

-
-
-

LEXUS a flexible web based lexicon tool

Marc Kemps-Snijders

Peter Wittenburg

www.mpi.nl/lexus

lexus@mpi.nl





Topics

- LEXUS features
- LEXUS and DCR
- Current development
- Conclusion/future development



LEXUS features – Authorization and authentication

Please enter your username and password.

Username:

Password:

[register as guest..](#)

LEXUS

lexical resource tool

Users must authenticate before login onto the application.



LEXUS features – Personal Workspace

Workspace

Lexical Resource

Search

DCR

willem wever



Frisian Dictionary



lexus_RUS.cut

Each user has his/her own personal workspace
where private lexica are stored

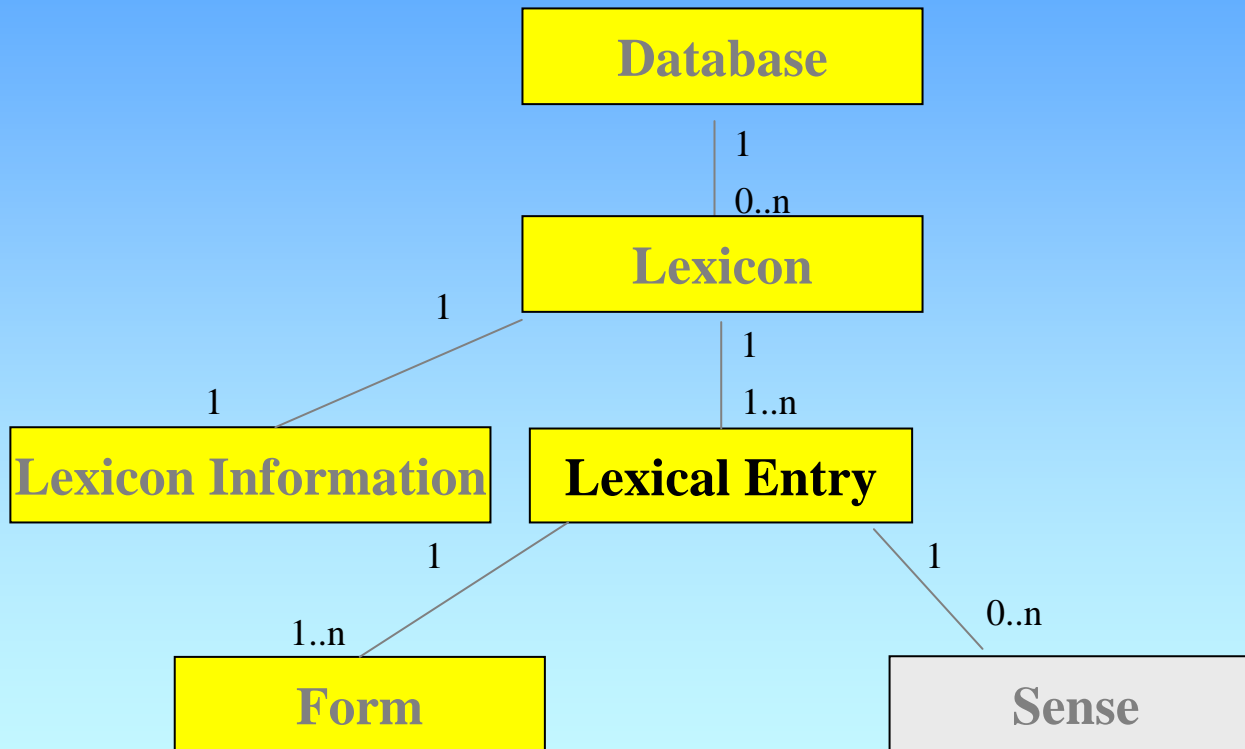


LEXUS features – LMF model

- LMF model
 - Supports LMF core model
 - Components and data categories may be added at every level
 - Data categories can be selected intuitively from a number of well established DCR's
 - Restrictions may be placed by the user on how components and data categories are to be used.

