Member Node Description: KU Biodiversity Institute

Version 1.0   10/21/2013   Aimee Stewart

General

Name of resource:	University of Kansas Biodiversity Institute
URL(s):	http://bidataone.nhm.ku.edu/
Institutional affiliation(s):	Biodiversity Institute, University of Kansas
Primary geographic location:	Lawrence, Kansas, USA
Project Director & contact info:	Jim Beach, beach@ku.edu
Technical Contact & contact info:	CJ Grady, cjgrady@ku.edu
Age of resource:	Since 2005
Funding support:	US NSF and NASA
Proposed Unique Identifier:	urn:node:KUBI

Content

Content description/collection policy (1 paragraph, domain and spatial/temporal coverage, uniqueness of content, exclusions, as applicable):

Global extent predicted future climate data calculated from current and predicted monthly temperature and precipitation data. Layers contain bioclimatic variables defined by Anuclim. The data inputs for these calculations include GCM outputs from NIES for the Fourth Assessment Report of the IPCC, and Worldclim data. The data cover three aggregated future time periods, 2010-2039, 2040-2069, and 2070-2099. As we extend our DataONE integration, researchers will be able to 'publish' experiment inputs, outputs, and workflows to DataONE. The KU Biodiversity Institute will provide the mechanism, and globally unique identifiers, for cataloging data and metadata in DataONE.

Types of data (complex objects, text, image, video, audio, other):

Experiment workflow metadata, raster and vector spatial data, and tabular experiment analysis outputs.

Data and metadata availability (rights, licensing, restrictions):

http://lifemapper.org/?page_id=1096

Option for embargo (yes/no, duration):

No

Size of holdings (number and size of datasets, mean and median granules (files) per dataset):

Initially, we will catalog nine datasets of bioclimatic variables for predicted future climates under different IPCC-defined scenarios. The bioclimatic variables are derived from monthly temperature and precipitation values and follow the scheme defined by ANUCLIM spatial climatic and bioclimatic modeling software. Each dataset contains 19 files, approximately 1.5 GB in size.

Please describe recent usage statistics, if known, including information on annual data product downloads, annual number of users, annual number of data products used in publications:

User interactions

How does a user contribute data? (what can be deposited, how are data prepared, are specific software required, documentation/support available)

Users will use the Lifemapper system to build species distribution modeling or range and diversity experiments using Lifemapper archive data or their own data as inputs. Experiment inputs, outputs and metadata will be cataloged in the Lifemapper system for a limited period of time. At any time before the data expiration, will be given an option to "publish" their results which will trigger Lifemapper to generate unique identifiers, and use the DataONE APIs to deposit the data and metadata into the DataONE system.

How does a user acquire / access data?

Lifemapper will provide some inputs for users or they can upload their own. They may acquire their input data in any fashion as long as it is in a format that Lifemapper can use. They may collect their own occurrence point data or they may use data from a source such as GBIF. Climate data can come from any number of sources as well.

What user support services are available (both for depositing and accessing/using data)?

Support will be available through online documentation for using Lifemapper connection to GMN API as well as through e-mail at: lifemapper@ku.edu.

How does the resource curate data at the time of deposit?

Using the DataONE Member Node API.

Technical characteristics and policies

Software platform description, incl. data search and access API(s):

Software stack suggested by DataONE.

Service reliability (including recent uptime statistics, frequency of hardware refresh, if known):

Our data access servers have averaged about two days of downtime per year as a result of maintenance and upgrades. Hardware is refreshed every 3-4 years.

Preservation reliability (including replication/backup, integrity checks, format migration, disaster planning):

Copies of the archive are held on multiple servers, and full monthly backups are stored in a secure, tornado-proof vault managed by the University of Kansas.

User authentication technology (incl. level of create/modify/delete access by users):

We will use CI-logon for authentication. At least initially, users will only have read access.

Data identifier system and data citation policy, if available:

User of the data should cite: "Data Set Name", Lifemapper Project, University of Kansas.

Metadata standards (including provenance):

Initially we will provide Ecological Metadata Language (EML). We will extend our metadata offerings to PROV.

Capacity/services to DataONE

At what functional tier will you initially be operating? (see http://bit.ly/MNFactSheet for definitions)
- ☒ **Tier 1: Read only, public content**
- ☐ **Tier 2: Read only with access control**
- ☐ **Tier 3: Read/write using client tools**
- ☐ **Tier 4: Able to operate as a replication target**

If you can host data from other member nodes, what storage capacity is available?
[We have the capability to store up to 3 TB of data from other member nodes, negotiable.]

Can you provide computing capacity to the broader network? If so, please describe.
Lifemapper web services provide mapping and spatial data through the Open Geospatial Consortium (OGC) standards Web Mapping Service (WMS), Web Feature Service (WFS) and Web Coverage Service (WCS). Species Distribution Modeling (LmSDM) and macro-ecological Range and Diversity (LmRAD) analysis are accessed through the OGC Web Processing Service (WPS) interface standard. LmSDM services allow species distribution modeling using user or Lifemapper-provided data inputs, with a choice of Maxent or modeling algorithms provided by the openModeller framework. LmRAD services allow multi-species macro-ecological analysis of landscapes at multiple scales of time, space and species classification.

All web services are accessible through the HTTP web protocol; non-spatial outputs are retrieved through REST services. Lifemapper services can be accessed programmatically, through user or project-developed interfaces, a downloadable python client library, or a QuantumGIS (QGIS) plugin, lifemapperTools. Lifemapper web services will also allow users to catalog Lifemapper experiment metadata and data inputs, outputs, or workflows in to DataONE.

Other Services

What other services or resources (such as expertise, software development capacity, educational/training resources, or software tools) can be provided of benefit to the broader network?
Lifemapper is an open-source software project. We would be happy to consult on the use of Lifemapper client applications for species distribution and biodiversity pattern modeling and analysis.