LIRICS

Deliverable 5.4.B

Data Category usage platform - documentation

Project reference number	e-Content-22236-LIRICS
Project acronym	LIRICS
Project full title	Linguistic Infrastructure for Interoperable Resource and Systems
Project contact point	Laurent Romary, INRIA-Loria 615, rue du jardin botanique BP101. 54602 Villers lès Nancy (France) romary@loria.fr
Project web site	http://lirics.loria.fr
EC project officer	Erwin Valentini
Document title	Data Category Usage Platform
Deliverable ID	**D5.4.B**
Document type	Report
Dissemination level	Public
Contractual date of delivery	M18
Actual date of delivery	30 June 2006
Status & version	Draft
Work package, task & deliverable responsible	WP5, USFD
Author(s) & affiliation(s)	Marc Kemps-Snijders (MPI)
Additional contributor(s)	
Keywords	Data Category Usage Platform

Document evolution

version	date	version	date
0.1	25/01/06		
1.0	15/12/06		
2.0	13/07/2007		

Table of contents

1 Introduction.

This document describes the how the open source tool LEXUS will be extended to allow users to find and re-use existing data categories and to offer a framework to define new ones. This deviates from the workpackage description which states that the open source tool ELAN would be extended. The reason for this deviation is that the timeline for the development of LEXUS was better suited for integration with the Data Category Registry server. The results for this project will appear in future releases of ELAN.

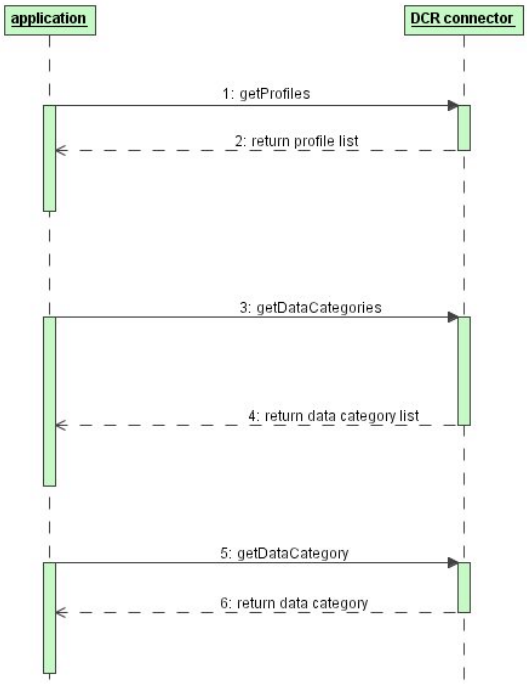
The software for connecting to the Data Category Server allowing users to search, browse and retrieve data categories of their interest is delivered as a standalone software package allowing other application developers to integrate this functionality into other applications. It shields the developer from the intricacies of communicating directly with the web based service provided by the Data Category Registry Server. Since this server implements a parameterized HTTP request based operation approach rather than a web service approach (WSDL and SOAP) transformation of the operation and parameters into the appropriate request needed to be shielded from the developers using a custom approach. In normal web service operation this is done by appropriate tools generating the necessary client code.

The software implements the full interface as defined in task 5.1 as set of Java classes describing the interface, its implementation and supporting classes.

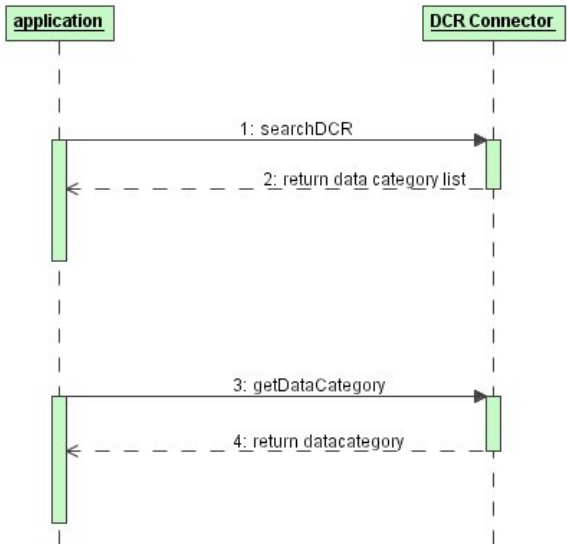
2 Use case summary.