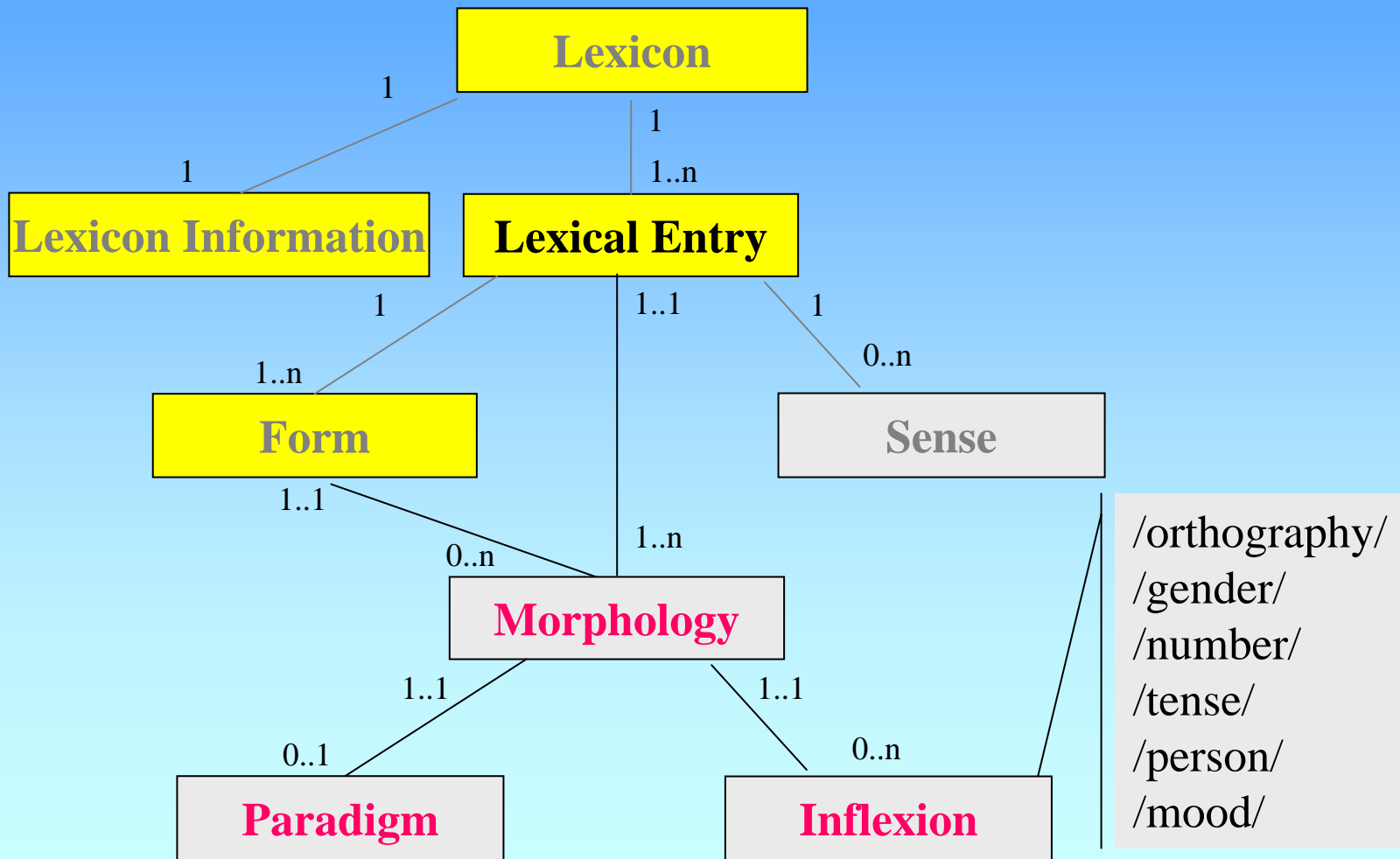


LMF Core model





Flexible extension of the Core model





LEXUS features – LMF model

Workspace Lexical Resource Edit Search View DCR willem wever

PACE) LEXUS-Lexicon Scheme Viewer

Copy
Paste
Delete
New Data Category
New Component

LEXUS
LexicalDatabase
GlobalInfo
LexicalEntry
Form
Sense
New Data Category

New Data Category

user defined
 shoebox
 iso 12620

General Layout Open issues

Creation date: Thu Apr 28 15:36:25 ICT 2005 Last modified: Author: willem wever

Description:
to be filled out

Admin info:
to be filled out

Model name: user defined

mandatory
 multiple values allowed

Reference:

Save

The LMF core model can be identified in this simple structure.
Components and datacategories can be identified using different icons.
All may be dynamically created or modified.



LEXUS features – Import

Workspace Lexical Resource Search DCR willem wever

Please select the type of lexicon you would like to import

- Shoebox
- Clan
- Wichita XML

Please select whether you want to add a new lexicon or import your data into an existing one

- Create new lexicon
- Import data into existing lexicon..

Shoebox lexicon schema and data import

This page allows you to import a shoebox lexicon. A new Lexicon

Please upload the Shoebox .typ structural file here.

Please upload the Shoebox lexicon file here.

Import

Choose file

Look in: chat

- ClanGerPolLat
- ClanGerPolLat.zip
- det.cut
- KathrinTest.cut
- lex.cut
- lex_VAN_PAS_ANJ_ALJ_end.cut
- n.cut
- test.cut
- v.cut

File name:

Files of type: All Files (*.*)

Open Cancel

New lexica may be imported from a lexical resource...



LEXUS features (continued)

- Is fully unicode enabled
- Supports use of multimedia content
- Advanced search capabilities
- Customizable layout



Interaction LEXUS and DCR

Workspace Lexical Resource Edit Search View DCR willem wever

(WORKSPACE) LEXUS-Lexicon Scheme Viewer

Shoebbox data categories are listed below. Please select a data category to be inserted into your scheme.

LEXUS

- LexicalDatabase
 - GlobalInfo
 - LexicalEntry
 - Frisian Word
 - Part of Speech
 - Frisian Word
 - Environment
 - Morpheme Property
 - Comment
 - Frisian Alternate
 - Gloss
 - English Alternate
 - test
 - English
 - Morpheme Cooccurrence Constraint
 - Underlying Form

Data category	Description	Reference
First dual	Used to give the vernacular for this particular paradigm form. (It is better to use the new \pdl field set for this.)	Shbx:MDF:1d
