

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”

C6.18 A report on the sustainability of exemplar groups and their data.

Due date of component: Month 50

Actual submission date: Month 50

Start date of project: 01/03/2006

Duration: 5 years

Organisation name of lead contractor for this component: 10 - NHML

Revision [final]

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level (“X” in the relevant box)		
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)	

Cichorieae exemplar group

During the course of the EDIT project, the Cichorieae exemplar group has launched the ICN (International Cichorieae Network), a new world-wide initiative which brings together the specialists working on this taxonomic group. The main publicly-visible product of the ICN is the Cichorieae Data Portal (<http://wp6-cichorieae.e-taxonomy.eu/portal/>), aiming at a complete online taxonomic revision of this group of plants.

Current state

The Cichorieae Data Portal holds data on all of the ca. 1400 currently accepted species belonging to 93 genera, of the tribe Cichorieae of the sunflower family, the Compositae or Asteraceae. It, moreover, includes most of the some thousand microspecies of the genera *Hieracium*, *Pilosella* and *Taraxacum*. The number of all names incorporated is about 30.000. For all accepted species a taxon page with nomenclature, complete synonymy and source-referenced distribution maps is provided. Currently ca. 4100 protologues (original descriptions) are linked to the nomenclatural references. The descriptive part of the taxon pages currently provides ca. 1000 photographs, hundreds of descriptions (among them the first ca. 70 data-matrix-based, structured, standardised descriptions). Additional data types will be added until EDIT ends, such as common names, palynological data, links to specimen images, and prototypes of interactive keys for selected taxonomic groups. It has, however, always been clear that the ultimate goal of a complete online revision and a dynamic web monograph of the tribe Cichorieae is beyond the scope of EDIT. Therefore, sustainability of the International Cichorieae Data Portal is crucial to maintain the eTaxonomy approach for this group.

Long-term sustainability

The International Cichorieae Network will continue to exist beyond the end of EDIT. It holds an important and advanced role within TICA (The International Compositae Alliance), a larger network of Compositae researchers. TICA fosters the Global Compositae Checklist, plans a Virtual Key to the Compositae, and coordinates the contributions from the Compositae Community to the Encyclopedia of Life. However, the ICN can only continue to play an important role within the Compositae community if the sustainability of its core product, the Cichorieae Data Portal, is guaranteed.

The Cichorieae Data Portal is presently hosted at the web server of the Botanic Garden and Botanical Museum Berlin-Dahlem (BGBM) of the Free University of Berlin. This institution has guaranteed to continue hosting the website for the foreseeable future. Since the underlying Common Data Model of the EDIT WP5 platform for cybertaxonomy has been developed at the same institution, a certain amount of further support for the infrastructural components of the Portal can realistically be assumed after the end of EDIT. The development of new functions that might be needed in the future will require further funding.

The content in the Cichorieae Data Portal will be quite advanced by February 2011. Since active systematic research on the Cichorieae is ongoing, the data currently displayed in the ICN portal are subject to change. The corresponding permanent small-scale updating of the existing taxonomic and additional data in the portal will be done by the Cichorieae working group of the BGBM. Filling of existing data gaps, incorporating of new data types or of larger amounts of data, which may be provided in future by the Network, and, in particular, a further development of the portal's content towards a dynamic web monograph of the group, will, however, be dependent on further funding.

Palm exemplar group

The palm exemplar group has focused on creating a web portal for palm taxonomy by bringing together the extensive information on the subject that already exists either in print or in the web. This portal is the main product of the palm exemplar group and this report outlines its current state and the requirements for its long-term sustainability.

Current state

The content in Palmweb is structured around a list of palm names, from genera to infraspecific ranks, which can be browsed or searched to access the page for a particular taxon. All accepted names in the palm family have a basic taxon page that includes at least nomenclature, synonymy, a distribution map of TDWG areas and links to other databases. When available, the taxon page also displays descriptive content and images. Currently, Palmweb includes descriptive content for all genera and 40% accepted taxa, protologues for over 60% of the species and images for around 30% of them.

Palmweb is part of the platform for cybertaxonomy being developed by EDIT's WP5. Its current functionality is acceptable and more than basic. The reception of the latest version by the palm scientific community has been

positive and encouraging, and their comments indicate that the portal is useful as it is. Improving existing features (e.g. the display of references) and creating new ones (e.g. dot maps, content downloads) are necessary developments to bring Palmweb closer to the initial vision of the palm exemplar group.

Long-term sustainability

The long-term survival of Palmweb depends on the availability of resources to maintain its content and to support and further develop the infrastructure it is based on. These two aspects are discussed below.

Content

The content currently available in Palmweb has been almost entirely compiled or generated by the palm exemplar group, which consists of the team leader and the eTaxonomy coordinator, a full time post funded for 90% of the duration of EDIT. External contributions so far have been sporadic.

