



DCR and LMF services

Lirics

Marc Kemps-Snijders

MPI for Psycholinguistics

Adam Funk

University of Sheffield



Topics

Lirics

- Data Category Registry (DCR) web service interface
- LMF web service interface
- GATE LMF plugin



DCR web service

Lirics



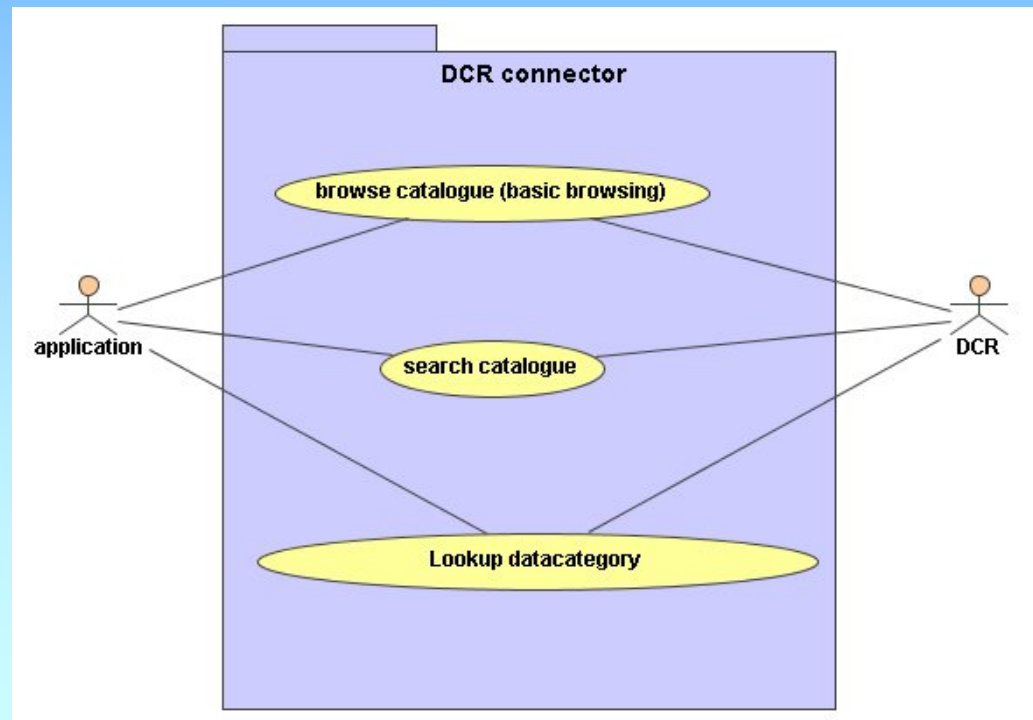
Interface to DCR service

Lirics

- Purpose is to achieve a higher degree of interoperability between linguistic resources by providing an interface for external applications.

