

A Pan-European Species-directories Infrastructure (PESI)

Charles Hussey¹ and Yde de Jong²

¹ Natural History Museum, Cromwell Road, London SW7 5BD, U.K.

² Zoological Museum Amsterdam, Faculty of Science - University of Amsterdam, P.O. Box 94766, NL - 1090 GT, Amsterdam, The Netherlands

¹ c.hussey@nhm.ac.uk; ² yjong@science.uva.nl

Abstract. This communication introduces the rationale and aims of a new Europe-wide biodiversity informatics project. PESI defines and coordinates strategies to enhance the quality and reliability of European biodiversity information by integrating the infrastructural components of four major community networks on taxonomic indexing, namely those of marine life, terrestrial plants, fungi and animals, into a joint work programme. This will include functional knowledge networks of both taxonomic experts and regional focal points, which will collaborate on the establishment of standardised and authoritative taxonomic (meta-) data. In addition PESI will coordinate the integration and synchronisation of the European taxonomic information systems into a joint e-infrastructure and the creation of a common user-interface disseminating the pan-European checklists and associated user-services results.

Keywords. biodiversity, infrastructure, taxonomy, nomenclature, standards, Europe, checklists

1 Rationale for the project

PESI (<http://www.eu-nomen.eu/pesi>) is a new initiative, funded by the European Union under the Framework 7 Capacities Work Programme: Research Infrastructures. It will be a three-year project, starting in May 2008. Led by the University of Amsterdam, it will involve 40 partner organisations from 26 countries.

PESI will take forward and build upon previous European initiatives and, in doing so it will help sustain the momentum, ensuring that data is kept up to date and that valuable social networks are maintained and strengthened. This re-invigoration will help build the capacity necessary to support the growing international initiatives such as the *Global Biodiversity Information Facility (GBIF) Lifewatch*, the *Catalogue of Life (CoL)* and the *Encyclopedia of Life (EoL)*. At the same time, it will provide a robust infrastructure to support the nomenclatural needs of European users and stakeholders. PESI will specifically address the issues of: pooling resources, standardisation, sustainability, accessibility and international cooperation. The following sections each represent a Work Package within PESI.

2 Coordination and integration of European expert networks

The strengthening and integration of European taxonomic communities has been progressing since the start of the taxonomic indexing EU framework programmes *Fauna Europaea*, *European Register of Marine Species (ERMS)*, and *Euro+Med PlantBase*. These initiatives built up expert networks to fulfill the project objectives and played an important role in helping to identify and consolidate the European taxonomic expertise. In addition *Fauna Europaea* and ERMS organised their expert resources in the *Society for the Management of European Biodiversity Data (SMEBD)*, which was established to ensure the long-term ownership and management of the intellectual property rights (IPR) of the contributors of these databases. PESI will review and develop long-term strategies and plans for the sustainability of the involved systems in terms of ownership, IPR, and support of individual experts and institutions.

A significant addition to the species registers from PESI will be the expansion of the network of expertise in Eastern Europe. This is overdue considering the expansion of the EU. In addition to providing more comprehensive data for the pan-European species registers, this expansion will reverse the separation of knowledge and practice in taxonomy for many decades. Another addition will be the closer involvement of taxonomic societies, especially on adopting relevant taxonomic sectors for long-term maintenance and upgrading.

3 Coordination and integration of European Focal Point networks

In addition to creating a network of taxonomic experts, PESI will also develop a network of regional (often national) focal points. These focal points (either individuals or organisations) will complement the taxonomic network through: (1) liaison with national governmental bodies on the implementation of European standards relevant to, for instance, national and European regulations and environmental monitoring, (2) collection and transfer of local expertise and applied tools, (3) lobbying and public policy input at the national and European level, and (4) support for closer collaboration of scientific contributor and user communities across Europe. Focal points contribute country-specific information about species, relevant databases, local literature, experts, professional societies and major users such as government organisations.

Building on existing focal point networks, PESI investigates models to establish a more formal, integrated structure of pan-European Regional Focal Points to enable joint activities. One of these activities will be the enrichment of the European taxonomic information systems with vernacular (common) names. The absence of vernacular names indices of all European languages is currently hindering the general application of biodiversity information by non-expert users.

4 Coordination of taxonomic metadata standard assessment

In the field of biology taxon names provide the anchor which allows information about organisms to be linked. A taxonomic name, typically a species name, is attached to every primary data object (field observation, specimen, genetic data, etc. Therefore names, together with their organisation into taxonomic classifications are understood as crucial metadata for biological information systems.

In actual use there are many challenges to integrating data sources that contain taxonomic names and classifications, particularly where the sources span different organismal kingdoms or national boundaries. Names may be badly formed or incomplete and so fail against searches based upon character matching. Names that are synonyms or no longer in current use may occur in museum and herbarium specimen catalogues or in legislative lists. There may also be disagreement amongst experts on the identity of specimens and on the taxonomic constituents of genera and the arrangement of classifications. The partners involved in PESI have long experience of such problems and it is the purpose of PESI to produce practical solutions to many of these issues.

The availability of authoritative taxonomic metadata standards will be of particular importance where organisms are directly important to society or because they are subject to conservation and environmental control. PESI will promote harmonisation and certification of taxonomic metadata standards of prioritised taxa that are listed in various EU regulations and legislative lists.

PESI will work closely with relevant standards organisations to identify appropriate authoritative standards and schemas and then ensure their adoption within the European biodiversity community. Work will include development of a management classification scheme, utilisation of globally unique identifiers for names (GUIDs) and support for nomenclators (such as *ZooBank*, *International Plants Names Index*, *Index Fungorum* and *Algaebase*) to help implement a practical name resolution service.

5 Coordination and integration of information e-infrastructures

PESI technically integrates the pan-European species registers into an interoperable *e*-infrastructure by harmonising the respective information infrastructures and by creating a joint access (middle) layer. PESI will establish mirror sites for ERMS, Fauna Europaea, and Euro+Med Plant Base, and thus elevate their services and content to a new level. PESI will link the planned joint European taxonomic *e*-infrastructures middle-layer to the global *e*-gateway.

6 Integrated e-Services for users and dissemination

PESI will build an interactive, multilingual web portal to carry out the dissemination of the developed species names service and to support the use of the pan-European species data in the *e*-science domain. This will include relevant supplementary data, like occurrence details by applying dynamic links to pertinent *e*-data services.