action: navigate
  memberRole: subentry
    memberType: entry
    min: 1
    max: 1
    action: embed

XML: TBD

JSON: TBD

Suggested rendering of the entry “bible” for human users:

bible

1. the Bible the book considered holy by Christians
2. a book considered important for a subject

Suggested rendering of the entry “the Bible” for human users:
the Bible

- the book considered holy by Christians

see also: bible

**Example 19: Using placeholderMarker**

NVH

```xml
<lexicographicResource>
  <language>en</language>
  <entry>continue-studies</entry>
    <headword>continue your studies</headword>
    <placeholderMarker>your</placeholderMarker>
    <sense>...</sense>
</lexicographicResource>
```

XML: TBD
JSON: TBD

**Example 20: Using placeholderMarker in a bilingual lexicographic resource**

NVH

```xml
<lexicographicResource>
  <language>en</language>
  <translationLanguage>de</translationLanguage>
  <entry>beat-up</entry>
    <headword>beat sb. up</headword>
    <placeholderMarker>sb.</placeholderMarker>
    <sense>beat-up-1</sense>
    <headwordTranslation>jemanden verprügeln</headwordTranslation>
    <placeholderMarker>jemanden</placeholderMarker>
</lexicographicResource>
```

XML: TBD
JSON: TBD
Example 21: Using headwordMarker

NVH

lexicographicResource:
  language: en
  translationLanguage: cs
  entry: autopsy
    headword: autopsy
    sense: autopsy-1
    headwordTranslation: pitva
    example: The coroner performed an autopsy.
      headwordMarker: autopsy
      exampleTranslation: Koroner provedl pitvu.
      headwordMarker: pitvu

XML: TBD
JSON: TBD

Example 22: Using collocateMarker

NVH

lexicographicResource:
  language: en
  translationLanguage: cs
  entry: autopsy
    headword: autopsy
    sense: autopsy-1
    headwordTranslation: pitva
    example: The coroner performed an autopsy.
      collocateMarker: performed
      lemma: perform
      exampleTranslation: Koroner provedl pitvu.
      collocateMarker: provedl
      lemma: provést

XML: TBD
JSON: TBD