• Use cases

- Browse DCR
- Search DCR
- Lookup datacategory

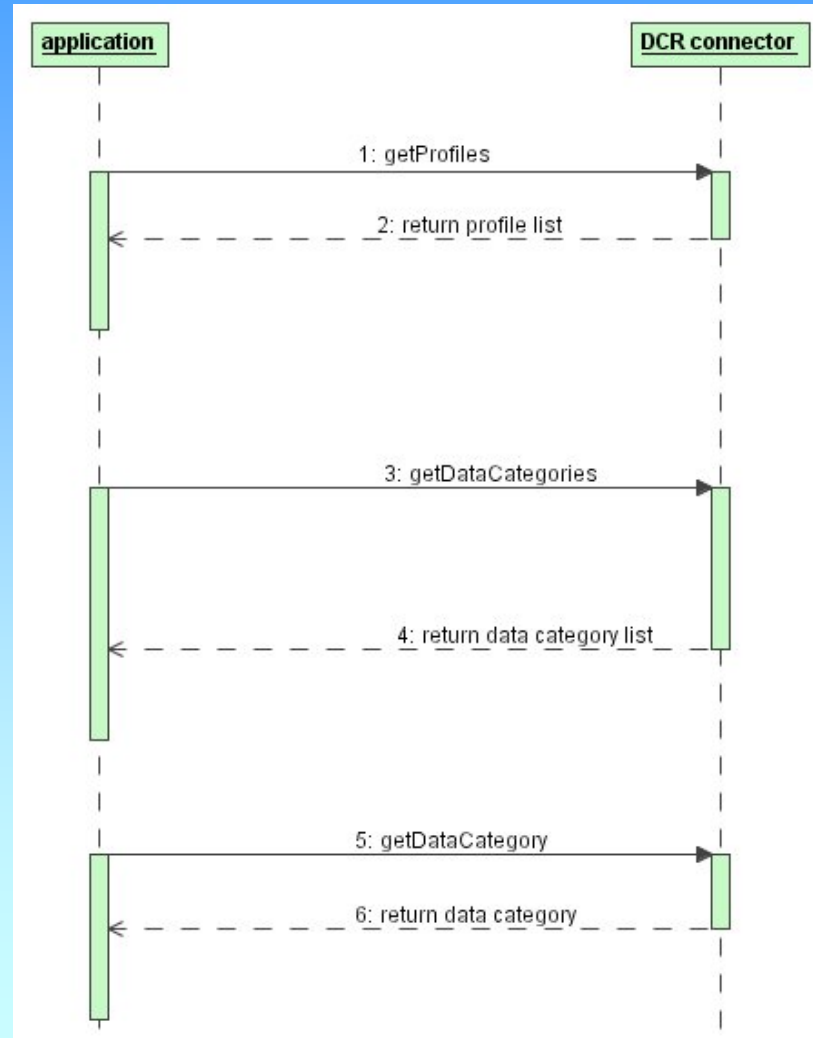




Browsing the DCR

Lirics

- Application accesses DCR and is retrieves a list of available profiles
- User selects a profile of interest.
- User selects datacategory from the list of datacategories for the selected profile





Browsing the DCR

Lirics

- function List getProfiles()
 - Returns a list of all profiles (e.g. metamodel, terminology, morphosyntax etc.)
- function List getDataCategories(a_profile,a _registrationStatus)
 - Returns a list of all datacategories associated with the specified profile having the specified registration status
- function DataCategory getDataCategory(URID)
 - Returns a datacategory identified by the specified ID



Browsing the DCR

Lirics

LEXUS Workspace Lexical Resource Search DCR gaby c

Data Category Repository:

Search >>

Profile: Private Registration status: All

Ident	Category	Status	Registration Authority
	Private		
	CategoryAdministration		
	MetaModel		
	Level		
	Terminology		
	LanguageDescription		
	MorphoSyntax		
	Semantic		
	Personal		
	Dialog		
	BJProfile		



Browsing the DCR

Lirics

LEXUS Workspace Lexical Resource Search DCR gaby c

Data Category Repository:

Search >>

Profile: MorphoSyntax Registration status: All

Identifier	Registration status	Registration Authority	
logicalValue	candidate	Private	
infinitiveParticle	candidate	Private	
addComponentStem	candidate	Private	
participle	candidate	Private	
numeralLetter	candidate	Private	
demonstrativePronoun	candidate	Private	
superlativeParticle	candidate	Private	
interrogativeParticle	candidate	Private	
nonSpacedPinyin	candidate	Private	
dativeCase	candidate	Private	
closeParenthesis	candidate	Private	
unclassifiedPunctuation	candidate	Private	
voice	candidate	Private	
circumposition	candidate	Private	
imperfective	candidate	Private	
definite	candidate	Private	
indefinitePronoun	candidate	Private	
definiteArticle	candidate	Private	



Browsing the DCR

Lyrics

DataCategory details homophone - Microsoft Internet Explorer provided by MPI Nijmegen

Data category information

homophone
(*MorphoSyntax*)

Administration Identification

Identifier: homophone Creation date: 2004-07-09
 Version: 0.0.0 ?
 Registration authority: Private
 Registration status: candidate Last change date: 0000-00-00
 Administration status: Private ?
 Origin: ?