First plural exclusive	Used to give the vernacular for this particular paradigm form. (It is better to use the new \pdl field set for this.)	Shbx:MDF:1e
First plural inclusive	Used to give the vernacular for this particular paradigm form. (It is better to use the new \pdl field set for this.)	Shbx:MDF:1i
First plural	Used to give the vernacular for this particular paradigm form. (It is better to use the new \pdl field set for this.)	Shbx:MDF:1p
First singular	Used to give the vernacular for this particular paradigm form. (It is better to use the new \pdl field set for this.)	Shbx:MDF:1s
Second dual	Used to give the vernacular for this particular paradigm form. (It is better to use the new \pdl field set for this.)	Shbx:MDF:2d
Second plural	Used to give the vernacular for this particular paradigm form. (It is better to use the new \pdl field set for this.)	Shbx:MDF:2p
Second singular	Used to give the vernacular for this particular paradigm form. (It is better to use the new \pdl field set for this.)	Shbx:MDF:2s
Third dual	Used to give the vernacular for this particular paradigm form. (It is better to use the new \pdl field set for this.)	Shbx:MDF:3d
Third plural	Used to give the vernacular for this particular paradigm form. (It is better to use the new \pdl field set for this.)	Shbx:MDF:3p
Third singular	Used to give the vernacular for this particular paradigm form. (It is better to use the new \pdl field set for this.)	Shbx:MDF:3s
Non-animate dual	Used to give the vernacular for this particular paradigm form. (It is better to use the new \pdl field set for this.)	Shbx:MDF:4d
Non-animate plural	Used to give the vernacular for this particular paradigm form. (It is better to use the new \pdl field set for this.)	Shbx:MDF:4p
Non-animate singular	Used to give the vernacular for this particular paradigm form. (It is better to use the new \pdl field set for this.)	Shbx:MDF:4s
Antonym	Used to reference an antonym of the lexeme, but using the \lf (lexical function) field for this is better practice.	Shbx:MDF:an
Bibliography	Used to record any bibliographic information pertinent to the lexeme. MDF adds the label 'Read:' to this field.	Shbx:MDF:bb
Cross-reference	This is a generic reference marker used to link together any two related entries in the lexicon. The content is a vernacular lexeme. If the relationship is known, the lexical function \lf field is a better way	Shbx:MDF:cf

