WWARN Informatics

Chassis/WWARN Data Submission and Curation Workflow

Chassis/WWARN

The WWARN Application

data

submission

submitter

curation

curator

metadata (SSQ)

- system for **data management** = submission + curation of data & metadata

- system for analysing, visualising and reporting curated data

Chassis/WWARN Design Goals

- Used by submitters & curators
- Enables upload of data in any format
- Provides framework for curation
- Captures complete audit trail
- Manage privacy
- Manage quality

The WWARN Application

- Enables visualisation, analysis & reporting for curated data
- Used by submitters
- Used by other researchers, public health officials, ...

(Out of scope for today)

Roadmap - Chassis/WWARN 1.0

- Release first production quality version of Chassis/WWARN, **ASAP** (ideally early 2010)
- First release - limited functionality
- Add more functionality **incrementally**, in subsequent releases, **based on feedback**
- **Chassis/WWARN 1.0 - bare minimum to enable WWARN to begin operations**

This Meeting

- Illustrate the data submission & curation **workflow model** that we are proposing for Chassis/WWARN 1.0
- (i.e., who does what, when)
- does it make sense?
- have we missed anything vital?

Use Case - shortest path, end-to-end

- submitter asks for a user account to submit their data to WWARN
- administrator creates submitter user account
- submitter logs in to Chassis/WWARN
- submitter creates a new study (title, summary, module(s))
- submitter creates a new submission (title, summary, module(s), study)
- submitter uploads submitted data
- submitter creates a data dictionary (title, summary, module(s))
- submitter uploads data dictionary file
- submitter links submitted data to data dictionary
- submitter requests gatekeeper review
- gatekeeper reviews submission
- gatekeeper accepts submission for curation
- gatekeeper assigns curator
- curator downloads submitted data & data dictionary
- curator [offline] transforms data to standard format
- curator uploads new revision of submitted data
- curator enters log message describing changes, links new revision to previous revision
- curator creates a standard data dictionary (title, summary, module(s))
- curator links new revision of submitted data to standard data dictionary
- curator creates release criteria specification (title, summary, module(s), file)
- curator reviews data, approves for release

Study Questionnaire

- ... can bypass?

Submitted Data

- ... any format?
- ... can live without on-line tools for preview, at least for Chassis/WWARN 1.0?

Gatekeeper Review

- ... can decide not to accept a submission for curation?
- ... prioritise?
- ... each module can define their own acceptance criteria?
- ... what happens if submission has data for >1 module?
- ... need to request?

Data Dictionaries

- ... required prior to gatekeeper review?
- ... files or forms?

Revising Submissions

- ... submitters can revise submitted data prior to gatekeeper review?
- ... after gatekeeper review?

Studies and Submissions

- ... one study, one or more submissions?
- ... one submission, one or more studies?

Submission Permissions

- ... grant ownership (read/write)?
- ... grant read-only access?

Study Permissions

- ... visible, discoverable, private?
- ... grant ownership (read/write metadata, submit)
- ... grant submission authority (link submissions to study)
- ... link submission without authority?
- ... managing complexity (knock-on permissions)?

Dictionary Permissions

- ... keep some dictionaries private?
- ... discoverable?

Curation & Derivation

- ... off-line curation, download and upload of files ... can live with it?
- ... how capture the audit trail (derivation)?

Validation & Transformation

- ... can live with off-line validation & transformation, at least for Chassis/WWARN 1.0?

Release Criteria

- ... review w.r.t. release criteria?
- ... finer grained tagging w.r.t. individual criteria?

Curator Collaboration

- ... can any curator work on any submission? If not, how manage permissions?
- ... need to keep track of who has the ball?
- ... parallel/series?

Publishing Data (Open Access)

- ... making a submission public?
- ... making a study public?
- ... who decides? Study owner? Submission owner?