Description

Profile: MorphoSyntax

Definition	Source	Note
Un mot qui se prononce de la même manière qu'un autre mot mais qui s'écrit différemment.	Gil Francopoulo	

Language section english

Name section		
Name	Status	Note
homophone	standardizedName	

Language section french

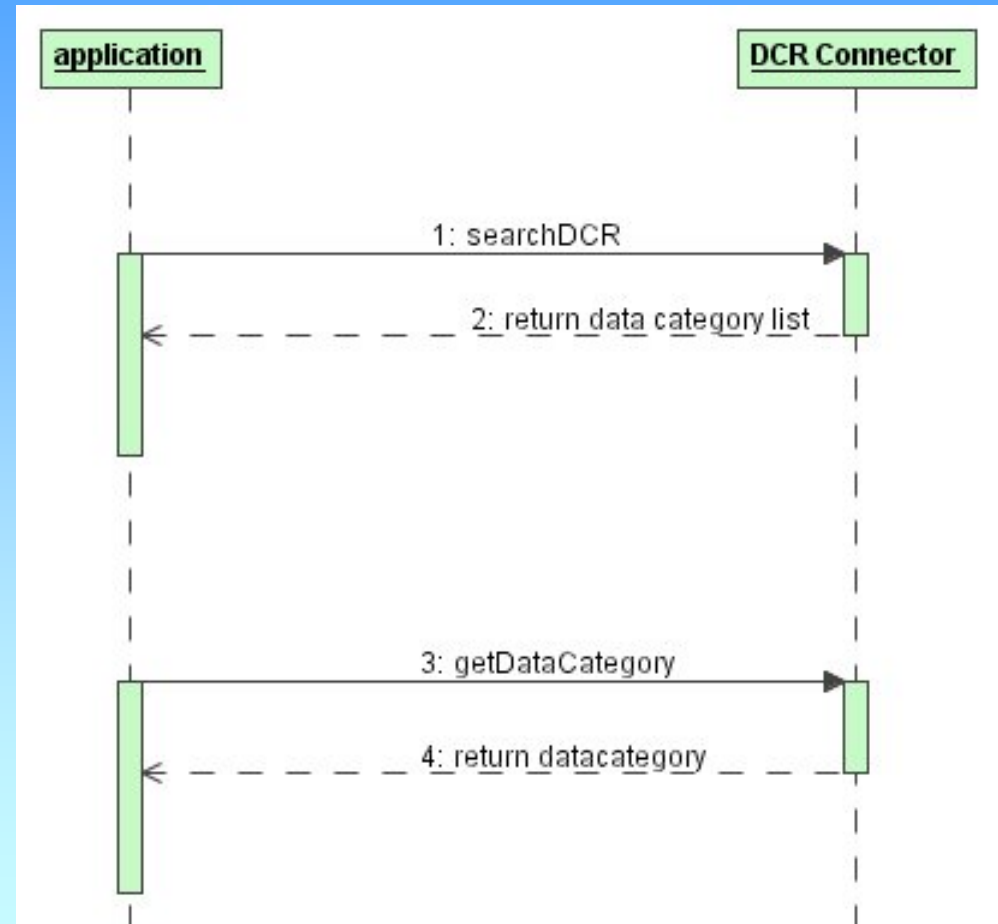
Name section		
Name	Status	Note
homophone	standardizedName	



Searching the DCR

Lirics

- User searches the DCR specifying the keywords to search for in selected fields (Identifier, definition, explanation, example, note)
- User selects datacategory from the result set to view results





Searching the DCR

Lirics

- function List searchDataCategories(List a_listOfKeywords, List a_listOfFields, String a_profile, String a_registrationStatus)
 - Returns a list of datacategories matching the specified query parameters
- function DataCategory getDataCategory(URID)
 - Returns a datacategory identified by the specified ID



Searching the DCR

Lirics

LEXUS Workspace Lexical Resource Search DCR gaby c

Data Category Repository:

Search:

Fields:
 identifier definition explanation example note

Profile:

RegistrationStatus:

Identifier	Registration Status	Registration Authority	
partitiveConceptSystem	candidate	Private	
partOfSpeech	candidate	Private	
partNumber	standard	ISO/INRIA-LORIA	
broaderConceptPartitive	standard	ISO/INRIA-LORIA	
partitiveDefinition	standard	ISO/INRIA-LORIA	
coordinateConceptPartitive	candidate	Private	
superordinateConceptPartitiv	candidate	Private	
partitiveRelation	standard	ISO/INRIA-LORIA	
subordinateConceptPartitive	candidate	Private	



DCR web service implementation

Lirics

- DCR service is implemented as REST web service
 - E.g. http://syntax.inist.fr/mod_webservice/call.php?fct=getProfiles
- Response is provided in RelaxNG format.
 - E.g.

```
<element name="STRUCT">
  <attribute name="type">
    <value>ListofProfiles</value>
  </attribute>
  <zeroOrMore>
    <element name="feat">
      <attribute name="type">
        <value>profile</value>
      </attribute>
      <text/>
    </element>
  </zeroOrMore>
</element>
```



