

Member Node Description: PPBio-CENBAM

General

Name of resource:	Brazilian Program for Biodiversity Research - PPBio - Western Amazon & Center for Integrated Studies of Biodiversity in the Amazon - CENBAM
URL(s):	http://ppbio.inpa.gov.br/knb/style/skins/ppbio/
Institutional affiliation(s):	National Institute for Amazonian Research – INPA
Primary geographic location:	Manaus, Amazonas, Brazil
Project Director & contact info:	William Ernest Magnusson - bill@inpa.gov.br
Technical Contact & contact info:	Lívia Naman - livianaman@gmail.com David Valentim Dias - dvdscripter@gmail.com Flávia Fonseca Pezzini - flaviapezzini@gmail.com Debora Drucker - deboradrucker@gmail.com
Age of resource:	since 2004
Funding support:	Brazilian National Council for Scientific and Technological Development - CNPq
Proposed Unique Identifier:	urn:node:ppbio

Content

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

Data from ecological surveys, mainly from western Amazon, but also from other regions of Brazil since 2001. Some of these data packages are also available through LTER-Brasil-Site 1 repository - <http://peld.inpa.gov.br/knb/style/skins/peld/>.

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

Mainly tabular data, rarely images

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

Data and metadata are available according to PPBio's data policy - <http://ppbio.inpa.gov.br/politicadados>. Acknowledgement is required.

Option for embargo (yes/no, duration):

Yes, up to 12 months for preliminary data and 24 months for consolidated data. Sensitive data such as threatened species or species with economic value can last longer, according to the data policy.

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

435 data packages (access at 11.Jul.2014), data files of variable sizes, ranging from 2 Kb to 2 Mb.

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:

Approximately 16000 accesses/year.

User interactions

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

Data and metadata are submitted by the researcher to the data curator and go through a revision process. When its quality is assured, is submitted to the PPBio Metacat instance through Morpho software by the curator.

How does a user acquire / access data?

Data and metadata are available on the website through Metacat search services.

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

Interaction through email: ppbio@inpa.gov.br

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

Data is curated by a human before the deposit, who assures its quality and consistency.

Technical characteristics and policies

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

Metacat software.

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

Server is out for scheduled maintenance few times a year. However, unpredicted outages can happen due to the electric power instability in the Amazonian region.

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

Replication among PPBio and PELD Metacat instances at INPA. Backup is performed twice a week.

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

LDAP system. Only system administrators and curators have permission to create/modify/delete content.

Data identifier system and data citation policy, if available:

Unique id provided by Metacat system

Metadata standards (including provenance):

Ecological Metadata Language - EML

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?

2 TB

Can you provide computing capacity to the broader network? If so, please describe.

No.

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?

Translation of tools to Portuguese. Training on ecological data documentation using Morpho.