Content aggregation and information re-use using Fedora

Ana Macario
Alfred-Wegener-Insitute for Polar and Marine Research
Bremerhaven, Germany
amacario@awi-bremerhaven.de

Abstract submission for the DORSDL workshop at ECDL 2006

Digital objects in the field of earth and biological sciences are known to be often compound and complex. If one takes biodiversity as example, an exhaustively long list of information systems can be found on-line. These systems often contain valuable and historically relevant information gathered slowly over many decades – metadata/data models and transport protocols are also consequently distinct.

In order to assure long term preservation of this distributed and not yet networked mass of information on world's biota, we need to create an abstraction layer in which content is aggregated in order to allow for information transformation and re-use . Using Fedora's flexible digital object model, we will propose a strategy for creating an information network on plankton biodiversity. This network will be built upon existing taxonomic naming and classification systems (uBIO Taxonomic Naming and classification services), digitised literature on biodiversity (Biodiversity Heritage library Project or any other open access repositories), on-line taxon descriptions (wikipedia-styled), and, mostly important, environmental data archived in world data centres. Because XML schemas for metadata description(s) and for expressing relationships among the objects will be made available, interoperability with other federated networks will be assured.

This work is being currently funded by the EU-proposal "An open-access system for developing and supporting distributed knowledge centres for taxonomic data – a pilot study targeting EU phytoplankton" under the acronym "Plankton-Net".