Conference Abstract

International Image Interoperability Framework: A unified approach to sharing images of natural history specimens?

Roger Hyam ‡

‡ Royal Botanic Garden Edinburgh, Edinburgh, United Kingdom

Abstract

Researchers have become accustomed to online access to data about specimens held in natural history collections. Over several decades, metadata standards have been developed to facilitate the sharing and aggregation of this data, notably Darwin Core and ABCD (Access to Biological Collections Data) developed under the auspices of TDWG but other standards developed in other communities, have also proved useful notably EML (Ecological Metadata Language) and GML (Geography Markup Language).

Data aggregators have arisen to both, drive standards development and take advantage of the vast number of records made available through this community effort. Examples include Atlas of Living Australia and spin off Atlas projects, EoL (Encyclopedia of Life), iDigBio, Global Biodiversity Information Facility (GBIF), WFO (World Flora Online).

There are still many "dark specimens" that are not visible to the web and efforts continue to digitise metadata on these objects and make them available. The vast majority of the data that has been liberated so far, has therefore been text based and the standards reflect this, although many institutions and projects are also producing large numbers of images and other media.

There have been media extensions to some standards to accommodate the sharing of images and other multimedia formats. However, these are restricted to metadata about media objects rather than the exchange of media objects themselves. For example, two extensions to Darwin Core are Audubon Core, which is designed to "determine whether a particular resource or collection will be fit for some particular biodiversity science application before acquiring the media." and the Simple Multimedia extension, which is a "simple extension for exchanging metadata about multimedia resources". Therefore image exchange, in particular, has not used open standards. Projects have relied on transferring high resolution versions of images (e.g. submission of type specimen images to JSTOR) or cut down compressed versions (e.g. many herbarium specimens submitted to GBIF or Europeana).

The network has not allowed access to high resolution versions of images as curated by the host institutions themselves beyond basic links to web pages. If high resolution images have been published in online catalogues, they have been made available using a hotchpotch of different technologies including the now defunct Java Applets and Adobe Flash player. The network has not supported different views of the same specimen or annotations of those views, or integration of audio and moving images.

In an ideal world a researcher should be able to view and annotate images of specimens held across multiple collections in a unified way, and the host institutions should have access to those annotations and statistics on how their specimens are being used. How can we achieve this?

The sharing of multimedia representations of objects online is not a problem unique to the biodiversity community. Scholars in museums and archives of all kinds are facing the same issues. In 2011 the British Library, Stanford University, the Bodleian Libraries (Oxford University), the Bibliothèque nationale de France, Nasjonalbiblioteket (National Library of Norway), Los Alamos National Laboratory Research Library, and Cornell University came together to develop an exchange standard called IIIF (International Image Interoperability Framework). This framework now consists of six APIs (Application Programming Interface), four stable and two in beta, to publish and integrate image and other multimedia resources in a uniform manner and has been adopted by many institutions and commercial partners in the digital humanities. Applications based on IIIF enable many of the features desired by biodiversity researchers.

The notion of sharing and annotating specimen images is not new to the natural history community. MorphBank, founded in 1998, has grown to allow much of this desirable functionality but at the cost and fragility of being a centralised database. The question we should perhaps be asking is: how can we make the biodiversity data sharing network as a whole more like MorphBank?

From 2019 to 2021, part of the EU-funded Synthesys+ programme will support the adoption of IIIF as a unified way to publish images of natural history specimens. We aim to have a set of exemplar institutions publishing IIIF manifests for some millions of specimens by the end of the project and one or more demonstration applications in place. We hope

this will act as a catalyst for wider adoption in the natural history community. A key goal is to integrate image data served using IIIF with metadata available via CETAF (Consortium of European Taxonomic Facilities) specimen identifiers. If IIIF were ubiquitous in the natural history community, building tools that implemented this functionality would be feasible.

A brief demonstration of a herbarium specimen browser, Herbaria Mundi, will be given. It will illustrate how specimens hosted in different institutions can be manipulated in a single interface. The architecture that supports this behaviour will be explained and its challenges, by implementing the institutions discussed.

Keywords

CETAF ID, IIIF

Presenting author

Roger Hyam

Presented at

TDWG 2020

Funding program

Task 4.3 of the EU Commission funded Synthesys+

Hosting institution

Royal Botanic Garden Edinburgh

Ethics and security

n/a

Conflicts of interest

None declared