Data categories can easily be selected from data category registries. .

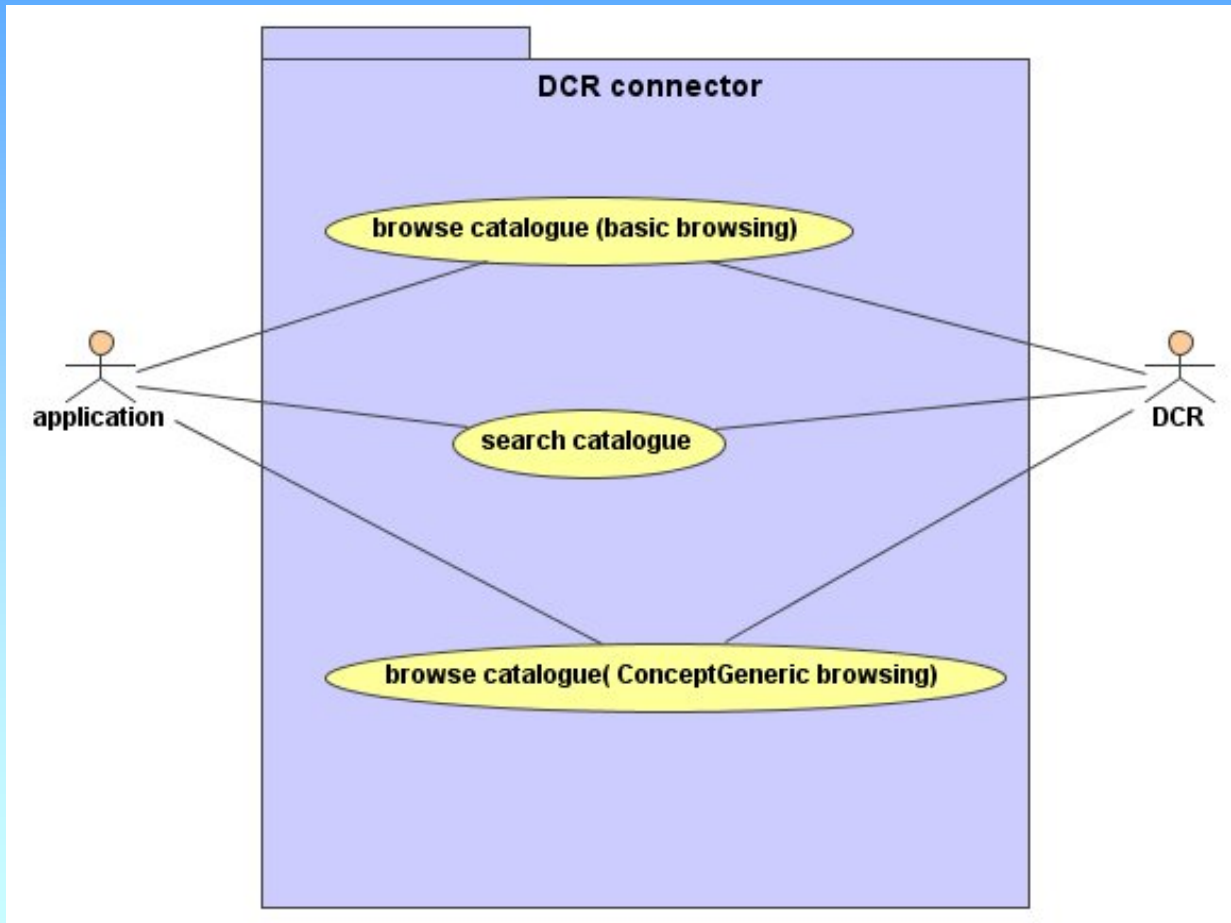


LEXUS and DCR

- DataCategories may be selected from DCR
- Currently static DCR list is used