The following use cases have been specified and implemented on the SYNTAX DCR server

- Browse catalogue(basic browsing)
    - Standard navigation over the DCR is done by selecting a profile of interest after which a list of data categories are retrieved from the DCR that are part of this profile. Optionally, a registration status may be specified to limit the number of data categories to be retrieved from a profile to only those with the specified registration status.



- Search catalogue
    - The DCR may be also searched by specifying the search terms and optional parameters such as profile to search or data category sections(title, description etc) to search. An overview of the interaction is shown below.



3 Implementing DCR access.

A user may browse the ISO 12620 DCR by selecting a profile. The data categories from the selected profile are displayed allowing users to either view the details of the data category or insert the data category specification directly into their LMF model.

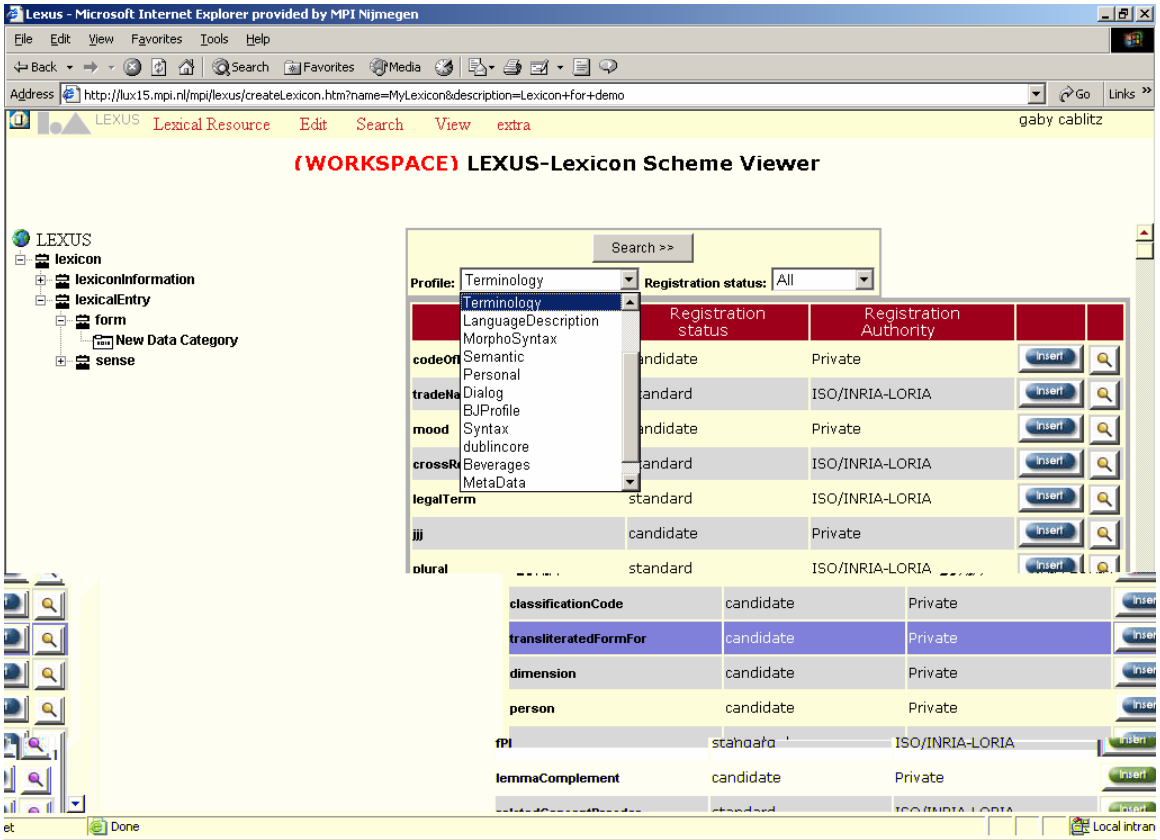


Figure 1: Profile browsing.

