Development Considerations for DESC

Jefferson Heard and Mary Whitton
RENCI, University of North Carolina

Webinar
June 6, 2013

renci

TECHNOLOGY CONSIDERATIONS

Information System

- Data Model
- Presentation
- Access control & methods
- Business logic & validation
- Policy enforcement
- Federation

Authentication and Authorization

- Authentication certifies that you are you
  - Technologies: OAuth, OAuth2, OpenID
- Authorization determines your permission to read or make a change
  - UNIX style groups, permissions, ACLs, etc.

Validation

- Validation strategies ensure data in canonical records is correct.
- Automatic validation – ensures data types are correct, required fields are filled in, and won't break the system
- Manual validation – queues new entries for review and sign-off

Access methods/RESTful

- REST is an access paradigm that maps common operations to HTTP in a known way
- Create
- Read
- Update
- Delete
- List
- Search

Collection Organization

- Federated collections allow organic growth and can exist without central authority and smaller funding.
- Centralized collections allow better consistency of information and information system versioning

Approach and Technology Choice

- Information system architecture
  - Web app framework such as Geoanalytics, Django, Rails, GeoNetwork, OpenGeoPortal
- Metadata store
  - Triple store?
  - RDMS?
- Digital files storage
  - Association with samples

POLICY CONSIDERATIONS

Governance

- Who decides policies?
- What data to be collected about users?
  - To understand usage patterns to guide future development
  - ….

Curation and Archiving

- Accession of samples
  - Approving entries (gatekeeper on samples/data)
    - Samples and associated digital data files
      - Digital file standards, e.g. image format, resolution, and allowable compression algorithms
  - Metadata
  - Creating entries for progeny samples (provenance)
- How to establish were a sample (and any progeny) is now?
  - Within repository
  - Outside repository
- Deaccession of samples

Business Rules

- Who can initiate <*service*>?
  - Access
  - Search
  - …..
- What's different between physical and digital objects?

Provenance tracking

- Definition of what information needs to be included

Sustainability Model

- What resources do you expect to expend on
  - System maintenance
  - Sustaining engineering: updates
  - New features (next slide)
- Avoid fragmentation (if federated….)

Extensibility for new features

- Support larger volume
- Add new data types
- Add new metadata
- Interoperability with other sample data management systems