Universal Collaborative Annotations with Thin Clients:
Supporting User Feedback to the Atlas of Living Australia

Ron Chernich
Principal Research Fellow
University of Queensland, Australia

Background

- **Atlas of Living Australia (ALA)**
  a five-year project funded under the Australian Government’s National Collaborative Research Infrastructure Strategy (NCRIS).

- Its mission is to develop a **biodiversity data management system** which will link Australia’s biological knowledge with its scientific and agricultural reference collections and other custodians of biological information.
Interlude

- The ALA does not:
  - own the data
  - provide the data
  - curate the data

- The ALA does:
  - facilitate data discovery through a single portal
    - Assist users with Biosecurity, Land-use planning, Conservation, Climate-change, Taxonomy, Leisure…
  - assist the data providers improve the quality of their data

Australian GBIF* Data: Intersecting

* Global Biodiversity Information Facility:  http://www.gbif.org/
Babies and Bathwater?

- Misspellings:
  - Ornithorynchus / Ornithorhynchus?
  - Mt. Tambourine / Mt. Tamborine?
- Coordinate problems:
  - Positive values for South or West
  - Latitude / Longitude transposed
  - Coordinates not near Locality
  - Unknown precision
- Other issues:
  - Same record shared through different routes
  - Unknown collecting strategy

...with thanks to Donald Hobern – ALA
Background (resumed)

Data Integration and Annotation Services – Biodiversity (DIAS-B)

- Services to support a metadata portal like GBIF that facilitates discovery of existing and new Australian biodiversity data collections
- Services that enable annotations of data and metadata to be made by researchers and the public that improves the quality of the underlying data

UQ eResearch Labs

- Working with collaborative annotations and metadata related services since 2004
- Experience with tools and techniques for annotating text and multi-media resources
- Began development of a “clean slate” annotation service for DIAS-B in October 2008
  - Prime user: the ALA
  - Open source and generic enough for others
DIAS-B Annotation Goals: Year 1

- Investigate existing collaborative annotation systems and select or create the most appropriate solution
- Investigate requirements for annotation services in other NCRIS capabilities and in ANDS
- Prototype candidate solutions
- Deliver pre-production quality (or better) service with client side widgets that can be incorporated into the ALA and other services.

What is an “annotation”? 

- Additional information attached to an existing resource (which may itself be an Annotation!)
  - Simple Example:
    - an email chain
    - A bookmark
  - Less Simple Example:
    - A comment about a web page
    - A highlighted selection on a web page with a comment
    - A structured correction to a data set
  - Complex Example:
    - An identified region of a multi-media resource
    - A specific view and zoom of a protein model
On the Server

- Store and retrieve structured annotations
  - Robust
  - Efficient
  - Scalable
  - Extensible
- Provide harvesting through OAI-PMH
  - including sets and resumption tokens
- Provide search capability
- Integrate with Australian Authentication Federation (AAF) when it comes on-line

W3C Annotea

- A W3C Semantic Web initiative (draft) that defines a protocol for creating, updating, and retrieving annotations using an extensible RDF XML schema having 7 properties:
  - Annotates (resource)
  - Author
  - Body
  - Context
  - Created/Modified (date)
  - Related
At The Client

- Create and display annotations
  - A whole web page
  - Selected text on a web page
  - A selected region of an image on a page
- Reply to annotations and other replies
- Support context sensitive annotation schemas
- Support specially authored pages and pages which are ignorant of the annotation service

The Ross Wilkinson Challenge:

- **NO PLUG-INS!**
  - They are browser and version specific
  - They require installation and configuration
  - They compromise user security and integrity
  - They are Verboten by many IT departments
Surprise! [1]

- All objectives met
  - Firefox
  - Safari
  - Google Chrome
  - Opera
  - Internet Explorer 8

- Now the Dreaded Live Danno Demo...

The developers were the ones surprised

Take-Away

- **Danno** and **Dannotate** allow viewers of specially prepared web pages to shared annotation at page, text string, and image region levels
- The **Danno Repeater** extends this ability to many unprepared web pages
- The components are open source (LGPL) and are being actively developed and refined
Availability

- **Danno** and **Dannotate** - open source (LGPL)
  - **Danno** (server)
    - 100% Java
    - Servlet (Apache Tomcat or similar)
    - Uses RDF triple store (Sesame or Jenna)
  - **Dannotate**
    - Java servlet talks to Danno and formats responses
    - JavaScript client provides user interface
  - **Danno Repeater**
    - Bookmarklet to Dannotate for unadorned pages


Key People

- **University of Queensland**
  - Prof Jane Hunter
    - Principal Investigator
  - Dr Stephen Crawley
    - Senior Research Fellow
  - Mr Ron Chernich
    - Principal Research Fellow
Sponsors

- Australian National Data Service (ANDS)
- Australian Research Collaboration Service (ARCS)
- Australian Collaborative Research Infrastructure Strategy (NCRIS)
- National eResearch Architecture Taskforce (NeAT)
- CSIRO and the ALA

Questions?

- chernich@itee.uq.edu.au
- scrawley@itee.uq.edu.au
- jane@itee.uq.edu.au