Moving to Sharing Relationships among Things

Greg Riccardi

Florida State University

iDigBio

griccardi@fsu.edu

Goals of presentation

- How to move from personal to shared data
- Relationship properties and identifiers
- Finding and inventing relationship properties

Imagining data from *The Beak of the Finch*

- Information about individual birds

tag	sex	mother	father	birthdate
1154	M			
1158	F			
1160	F			
1188	F			
1207	M	1158	1154	1/1/1975
1008	M	1188		2/12/1976
1330	F	1188	1207	3/15/1978
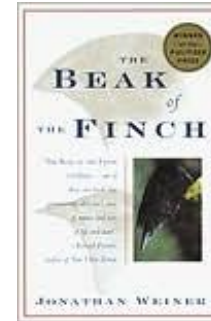
- Measurements

tag	date	beak1	beak2	wt	notes
1158	12/15/1974	25	12	30.5	
1158	1/15/1975	25	12	28.3	
1158	5/21/1978	25	12	32.1	
1207	1/1/1975			2.3	newborn
1207	2/5/1975	5	3	6	

Information for a single bird

- Example bird

Finch		
id	**property**	**value**
1207	sex	M
1207	mother	1158
1207	father	1154
1207	birthdate	1/1/1975

- Example measurements of the bird

Measurement		
id	**property**	**value**
	bird id	1207
	date	2/5/1975
	beak1	5
	beak2	3
	wt	6

  - No identifier for the measurement!

Exporting information

- Step 1: improve identifiers

Finches					
Id	**tag**	**sex**	**mother**	**father**	**birthdate**
http://beaks.org/finch/1154	1154	M			
http://beaks.org/finch/1158	1158	F			
http://beaks.org/finch/1160	1160	F			
http://beaks.org/finch/1188	1188	F			
http://beaks.org/finch/1207	1207	M	1158	1154	1/1/1975
http://beaks.org/finch/1008	1008	M	1188		2/12/1976
http://beaks.org/finch/1330	1330	F	1188	1207	3/15/1978

id	**finch**	**date**	**beak1**	**beak2**	**wt**	**notes**
http://beaks.org/measurement/113	http://beaks.org/finch/1158	12/15/1974	25	12	30.5	
http://beaks.org/measurement/114	http://beaks.org/finch/1158	1/15/1975	25	12	28.3	
http://beaks.org/measurement/115	http://beaks.org/finch/1158	5/21/1978	25	12	32.1	
http://beaks.org/measurement/116	http://beaks.org/finch/1207	1/1/1975			2.3	newborn
http://beaks.org/measurement/123	http://beaks.org/finch/1207	2/5/1975	5	3	6	

Exporting information

- Step 2: Improve properties
  - Find specific properties to use
    - *sex* replaced with *dwc:sex*
    - *mother* replaced with ?
  - Define new properties as needed
    - E.g. *beak1* and *beak2* are specific to this measurement
    - Create definitions and register with some repository

Moving to standard vocabularies

- Formal processes for defining properties
  - A property is an object, with identifier
    - Resource valued property is a type of relationship
    - Property object has its own (property, value) set
  - E.g. *beak1* is a measurement of beak height
    - Properties defined include
      - Units of measure
      - Morphological feature measured
- Social processes for agreeing on properties and values

iDigBio
Integrated Digitized Biocollections

Media Vocabulary: Audubon Core

- Properties of a media object
  - http://terms.tdwg.org/wiki/Audubon_Core_Term_List
- Examples
  - dcterms:identifier
    - Unique code of the media object
  - dc:type
    - Recommended terms are Collection, StillImage, Sound, MovingImage, InteractiveResource, Text
  - xmpRights:UsageTerms
    - The license statement defining how resources may be used
  - ac:associatedSpecimenReference
    - A reference to a specimen associated with this resource.