DCR web service wrapper implementation

Lirics

- A service wrapper is created to wrap service request details and return response as POJO objects shielding developers from messaging details.
 - DCR service available though standard Java API.
 - Message results are returned as Java objects
- This service wrapper is delivered as part of the Lirics project



LMF web service

Lirics



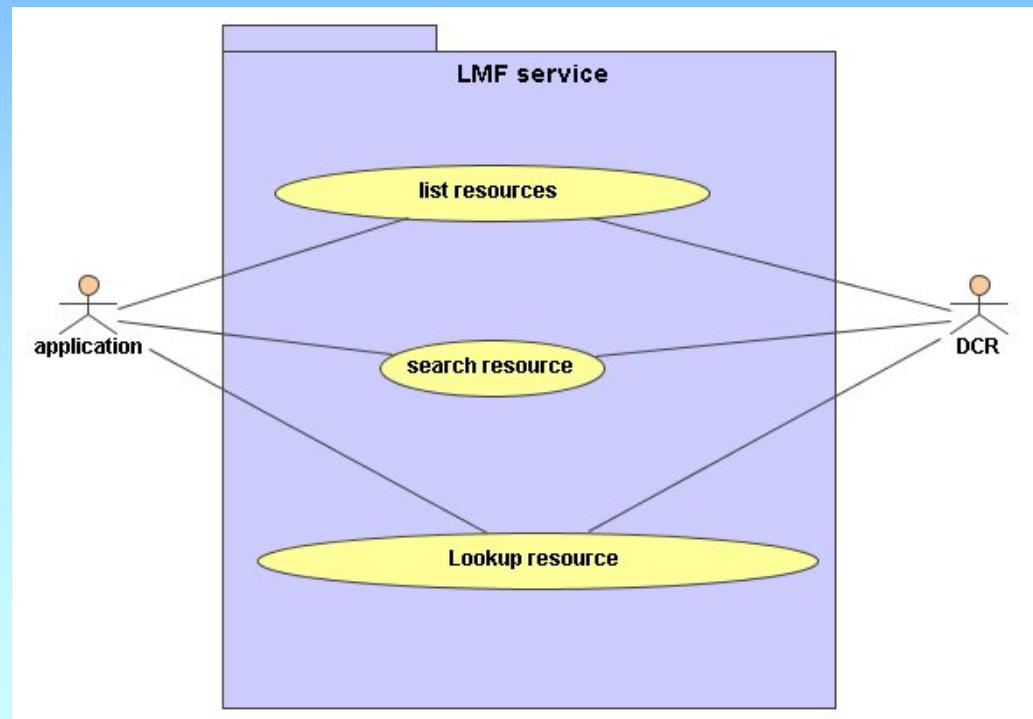
Interface to LMF service

Lirics

- Purpose is to allow access to content from lexical resources by external applications

