

Master Data Management and User Services of the Catalogue of Life

Viktoras Didžiulis, Yuri Roskov & Thomas Kunze

Catalogue of Life – i4Life, School of Biological Sciences, Harborne Building, The University of Reading, Reading RG6 6AS, UK

Email: v.didziulis@reading.ac.uk, y.roskov@reading.ac.uk, t.kunze@reading.ac.uk

The Catalogue of Life contains over 2 million names (accepted names, synonyms and common names) with higher classification of marine, terrestrial and freshwater species from a growing number of Global Species Databases. The Catalogue of Life is released monthly. Keeping database of this size up to date is a complex task because of big volume of data entities and multiple relationships that need to be managed. Making the resource available to users is a different kind of challenge because of wide range of user needs and preferences. We aimed to improve both master data management infrastructure, data aggregation (assembly) workflow and user services of the Catalogue of Life during the EU funded projects 4D4Life (2009-2012) and i4Life (EC-FP7-i4Life-project-261555). A notable improvement was creation of the Catalogue of Life Workbench – a master data management system that allows combining both manual and automated workflows in an assembly process. This development currently allows us to automate updates of marine sectors from 33 databases provided by WoRMS and the number is growing. We are also establishing more robust, standardized data sharing mechanisms between the Catalogue of Life and Global Biodiversity Programmes - GBIF, Encyclopedia of Life, IUCN Redlist, Barcoding for Life and the European Nucleotide Archive. The two projects delivered services like download service, species list matching service and webservices that users will benefit from.