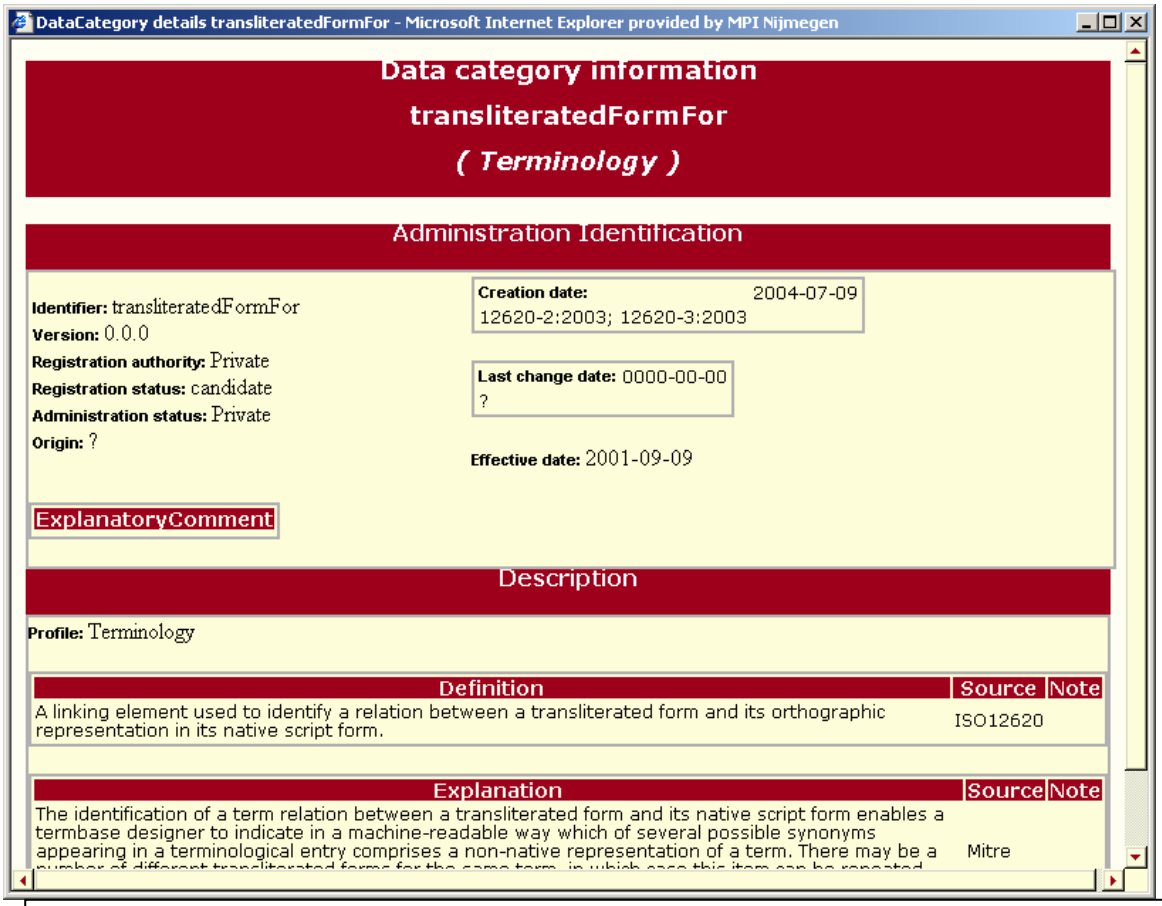


Figure 2: Datacategory details

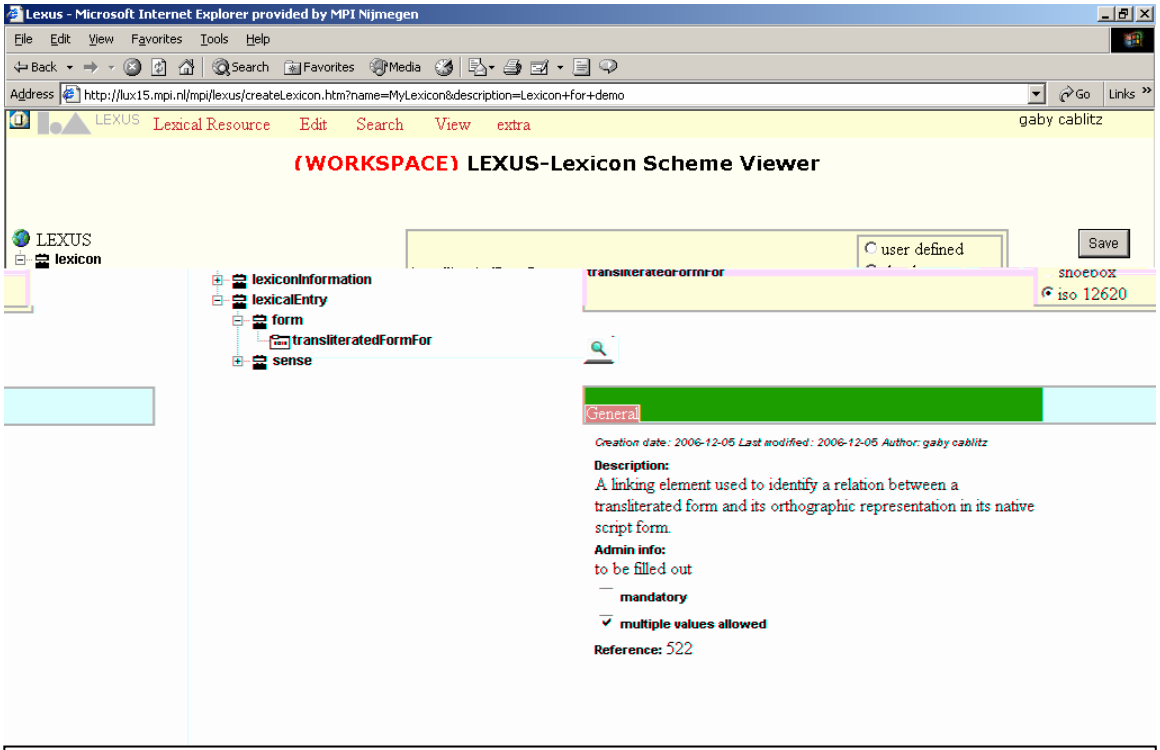


Figure 3: Inserting ISO 12620 datacategory directly into LMF model.

Another method of accessing data categories in the DCR is by searching the DCR for information matching the specified search criteria. Search across a number of fields is supported allowing the user vary the scope of the query. Details of the search result may be viewed or one of the results may be inserted directly into the LMF model.

