Accelerating ecological data management via software sharing:
The Information Management Code Registry

Colin Smith\(^1\), Kristin Vanderbilt\(^2\), Corinna Gries\(^1\)
University of Wisconsin-Madison\(^1\), University of New Mexico-Albuquerque\(^2\)

Overview

Ecologists and data managers write significant amounts of code for data manipulation in preparation for analysis and archive. However, publishing and sharing this code is not a common practice and is hampered by the lack of thematic code registries that are designed to make code easily discoverable and reusable. Here we introduce the new Information Management Code Registry (IMCR). Its focus is on code or workflow scripts developed for common procedures that ecologists and data managers encounter when organizing, cleaning, manipulating, and archiving data. It will live in the niche between scientific analysis and modeling code and short code snippets found in online community help boards.

Activities

- **Design and implementation** is informed through consultation with community stakeholder groups.
- **Code registration** is driven by software users and developers. Anyone may contribute.
- **Software development** is done by individuals and small groups, and is informed by the community.
- **Governance** implements and maintains resources.
- **Community building** is on-going.

Features

**Discover**

- Search on software features including function, language, dependency, and science domain.
- Find single functions, workflows, and mature software packages.
- Learn about current and emerging community needs and opportunities for collaboration.

**Use**

- Access rich software metadata to fully understand application and use cases.
- Download the software and users license.
- Install the software and any dependencies.
- Access operational instructions.
- Run tests and examples.

**Contribute**

- Create a user account for the IMCR portal.
- Upload your software to the portal.
- Document your software to facilitate discovery and reuse.

**Develop**

- Develop software for extant and emerging data management challenges.
- Participate in community discussions to guide software development efforts.
- Collaborate with others with similar data management challenges.

**Portal**

The IMCR is implemented in OntoSoft, a project of the EarthCube Initiative. Ontosoft is an intelligent system to assist scientists in publishing and sharing their code to make it more reusable. The upload wizard guides software contributors through a user friendly process to capture essential software metadata. Items registered in the portal link to the respective software repositories.

**Participate**

wiki: bit.ly/IMCRwiki
portal: bit.ly/IMCRportal
email: esip-imcoderegistry@lists.esipfed.org
monthly telecons: 1st Wednesday, 14:00 EDT
contact: krvander@fiu.edu, colin.smith@wisc.edu

Acknowledgements: This material is based upon work supported by the National Science Foundation under grants #1565103 and #1629233. Any opinions, findings, conclusions, or recommendations expressed in the material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.