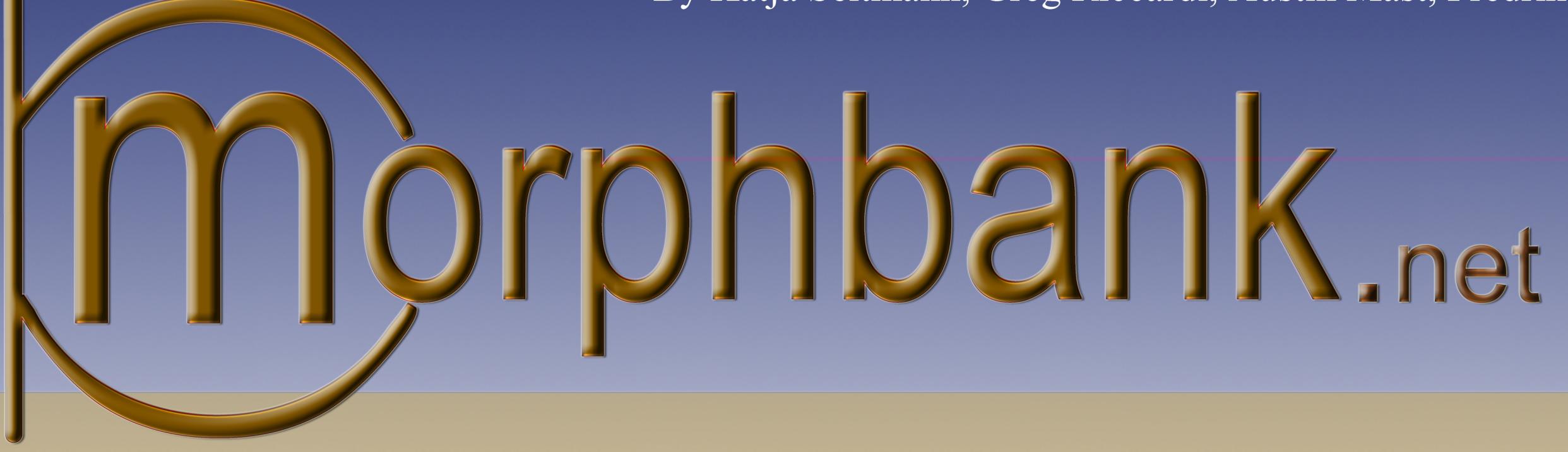


Encouraging Users to Share Biodiversity Information

By Katja Seltmann, Greg Riccardi, Austin Mast, Fredrik Ronquist, Neelima Jammingumpula, Karolina Mameva-Jakimoska, Steve Wimmer and Deborah Paul (Florida State University), and Andrew Deans (North Carolina State University).



Morphbank is an online database with image tools for scientific collaboration, research and education. Images deposited in Morphbank document a wide variety of research including: museum collections, specimen-based research in comparative anatomy, morphological phylogenetics, taxonomy and related fields focused on increasing our knowledge about biodiversity. A user base composed of researchers interested in images for biodiversity is a group with complex expectations of how an image database and collaboration platform should perform. We have seen that these underlying differences in expectations can create barriers to participation.

Who are the Morphbank Users?

Curators

Curators have large amounts of data to collect or preserve in an efficient manner. Many already have some sort of protocol in place and are interested in collaboration, backup and Web dissemination. Histories of sharing image data varies a lot and can strongly influence other groups.



Primary Taxonomists

Generally have a strong opinions about data sharing and Web dissemination, which varies depending on the cultural atmosphere related to the taxonomic group they are involved. This is often one of the most conservative sectors of the Morphbank user groups. Much of our ideals for good, robust metadata relies on their participation in online efforts.

Public Servants

Extension specialists/USDA employees and those interested in identification of agricultural or economically important organisms. This user group really wants to get the information out to the public and are often funded to disseminate information. They often must work across taxonomic groups.

Graduate Students

The future cyber taxonomists. They want total flexibility and to have a go at it themselves.

Public

Open access, attractive images and clear copyright information.

Barriers manifest themselves in how a user approaches sharing data online by an increase in desire to tightly control the use of the data after it is made public, an unwillingness to have imperfect or incomplete data, a feeling of being overwhelmed by the technology, or compete disinterest in participation. Generally by making access, upload and maintenance easier, users feel they have control over their data and will be more apt to share. Thus, effective user interface design and expanded flexibility to accommodate our complex user groups interested in participating is crucial. To encourage users we use two strategies:

1. Remove barriers in the system by moving toward increasing flexibility away from rigid requirements
2. Provide tools for data discovery, presentation and manipulation

1. Response to user concerns: removing barriers

User/institution credit

Every user agrees that data contributors should receive proper credit for their work. However ideas regarding watermarking, institution credit and copyright issues are often driven by the institution at which a researcher works or the institutions from which they receive specimens. Institution policies regarding images vary from considering them property of the public, whom originally funded the research, to a potential funding source for the museum itself.

