Annotate What?

- Curatorial and scientific metadata for an estimated 3 bn. specimens in the world's natural history collections. 1 bn in U.S.
- Digital record of about 1.5% at most. See gbif.org (Also includes non-captured observations).
• Ongoing programs in U.S., Europe, Australia,...to digitize all.

• In U.S., 10 year, $500 M effort now underway; automated; semi-automated;
  – QC issues; correlations to other paper and digital resources; intentionally changing data
The Problem: Data Quality

- Collections & occurrence data is all over the map
  - ... literally (off the map!)
  - ... even after “digitization” (sensu stricto)
  - ... and after “digitization” (sensu lato), a.k.a. “computerization”

Issues:
Lat/Long transposition, coordinate & projection issues
Data entry/creation, “fuzzy” data, naming issues, bit rot, data conversions and transformations, schema mappings, ... (you name it)
(1) Kvetch about data

(2) Push to interested parties

(3) Filter

(4) Change data in databases

(5) Store all assertions
Modeling distributed annotations of mutable, actionable, distributed data

- OA is *almost* adequate
- clear distinction between annotation model and domain models;
- provenance support
- web document-centric features are optional
- SpecificResource model fits data annotation well
- Use of CNT for machine-opaque resources
- (RDF is friendly to semantic pub/sub)
Modeling distributed annotations of mutable, actionable, distributed science data

- OA is *almost* adequate

- Model of *Evidence* for assertions in Body.

- Model to convey annotator's *Expectation* to consuming apps

- Model use of queries as Selectors

(-) Model “transcription” as an oa:Motivation

(?) Model domain associations between Target and Body
:anAnnotation a oa:Annotation;
   oa:hasTarget :nsptarget1;
   oa:hasBody <http://fp.org/rangeViolation> ;
   oa:hasBody :invalidLatitudeText.

:nsptarget1 a oad:AnySuchResource;
   oa:hasSelector :findInvalidLatitudes.

:findInvalidLatitudes a oad:SparqlQuerySelector;
   oad:hasQuery "select distinct ?x WHERE {
      ?x a dwcFP:Occurrence.
      ?x dwc:hasLatitude ?lat.
      filter (?lat > 90 || ?lat < 90).}"

   [...].
Thank you

http://filteredpush.org confusing page for now

http://wiki.filteredpush.org

Video linked there; links to pubs, svn repo, etc.

morris.bob@gmail.com

Also: Jim Hanken, James Macklin, Bertram Ludaescher, Paul Morris, Maureen Kelley, David Lowery