Service Oriented Metadata Management and Data Discovery System

Ranjeet Devarakonda, Giri Palanisamy, Jim M. Green*, Bruce E. Wilson

Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN

* On subcontract from Information International Associates, Oak Ridge, TN

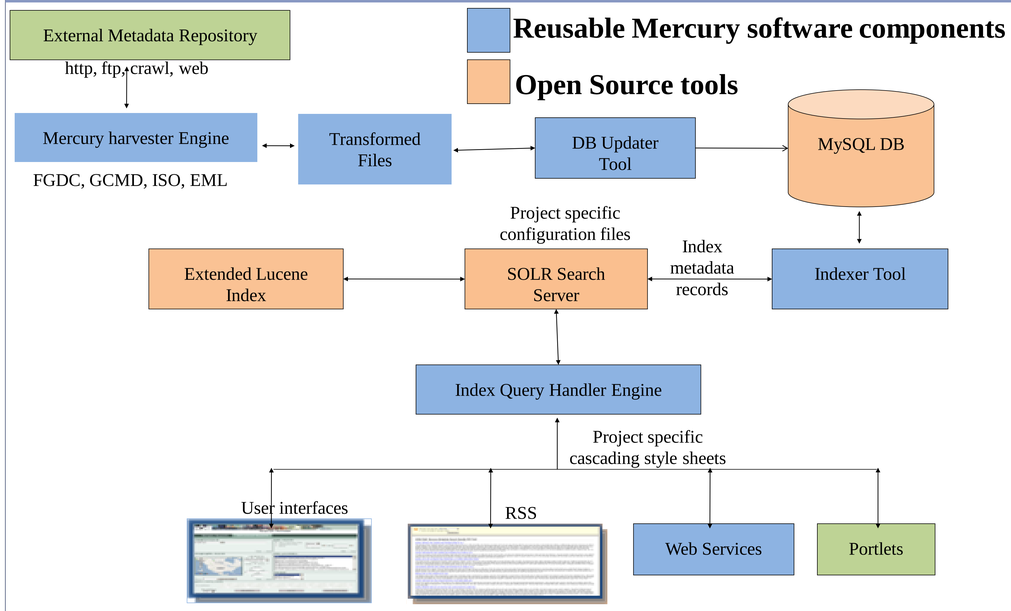
ORNL DAAC Regional and Global data Land validation Data LTER Long Term Ecological Research Network LBA Organization of Biological Field Stations bigfoot boreas boreas follow-on,canopy chemistry accp climate collections eos land validation FIFE FIFE follow-on fluxnet hydroclimatology collections model archive net primary productivity NPP NBII MAST-DC USA NPN IABIN DataONE WENDI

Introduction

Mercury is a federated metadata harvesting, data discovery and access tool based on both open source packages and custom developed software. Originally developed for single National Aeronautics and Space Administration (NASA) project, Mercury now used over fourteen different projects across three US federal agencies. It collects metadata from contributing project servers distributed around the world and builds a centralized index.

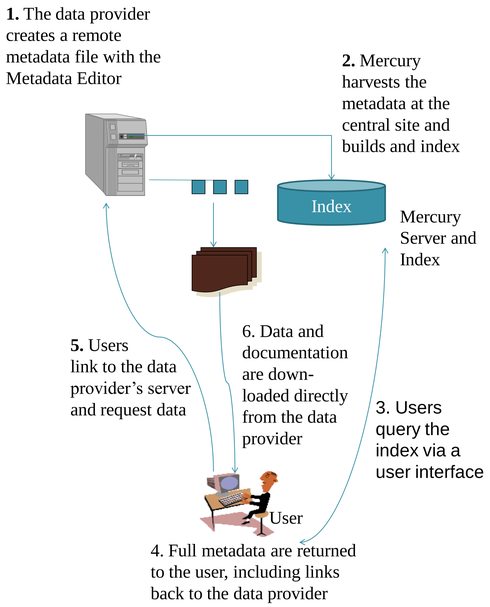
One of the major goals of the recent redesign of Mercury was to improve the software reusability across the fourteen projects which currently fund the continuing development of Mercury. Mercury has recently been extended to enable harvesting and distribution of metadata using the Open Archive Initiative Protocol for Metadata Handling (OAI-PMH).

