



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”

C3.2.7 CDM import tools for the pan-European species registers (European Register of Marine Species, Fauna Europaea and Euro+Med Plantbase)

Due date of component: Month 46

Actual submission date: Month 47

Start date of project: 01/03/2006

Duration: 5 years

Organisation name of lead contractor for this component: 9 BGBM

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)	

CDM import tools for the pan-European species registers (European Register of Marine Species, Fauna Europaea and Euro+Med Plantbase)

The Pan European Species directories Infrastructure (PESI, <http://www.eu-nomen.eu/pesi/>) is a 7th Framework funded initiative aiming at the integration of Europe's taxonomically authoritative species name registers and nomenclators. In its present phase, PESI provides access to the three major European taxonomic checklists Euro+Med, Fauna Europaea, and the European Register of Marine Species (ERMS). PESI will provide a common access system for all checklists via a web portal as well as web services for direct machine-to-machine communication.

Both portal and web services operate on a "PESI data warehouse" (Microsoft SQL-Server 2008) containing a merge of the participating checklists databases providing the information in a single unified taxonomic tree structurally optimized for output purposes. This data warehouse will be cyclically generated by software tools belonging to the Platform for Cybertaxonomy, which is being implemented in the context of the European Distributed Institute of Taxonomy EDIT (<http://wp5.e-taxonomy.eu/blog/>).

This component covers the import of ERMS, Fauna Europaea, and Euro+Med into EDIT CDM stores. We have implemented three separate applications, each performing a data import for one of the checklists. In a first phase, all imports are going into a single CDM store holding the three taxonomic trees within disjoint structures. In a next conversion phase, the merging is carried out leading to a single unified taxonomic tree. Taxonomic overlaps and conflicts are stored in a report to be resolved in the respective checklist source databases.

In a next step, we will provide a fully functional software module for exporting the merged checklist CDM store into the PESI data warehouse structure. Together with the import functionality the platform will then cover the entire information flow from the source checklists to their unified World Wide Web publication.

Both the import applications as well as the entire CDM software suite are freely available from the EDIT WP5 software repository (<http://dev.e-taxonomy.eu/trac/>).