• Use cases

- List resources
- Search resource
- Lookup resource

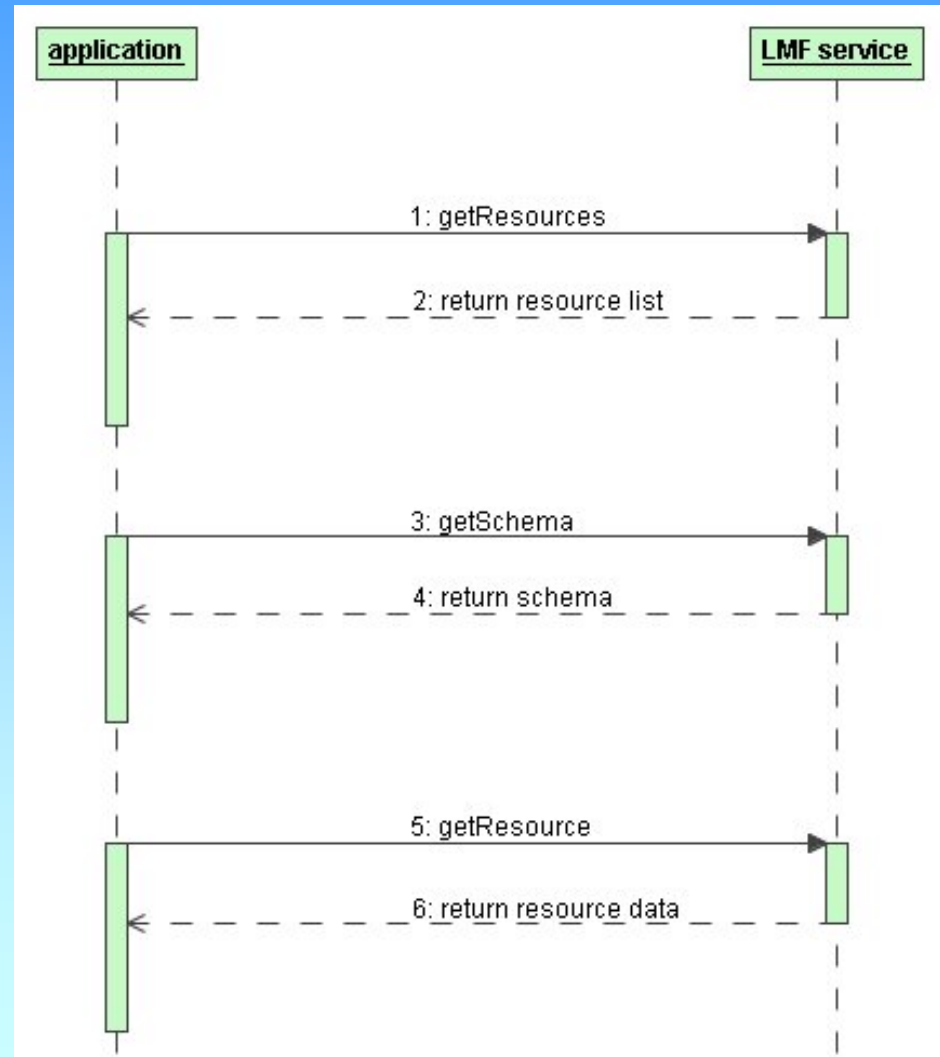




Lookup resource

Lirics

- The LMF service is contacted for a list of available resources
- A specific resource is selected.
- The schema for the resource is requested
- The data for the resource is requested