- In cooperation with developers of Syntax an API is being produced for dynamic access (first phase)

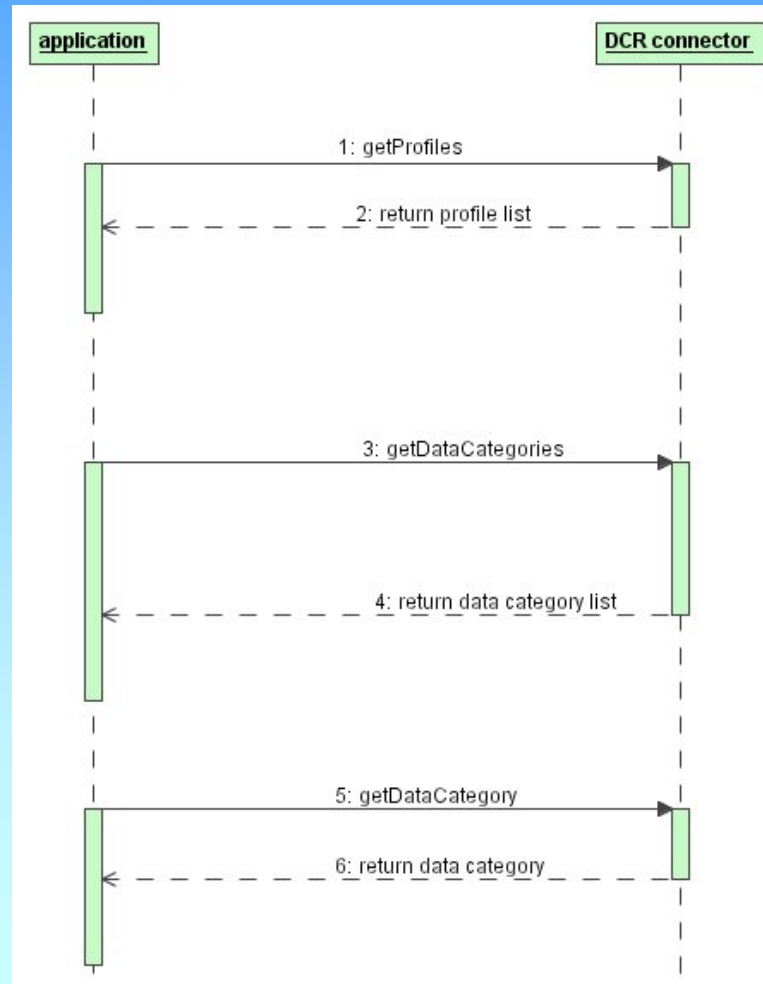


Interface to DCR services(first phase API)