Once EDIT ends, the palm exemplar group won't exist as such and the long-term maintenance of the content in Palmweb will hinge heavily on the palm scientific community and institutional support. Well before EDIT, the European palm scientists formed a network (EUNOPS), which also involves researchers from the rest of the world and will continue to exist independently of EDIT. The network has been supportive of the portal from the start of the project but their participation has been limited by the lack of an adequate content management tool. With the development of the Taxonomic Editor (the tool to edit the portal), it is expected that the involvement of the palm scientists will gradually increase. In the coming months and until the end of EDIT, the palm exemplar group will take the necessary steps to bring Palmweb to the palm community.

However, even if the palm experts take charge of the maintenance of Palmweb, an obstacle to their consistent and long-term participation, and hence on Palmweb sustainability, is the lack of mechanisms to recognise and reward the time invested by the scientists. After EDIT, funding would be necessary for an eTaxonomist to provide support to taxonomists at least until such mechanisms are in place and eTaxonomy becomes a mainstream approach to deliver the products of taxonomic research. The post would also speed up the completion of the palm taxonomy knowledge base, which will be well advanced but not finished by the end of EDIT in February 2011.

Infrastructure

It's crucial for an online resource to be user-friendly and for the tools to manage it to be well supported and easy to share among members of a community. As mentioned before, the portal Palmweb is useful as it is, but it does not deliver yet some of the content essential for taxonomic and biodiversity research (e.g. comprehensive references, specimen maps). The Taxonomic Editor is visually appealing but has a rather complicated set up and user management (required for a community-maintained website) is not operational.

Even though it is informative, the current portal falls short of the original vision for the website and needs more work to become a model tool for eTaxonomy. If improvements are not made to the portal and the editor, the interest of the general and expert users will quickly fade. Therefore, the sustainability of Palmweb requires the investment in further developments of the portal as well as its editing tool.

Currently Palmweb is hosted at the BGBM Berlin, where most of the tools for the platform are being developed. The BGBM is committed to keep Palmweb on their servers in the short term. Long-term sustainability then requires moving the portal to a permanent location, ideally RBG Kew, the institution where the palm exemplar group is based and that has supported the content generation for the portal. The RBG Kew has a strong IT department with the capacity to support and further develop the software for Palmweb and to provide the storage space needed for the images and pdf files delivered by the portal. Kew is also strongly involved in CATE, another eTaxonomy initiative which is, to a certain extent, compatible with EDIT's cyberplatform. A distributed software poses multiple challenges that hinder the transfer of Palmweb to another institution. Because of this, no commitments have been made by the RBG Kew, but Palmweb and eTaxonomy are clearly on its the agenda. Moreover, Kew is committed to the recently funded eMonocot project, an initiative that will make eTaxonomy a more prominent activity in the institution

Ultimately, the long-term sustainability of Palmweb, and any other eTaxonomy resource, depends on eTaxonomy moving from being project-funded to becoming part of the core activities of taxonomic institutions.

Summary of requirements for the sustainability of Palmweb

- Resources for an eTaxonomist during the gradual transfer of content maintenance to the palm scientific community and at least until recognition for doing eTaxonomy is well established.
- Transfer of Palmweb content and software to a permanent location.

- Resources to support and further develop the portal software and the editing tool with the flexibility to innovate and go beyond minimum requirements.
- eTaxonomy becoming an accepted approach to publish taxonomic research and part of the core activities of taxonomic institutions.

Insect exemplar group

In contrast to the other two exemplar groups, the insect exemplar group has focussed on using the Scratchpads system to enable expert networks as well as single researchers to publish their data on the web. In addition to this, data on the Diptera family Milichiidae are stored in the Common Data Model (CDM) and shown on the Diptera portal in order to test the software developed by WP5 for its suitability for zoological data.

Current State

As of May 2010 there were more than 30 Scratchpads focussing on insects. Insect Scratchpads offer diverse data, e.g. pages about entomological societies, species pages, bibliographies, taxonomic classifications, distribution maps, and images. Several Scratchpads have developed into community websites for their taxa. The experts or expert groups maintaining these sites are placed around the world but there is a large number of maintainers at the Natural History Museum, London (NHM). The focal Scratchpads for the insect exemplar group are the Milichiidae (www.milichiidae.info), Diptera (www.diptera.myspecies.info) and fungus gnat Scratchpads (www.sciaroidea.info) which together contain about 24,500 pages.

The Diptera data stored in the CDM mainly consist of a comprehensive catalog of the family Milichiidae with nomenclatural and taxonomic data as well as a complete bibliography with all citations for about 280 species. These data are the product of one researcher, not a network.

Long-term sustainability

The responsibility for the data in a Scratchpad lie with the maintainer of the site, not with EDIT. The NHM encourages its scientists to use the Scratchpad system and has committed itself to continue hosting the websites in the foreseeable future. Furthermore, the Scratchpad system itself will be further developed as part of the ViBRANT EU grant, which will start in late 2010 and run for three years.

The main amount of data generated within EDIT is present on the focal Diptera Scratchpads. The respective maintainers of these Scratchpads have agreed to continue caring for their websites after the EDIT project ends.

To avoid a duplication of effort by having two websites for the same group, the Diptera data in the CDM will be moved over to the Milichiidae Scratchpad (www.milichiidae.info) and the portal will be closed, eventually.