Mercury Architecture

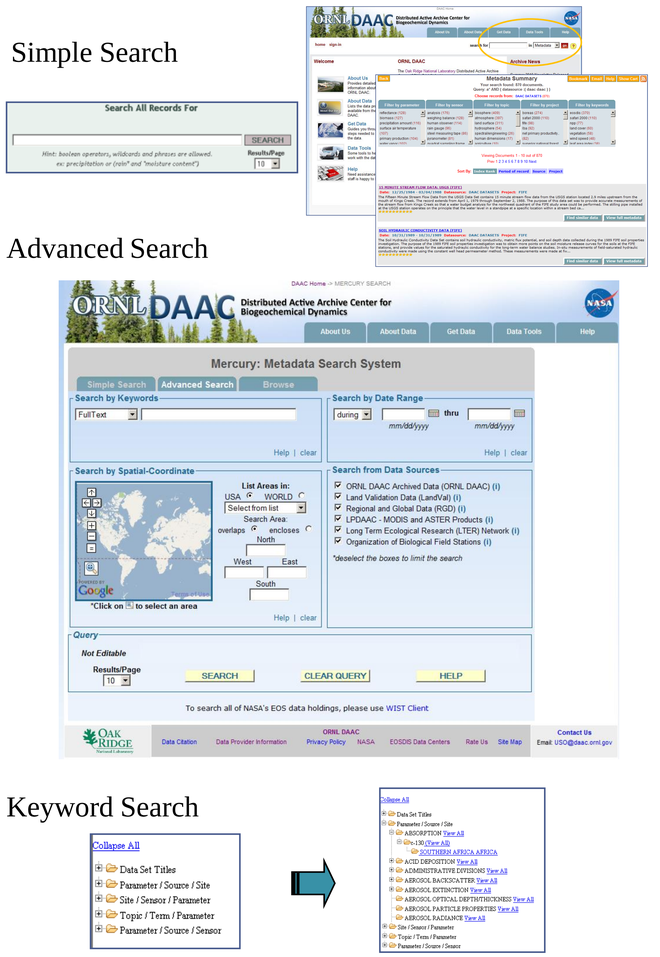


Mercury's architecture includes 1) a harvesting engine to collect various metadata records from publically available folders, web sites, ftp sites, and other network accessible locations; 2) a powerful indexing engine based on Apache Lucene and SOLR that can index billions of records; and 3) a service oriented architecture based search engine, which can perform searches and distribute results through web user interfaces, web services, RSS feed, and portlets. The harvester engine is responsible for harvesting metadata records from servers around the USA and the world. The harvester software was packaged in such a way that all the Mercury projects will use the same harvester scripts, but each project will be driven by a set of project specific configuration files. This backend component is supported by a very flexible, easy to use Graphical User Interface (GUI) that is driven by cascading style sheets. Mercury's GUI allows the users to perform simple, keyword, spatial and temporal searches.

Metadata Harvesting

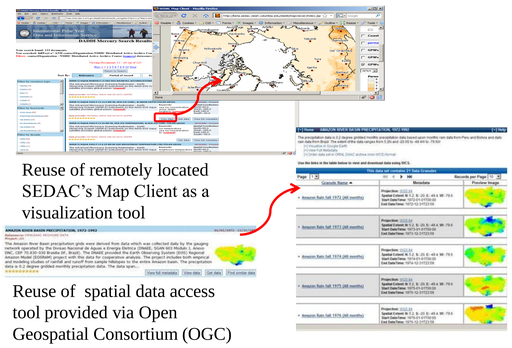


Mercury search options



External software tool reuse

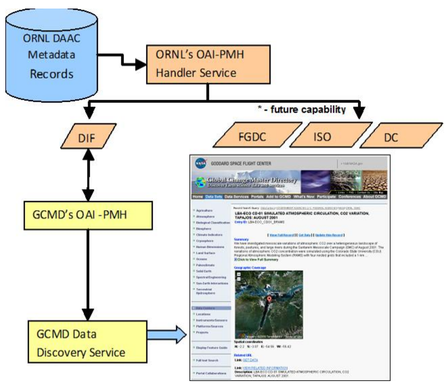
Mercury reuses several external software tools developed by ORNL DAAC and other organizations. This provides consistent results to end users, allows customization for project needs, and lowers cost to sponsors. Examples of reuse:



Reuse of remotely located SEDAC's Map Client as a visualization tool

Reuse of spatial data access tool provided via Open Geospatial Consortium (OGC) service

ORNL's OAI-PMH Overview



References

Devarakonda R, Palanisamy G, Green JM, Wilson BE, (2010) Mercury: reusable metadata management, data discovery and access system; Earth Science Informatics: Vol 3, 1-2 (87-94), DOI: 10.1007/s12145-010-0050-7

Useful Links:

Mercury home page
http://mercury.ornl.gov

ORNL DAAC Mercury
http://mercury.ornl.gov/ornldaac

NBII Clearinghouse
http://mercury.ornl.gov/nbii

Questions and Comments
mercury-support@ornl.gov