Moving to common properties

- iDigBio is an aggregator without a fixed schema
  - People send information to the repository using their own schemas
  - iDigBio import tools transform into common structure
  - Providers do not always use standard property vocabularies
- How can we enable integration and interoperability?
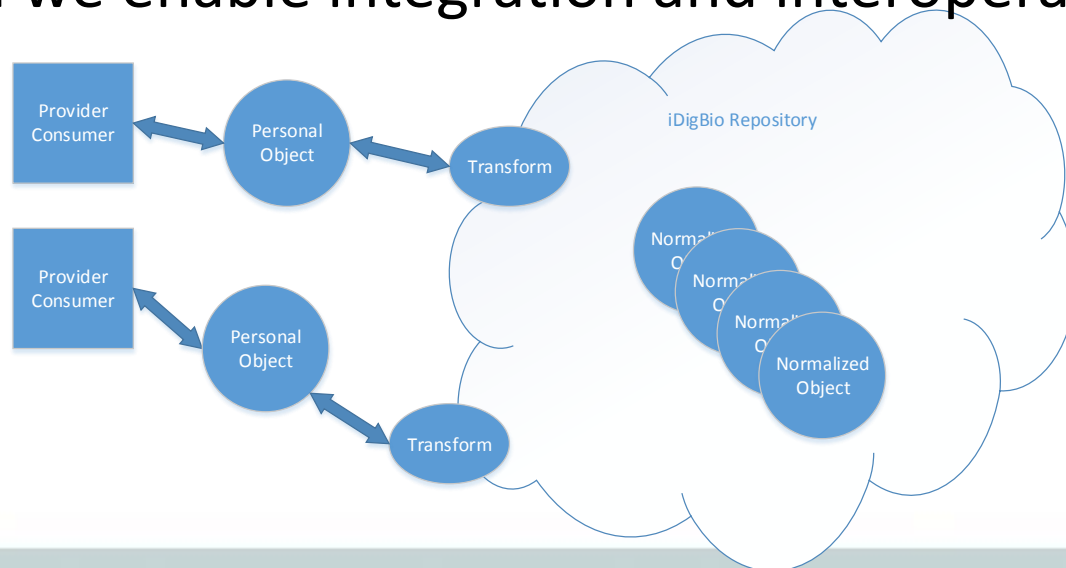
Integration of properties

- iDigBio keeps a list of properties
- Users agree on similarity of properties
  - Different names for same property
  - Different properties that are similar

Relationships as annotations

- A relationship that needs its own properties
  - When, who, why, what evidence
- An annotation (*url1*) is an assertion of properties for objects
  - **On** 4 October 2013, Joe **claims** that specimen (*url2)* is of species *url3* **because** he disagrees with the determination on the label, and for **evidence**, he offers a set of image annotations (*url4*) showing morphological features that can be seen in the photograph (*url5*).
- Many relationships are expressed in the annotation

Conclusions

- Must have identifiers for objects
  - Especially occurrences
- Must have agreement on properties
- Must have strategy for representing relationships
  - In provider databases
  - In repositories
  - In transit

Identifiers

- Lots of uncertainty in community
  - What form of identifiers, what services to provided, etc.
- We need to
  - Emphasize identification of specimens and other objects
  - Help providers to see value of specimen identifiers
  - Remove obstacles to adoption
    - E.g. validate and advocate standard practice in collections managers
  - Move forward in spite of problems
- Current suggestion
  - UUID as basis of identifier
  - URI with embedded UUID
    - urn:uuid:f47ac10b-58cc-4372-a567-0e02b2c3d47
    - ark:/87286/B2/f47ac10b-58cc-4372-a567-0e02b2c3d479

Constraints, needs and expectations: Suppliers

- All objects uniquely identified
- Suppliers help with data/property mapping
  - standard/s
- Suppliers need to populate ownership and repository fields
  - dwc:institutionCode, dwc:ownerInstitutionID
- Suppliers will need help with multiple identifiers to
  - store
  - share

Constraints, needs and expectations: Users

- Future expectation example: Morphbank
  - Download from iDigBio in Morphbank
    - annotate
  - Distribute annotations to all interested users
- Integration
  - feedback
  - attribution
- Community registration – who are you?
- Providers need to know when data is used.

Help with Specimen Information

- Multiple identifier resolution
  - Keep track of 'sameAs' relationships
  - Accept requests by any known identifier and return all data
- Reconciling fields
  - Have a community network to create best version of specimen information
  - Working with BiSciCol on keeping track of all the things that are related
    - Gut contents and metagenomics
    - DNA sequences
    - Analyses that use the specimen
    - Publications that reference the specimen