OPEN ACCESS

Conference Abstract

Semantically Linking Specimens and Images

Roger Hyam ‡

‡ Royal Botanic Garden Edinburgh, Edinburgh, United Kingdom

Abstract

Many of the world's natural history collections are creating high resolution digital images of their specimens. They often make these available on the web through some form or zoomable viewer. For historical reasons, a hotchpotch of technologies are used to achieve this. This diversity has lead to two issues. Firstly, maintenance becomes costly as technologies need replacing. Secondly there is little chance to share data between institutions or provide a unified user experience. A researcher visiting four different virtual collections may have four very different experiences.

Similar issues exist in the archives and libraries disciplines. They also need to share high resolution, annotated images of the physical objects in their care. In response to this issue many have coalesced around the International Image Interoperability Framework (IIIF).

IIIF is a set of shared application programming interface (API) specifications for interoperable functionality in digital image repositories. It separates the notion of a viewer, which may be used as part of a website or other application, and the web services that feed data to that viewer. By using a common API for serving data about images, different viewers can be used to view the same images, thus providing an upgrade path that does not require replacement of viewer and server software at the same time and allows different viewers to be used for the same image data. Potentially more importantly, it facilitates the construction of applications that view data from different collections as if they were in the same place. From the researcher's point of view, the experience could be the same whether the virtual specimen is hosted locally or in a museum on another continent.

There is one important thing that has been deliberately omitted from the IIIF standard. This has both enabled its rapid adoption but also makes it incomplete for building research applications. IIIF transmits no semantic data about the subject of the images, only labels. The IIIF data therefore needs to be bound to semantically rich data about the specimens being viewed, in some uniform way.

Consortium of Taxonomic Facilities (CETAF) specimen identifiers are now widely adopted by natural history collections in Europe. Each individual collection object is designated by a URI chosen and maintained by the institution owning the specimen (Groom et al. 2017, Güntsch et al. 2018, Güntsch et al. 2017, HYAM et al. 2012). Under Linked Data conventions, content negotiation is used at the server so that users accessing an object using a web-browser are redirected to a human-readable representation of the object, typically a web-page, whilst software systems requiring machine-processable representations are redirected to an RDF-encoded metadata record. CETAF specimen identifiers are therefore ideal partners for IIIF representations of specimens. But how should we join the two together in a semantically rich way that will be generally understandable?

SYNTHESYS+ is a European Commission funded programme that facilitates collaboration and network building among European natural history collections. It is concerned with both physical and virtual access to the 390 million specimens of plants and animals housed in participating institutions. Under Task 4.3 of this project, we have been working to create a reliable way to link between the RDF metadata about specimens and images of those specimens in IIIF as well as from images of specimens back to metadata of those specimens. By January 2021, we aim to have ten exemplar institutions publishing IIIF manifest files linked to CETAF identifiers for a few million specimens and for this to act as a catalyst for wider adoption in the natural history community. This presentation gives an update on the rollout of these implementations, paying particular attention to the challenges of semantically annotating specimens with images.

Keywords

linked open data, RDF, identifiers, CETAF, IIIF, images, multimedia

Presenting author

Roger Hyam

Presented at

Biodiversity Next 2019

Funding program

European Commission:SYNTHESYS+

Hosting institution

Royal Botanic Garden Edinburgh

Conflicts of interest

No conflicts

References

- Groom Q, Hyam R, Güntsch A (2017) Stable identifiers for collection specimens. Nature 546 (7656): 33-33. https://doi.org/10.1038/546033d
- Güntsch A, Hyam R, Hagedorn G, Chagnoux S, Röpert D, Casino A, Droege G, Glöckler F, Gödderz K, Groom Q, Hoffmann J, Holleman A, Kempa M, Koivula H, Marhold K, Nicolson N, Smith V, Triebel D (2017) Actionable, long-term stable and semantic web compatible identifiers for access to biological collection objects. Database 2017 https://doi.org/10.1093/database/bax003
- Güntsch A, Groom Q, Hyam R, Chagnoux S, Röpert D, Berendsohn W, Casino A, Droege G, Gerritsen W, Holetschek J, Marhold K, Mergen P, Rainer H, Smith V, Triebel D (2018) Standardised Globally Unique Specimen Identifiers. Biodiversity Information Science and Standards 2 https://doi.org/10.3897/biss.2.26658
- HYAM R, DRINKWATER R, HARRIS D (2012) <p class="HeadingRunIn">Stable citations for herbarium specimens on the internet: an illustration from a taxonomic revision of Duboscia (Malvaceae)</p>. Phytotaxa 73 (1). https://doi.org/10.11646/phytotaxa.73.1.4