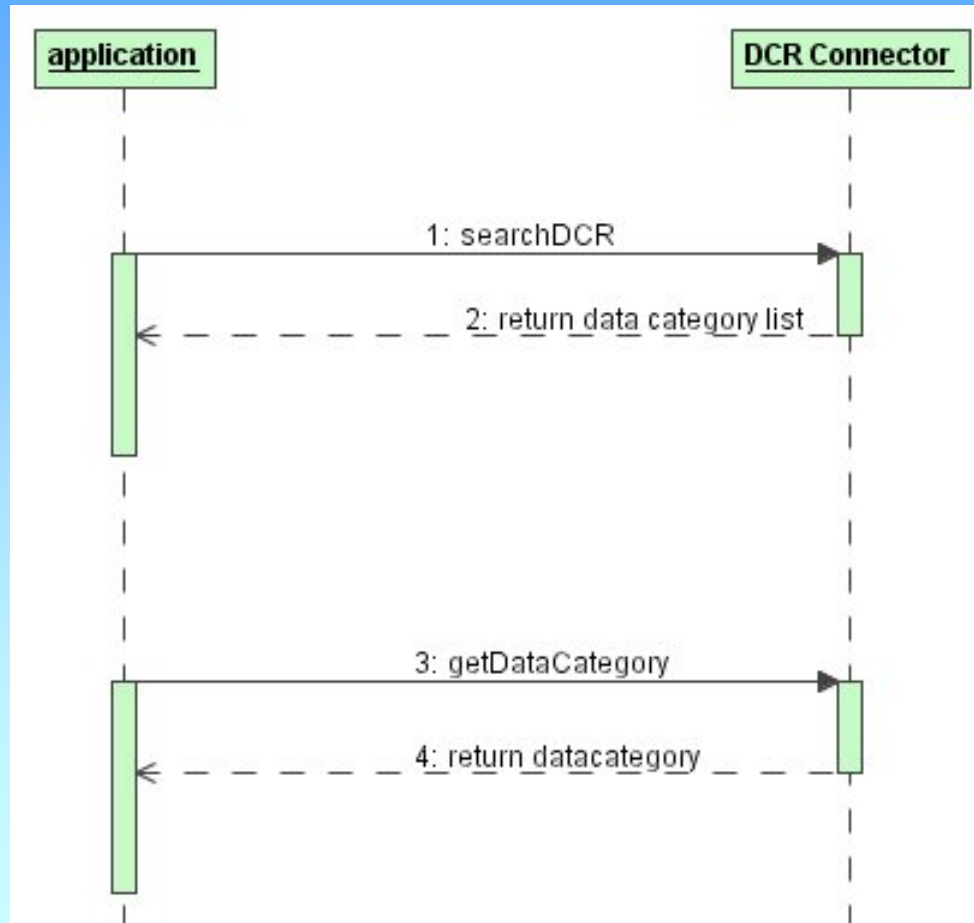


Browse Catalogue interaction (basic browsing)





Search Catalogue interaction





Interface to DCR services(methods) (*cooperation with LORIA*)

- function List getProfiles
 - Returns a list of all profiles (e.g. metamodel, terminology, morphosyntax etc.)
- function List getDataCategories(a_profile)
 - Returns a list of all datacategories associated with the specified profile
- 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
- function List searchDataCategories



Interface to DCR services(final phase)

- The DCR has to be offered finally as a web-service
 - UDDI description
 - WSDL description
 - Message exchange (SOAP)
- The UDDI registry must be searchable/browsable as one of the linguistic services
- A number of standard methods have to be supported



LEXUS worklist (current workpackage deliverables)

- Customizable print layout
- User authorization on lexical resources
- Cross browser interoperability
- Integration SYNTAX DCR API



Conclusion/ future development

- Support for component registry
 - Components (e.g. MRD extension, morphology extension) can be selected and used in lexicon.
- Support for Relations, including intuitive user interface
 - Relations are used in suggested extensions(e.g. semantics extension)
- Further development DCR API and web service
 - Work out remaining API methods in cooperation with LORIA
- Lexicon merging
- Improved user interaction