A screenshot of a Morphbank image record for Pseudochromis pylei. It shows the user logo (a blue building icon) and mirror logo (a white building icon). Below the image are sections for 'copyright information' and 'external links on many records'. The 'External links' section includes fields for 'Institution' (Bishop Museum) and 'Project' (Pacific Basin Information Node).

To address concerns regarding branding and copyright users may:

- add an institution or user logo
- add a mirror logo
- upload images of any file size. By not requiring a minimum image size for upload concerned institutions may upload a smaller than print quality image.
- add external links
- add copyright information displayed on every image record. Watermarking must be done prior to uploading, as Morphbank does not watermark images.

Watermarking must be done prior to uploading, as Morphbank does not watermark images.

User control over data (public or not?)

Users naturally feel they are the most qualified to curate their own data. As the user base of the system has grown, Morphbank has found it necessary to increase users ability to modify their own data, prior to publication. However the ability to edit is greatly reduced once a digital object is public. This way Morphbank can guarantee that a link to a Morphbank image or record will be consistent. This is very important as we maintain these links to insert into journal articles. Users control when a digital object is published.

A screenshot of the 'My Manager (Beta)' interface. It shows a list of digital objects with columns for 'Select Mass Operation...', 'View', 'Specimen', 'Publication', and 'Image'. A red arrow points to the 'Select Mass Operation...' button. Another red arrow points to the 'View' column, which displays thumbnail images of the objects.

With the new My Manager interface curation of large numbers of digital objects will become easier.

How much data entry is too much?

What is the point where increasing requirements results in less willingness from users to upload data? The answer varies greatly and often depends on how many images a user wants to upload. We have approached this barrier in four ways:

- reduced strict metadata requirements. Allowing users to upload the metadata they actually have or need
- making the process less form intensive. Use of mass changes, the ability to associate existing View, Specimen, Publication and Location records with a new Image, and increased use of AJAX is reducing the number of forms a user must complete.
- user manual, help desk and workshops. Having a contact person to answer questions and provide feedback concerning data management can prevent frustration.
- provide streamlined workflows for larger projects (using ftp, and bulk upload)

Data persistence

A screenshot of a 'Mirror Server Information' interface. It shows fields for 'Mirror URL', 'Local Path', 'Remote URL', and 'Current content'. A red arrow points to the 'Manage Mirror Content' button. Below the interface is a preview window showing a thumbnail of an image labeled '62247.jpg'.

Becoming a Morphbank mirror is relatively simple, only requiring an ftp account. This simplicity allows more users to maintain copies of their own images and actively participate in the archive process.

Entering data is time consuming and users weigh the impression of how persistent the data will be versus how much time it takes to work with the system. We have made efforts to show users their data will persist by:

- giving them the ability to access and maintain their own images through mirroring
- backup of images and data nightly
- housed in the School of Computational Sciences with trained systems administrators maintaining the machines
- persistent URLs with little Website down time

Taxonomic names

The management of names is one of the greatest challenges in digital archiving of biodiversity data. We use the Integrated Taxonomic Information System (ITIS) as our primary name and classification source. However, to accommodate names not in ITIS and names not yet published, we have created the ability to add Morphbank temporary names. Corrected names and new names with publication information are sent to ITIS for review, allowing users to directly contribute to the community management of taxon names. Names that users are not prepared to submit to ITIS can still be added and used in the Morphbank system.

Taxon names and specimen determinations can both be annotated, allowing for peer review of names and name associations.

A screenshot of the 'Taxon Name Search' interface in the Integrated Taxonomic Information System (ITIS). It shows a search bar and a table of results with columns for 'Name', 'Classification', 'Status', and 'Last Published'. A red arrow points to a 'New Annotation' button.

2. Tools users have found particularly useful

A screenshot of the 'My Collection' interface. It shows a grid of images with labels like 'Osteichthyes general body plan (117) images' and 'Check All' buttons. A red arrow points to a 'Copy Checked Objects...' button.

Collection software linking to publications

A screenshot of the 'Handbook of Nearctic Chalcidoidea' interface. It shows a search bar and a detailed record for 'Histeromerus elongatus', including images of the insect and its larva, and a 'Display Information...' button.

Using Morphbank ids to embed images in a user's own Web pages

A screenshot of a 'LinkOut provider' interface. It shows a detailed record for 'Histeromerus elongatus' with a 'Send Link' button.

LinkOut provider

A screenshot of an 'Annotation Record' for 'Antennal insertion'. It shows a close-up image of an antennal insertion with a red arrow pointing to it, and a detailed record with fields for 'Base Obj Info' and 'Object Info'.

Annotation software

A screenshot of an 'Annotation Record' for 'AGC Voucher Number'. It shows a large image of a butterfly with a red arrow pointing to its antenna, and a detailed record with fields for 'Base Obj Info' and 'Object Info'.

Examples of some tools users have found most beneficial. Being an image archive for museums and large collaborative projects presently attracts the bulk of our users with large image sets. The ability to link a collection of images to a journal publication is the primary reason users with small numbers of images presently participate in Morphbank.



Funding for Morphbank provided by the Biological Databases and Informatics program of the National Science Foundation (Grant DBI-0446224)