Distributed Open-Source Development in the DINA Consortium

Fredrik Ronquist & Karin Karlsson
Dept. Bioinformatics and Genetics
Swedish Museum of Natural History
Collection Management Systems

1. Develop your own system in-house
2. Acquire a commercial system
3. Partner with other institutions in distributed open-source development
The Case For Open Source

- **Market considerations.** Professional collection management systems not viable commercial products in a pluralistic market.
- **Long-term stability.** An open-source software solution developed by institutions with long-term focus will be more stable than a commercial solution.
- **Flexibility.** A distributed open-source system must by necessity conform to a modular design based on open API:s. This favors flexibility and adaptability in a way that a commercial product will not.
- **Cost effectiveness.** Although some overhead is associated with distributed development, more development teams involved in the effort will result in a lower cost to the individual institution compared to in-house or commercial solutions.
The Case For Open Source (cont’d)

- **Opt-in opt-out scheme.** Institutions can participate in the development when they have resources to do so, and can opt out when they do not. At any single point in time, it should be feasible to have enough institutions involved for development to move forward at an acceptable pace.

- **Community Control.** A distributed open-source solution means that the community retains control over both the information standards and the system architecture and web service/API designs.

- **Egalitarian.** A professional open-source collection management system offers a better way for developing countries to catch up than any commercial product.

- **Stable marketplace for extensions and services.** A community-supported de-facto standard for collection management systems architecture will ensure that there is a stable market for various plugins, extensions and services based on the system.
DINA Consortium
(Digital Information system for NAtural history data)

- **Core Member.** Required contribution 1.0 FTE to the project, of which at least 0.5 to the development effort. Voting member of the System Engineering Task Force, which controls deliverables and deadlines for the 1.0 FTE contribution.

- **Associate Member.** No contribution requirements. Non-voting member of the Steering Group.
DINA Consortium

- Core Members
  - Agriculture and Agri-Food Canada, Ottawa
  - Estonia (University of Tartu)
  - Denmark (University of Copenhagen)
  - Sweden (Swedish Museum of Natural History)

- Associate Members
  - Museum für Naturkunde, Berlin
  - Royal Botanic Garden, Edinburgh

- Open to Additional Members
  - Memorandum of Cooperation and more information at [http://dina-project.net](http://dina-project.net)
Welcome to PlutoF

PlutoF provides cloud database and computing services for the taxonomical, ecological, phylogenetical, etc. research. The purpose of the platform is to provide synergy through common modules for the classifications, taxon names, analytical tools, etc.

It allows to address integrated questions in ecology and coevolution of taxa. Different types of the species occurrences, viz. preserved specimens, DNA sequences, human observations, references can be stored in PlutoF as well. PlutoF has no restrictions on taxon and geographic coverage and therefore can be used for the databasing interacting taxa.

It also includes collection management module. Few examples of the public web outputs from PlutoF are Estonian eBiodiversity (http://elurikkus.ut.ee), and molecular key for fungi (http://unite.ut.ee).
Lessons learned

- **Commitment.** Formalization of the collaboration and a good governance model is essential.

- **Patience.** It may take an institution with long-term perspective several years from a decision to join the consortium to actively contributing to the development.

- **Respect.** Different teams come with different backgrounds, different skill sets, and different external pressures. Striking the right balance between the cathedral (centrally controlled) and the bazaar (locally controlled) approach to collaborative development is crucial.

- **Trust.** A team needs to trust the other teams in the consortium to deliver according to agreements, so that consortium membership pays off.