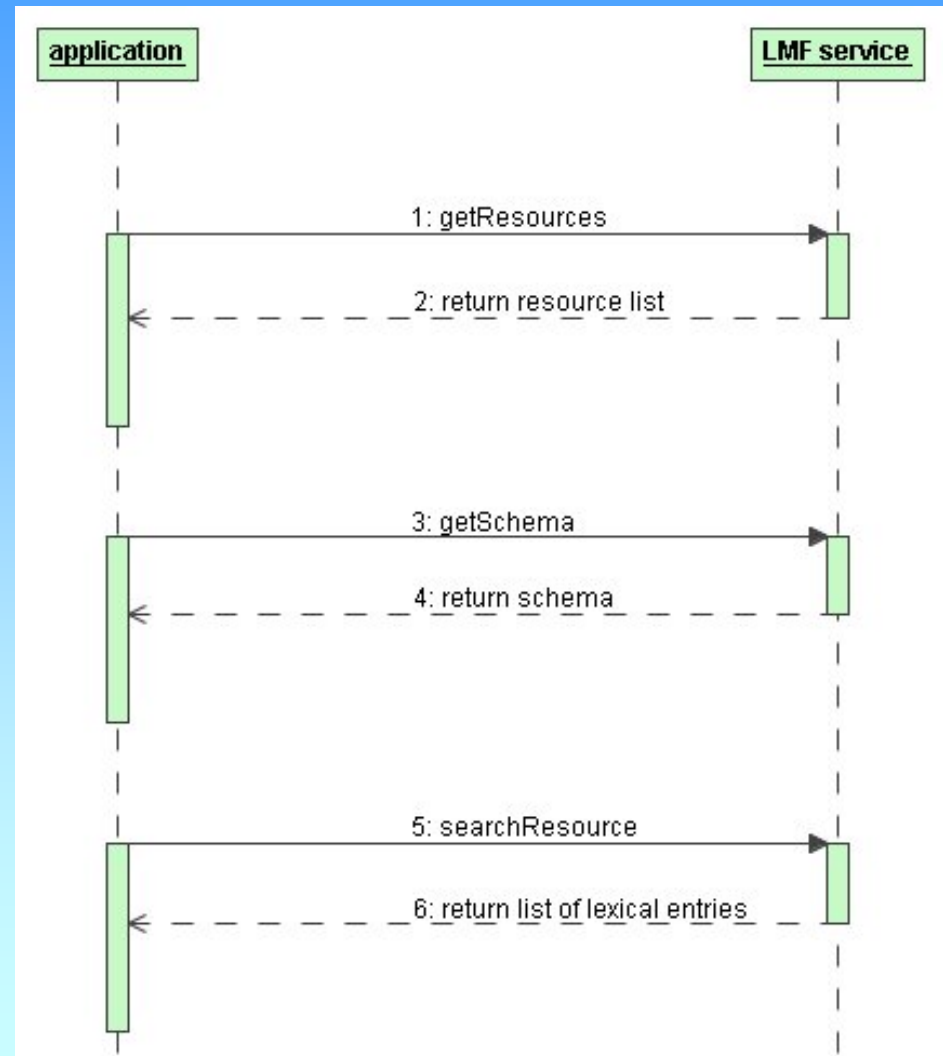




Search resource

Lirics

- The LMF service is contacted for a list of available resources
- A specific resource is selected.
- The schema for the resource is requested
- The selected resource is searched for lexical entries that match the specified criteria





LMF web service implementation

Lirics

- LMF service is implemented as SOAP web service
- WSDL available for client generation
- Lexicon documentation is generated at namespace identifier.
- Response format is well defined
 - E.g.

```
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" elementFormDefault="qualified">  
  <xs:element name="resources">  
    <xs:complexType>  
      <xs:sequence>  
        <xs:element maxOccurs="unbounded" ref="lexicon"/>  
      </xs:sequence>  
    </xs:complexType>  
  </xs:element>  
  <xs:element name="lexicon">  
    <xs:complexType>  
      .....
```
 - Schema information is returned in XML Schema (RelaxNG is also possible)
 - Resource information and search results validate against resource (sub) schema.



Gate LMF plugin

Lirics

- Visual Resource similar to the WordNet plug-in
- WSDL back-end according to the LMF standard
- Standards
 - ISO/TC 37/SC 4 LMF
 - Unicode
 - ISO 12620 DCR
 - GATE (one of the *de facto* standards in NLP)



WordNet plugin

Lyrics

Gate 2.0beta1 build 759

File Options Tools Help

Gate

- Applications
- Language Resources
 - WordNet
- Processing Resources
- Data stores

Messages WordNet

Search Word: bank Search

Searches for bank: Noun Verb Adjective Adverb

Hypernyms:

- Member Holonym
- Hyponym
- Meronym

Sense 1

=> depository financial institution, banking concern, banking company -- a financial institution that accepts deposits and channels the money into lending activities; "he cashed a check at the bank"; "that bank holds the mortgage on my home"

=> financial institution, financial organization -- an institution (public or private) that collects funds (from the public or other institutions) and invests them in financial assets

=> institution, establishment -- an organization founded and united for a specific purpose

=> organization, organisation -- a group of people who work together

=> social group -- people sharing some social relation

=> group, grouping -- any number of entities (members) considered as a unit

Hypernyms:

Sense 2

=> bank -- sloping land (especially the slope beside a body of water); "they pulled the canoe up on the bank"; "he sat on the bank of the river and watched the currents"

=> slope, incline, side -- an elevated geological formation; "he climbed the steep slope"; "the house was built on the side of the mountain"

=> geological formation, geology, formation -- the geological features of the earth

=> natural object -- an object occurring naturally; not made by man

=> object, physical object -- a physical (tangible and visible) entity; "it was full of rackets, balls and other objects"

=> entity, something -- anything having existence (living or nonliving)

Hypernyms:

WordNet 1.6 Features Editor

Closes the selected resources



Gate LMF plugin

Lirics

- GUI client will be useful on its own
- and it will demonstrate the usability of the LMF standard:
 - a programmatical client would also work with the same WS interface,
 - re-using much of the same API, and
 - possibly easier to implement!



Machine-Readable Dictionary

Lirics

- LexicalEntry
 - LemmatisedForm
 - InflectedForm
 - RepresentationFrame (alternative orthography)
- EntryRelation
 - Headword <-> RelatedForm
- Definition, Translation
- SenseRelation