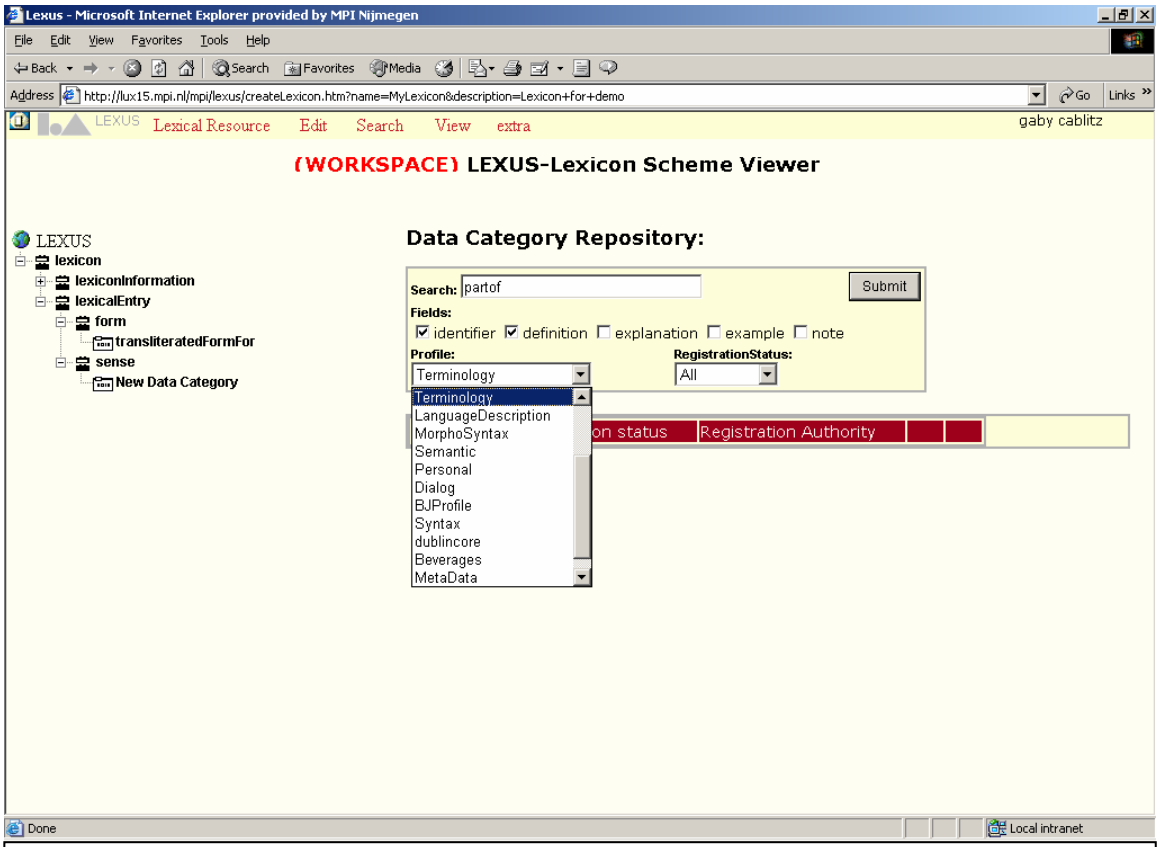


Figure 4: Specifying search criteria.

The LEXUS tool also allows users to define data categories directly. However, these may not be added to the ISO 12620 DCR directly since the ISO 12620 DCR does not support direct programmatic insertion of newly defined data categories into the DCR.

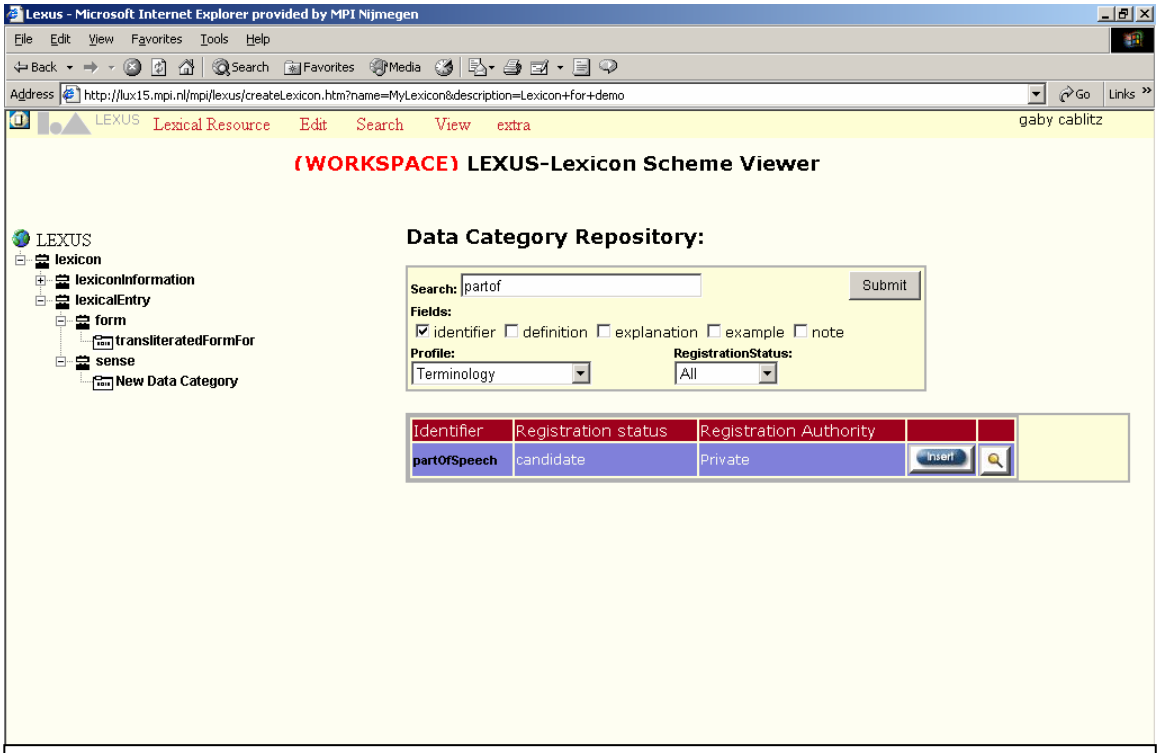


Figure 5: Displaying search results.