Project no. 018340

Project acronym: EDIT

Project title: Toward the European Distributed Institute of Taxonomy

Instrument: Network of Excellence

Thematic Priority: Sub-Priority 1.1.6.3: "Global Change and Ecosystems"

C5.79 SDD Import/Export module for CDM v1 (software)

Due date of component: Month 32
Actual submission date: Month 32

Start date of project: 01/03/2006 Duration: 5 years

Organisation name of lead contractor for this component: 2 MNHN

Revision: final

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	**X**
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

C5.79 SDD Import/Export module for CDM v1 (software)

TDWG-SDD to ensure compatibility of the CDM platform with other descriptive systems

To ensure full compatibility of the EDIT Platform with other initiatives, as well as to foment its use as a tool for data integration and access, it was decided to develop a CDM interface for descriptive data. It was decided to use a unique already-existing public standard exchange format for descriptive data: the TDWG-SDD XML schema. This choice was made in accordance with the recent efforts of the developers of several existing tools (DiversityDescriptions ["DeltaAccess"], Xper², Lucid, FRIDA) to allow interoperability with SDD.
In this context, preliminary tests were performed to validate the exchange possibilities between two existing tools (Lucid and Xper²) through SDD. The results were presented at the 2008 TDWG conference in a poster (http://www.tdwg.org/proceedings/article/view/416). These tests underlined the importance of interpretation of the standard but are nonetheless very encouraging for the taxonomic community, who will benefit from the complementary features provided by different existing tools.

Import/export SDD-CDM

To allow the CDM to communicate with other descriptive systems through the TDWG-SDD XML schema, import functionalities into the CDM java libraries were implemented. For the first version of the import, the following descriptive data can be retrieved from SDD and stored into a CDM database:

- metadata about the SDD file: how it was generated and when
- dataset information: name, author, editor, date of modification
- copyright
- taxon names
- descriptive system with the definition of characters: categorical, quantitative and text characters
- taxa descriptions using the descriptive system: associated with a taxon if the description refers to a taxon name
- publications and their association with descriptions
- images references.

The import of the following elements will be implemented in the next version:

- character dependencies and groups
- single access keys
- specimen descriptions.

We established a mapping between SDD XML elements and CDM description model Java objects and slightly modified some previously implemented CDM Java objects. This work brought forth a considerations about on the CDM description model that will fuel the revision of the model planned for February 2009.

The import of SDD data was tested on different data sets: a simple example included with the SDD schema with the description of the taxon *Viola hederacea* (http://wiki.tdwg.org/twiki/bin/view/SDD/Version1dot1), SDD examples available on the SDD wiki site ranging from 54 to more than 2,000 taxon descriptions (http://wiki.tdwg.org/twiki/bin/view/SDD/RealWorldExamples_SDD1dot1), and an SDD export from an Xper² application on freshwater aquatic insects (Figure 1: 16 taxa, 14 categorical characters, 35 states, 52 images). Both saving descriptive data into an empty base and adding descriptive data into an existing base have been tested.

Until a tool to view CDM descriptive data is available, we are using SQL queries to visualize the content of the descriptive part of the CDM database and test the consistency of the imported information.

Finally, as we enhance the import functionalities, an export function will be implemented creating SDD XML files from extracted CDM descriptive data, allowing bidirectional communication between tools using SDD and the CDM.

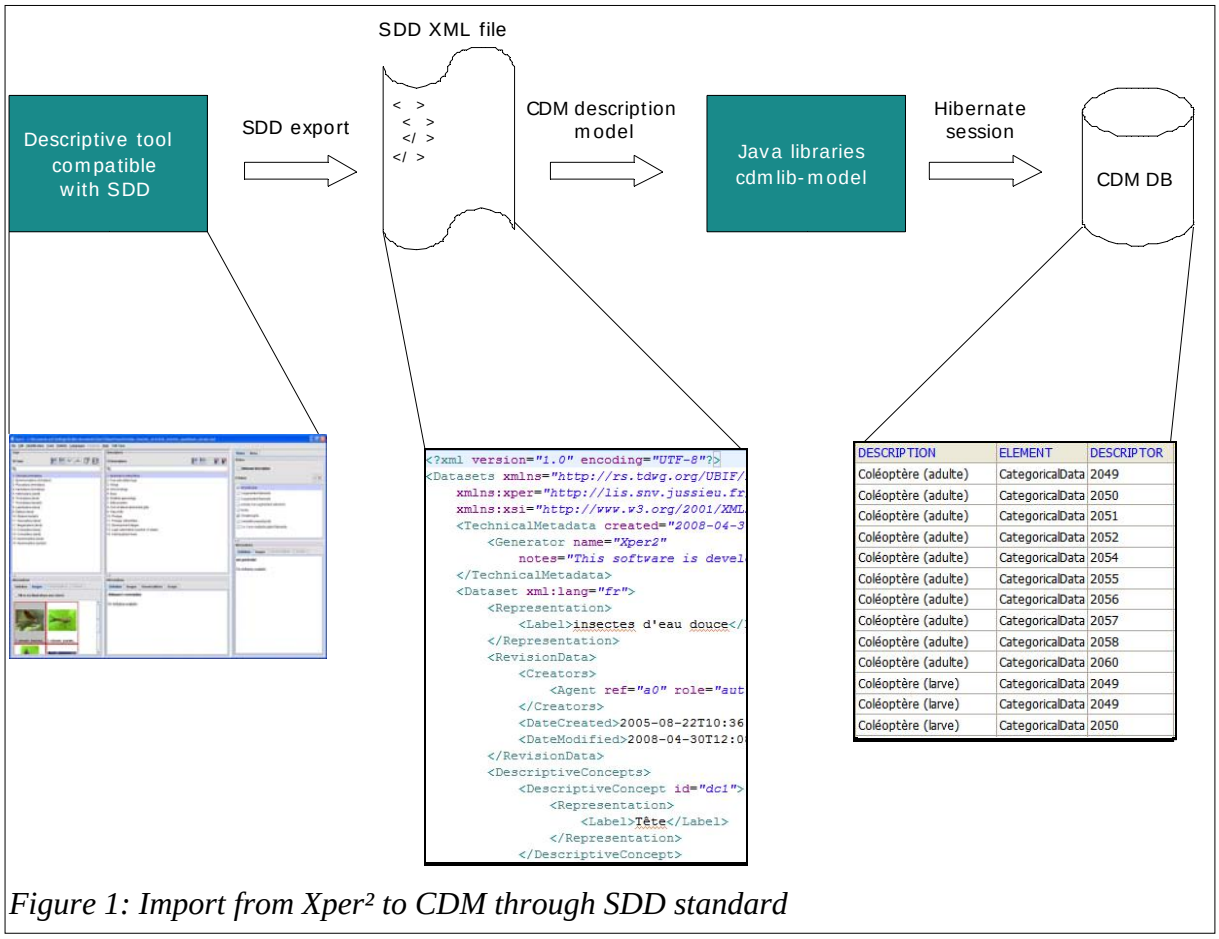

Figure 1: Import from Xper² to CDM through SDD standard