Bonolo*
A General Digital Library System for File-based Collections

Lighton Phiri Kyle Williams Miles Robinson
Stuart Hammar Hussein Suleman

Digital Libraries Laboratory
Department of Computer Science
University of Cape Town

November 13, 2012

*Sotho word meaning easy.

Digitization - martinwest.uct.ac.za

Holy United Methodist Church in SA (52). A young member.

Apostolic Full Gospel Mission of South Africa (2). Special Sunday service in Soweto - held in a marquee, to celebrate the 100th birthday of one of its members.

Digitization - lloydbleekcollection.cs.uct.ac.za

Motivation

- Preservation costs
  - Preservation lifecycle
  - Heritage funding model
- Technical skills and education
  - Content curators skillset
  - Steep learning curve for most solutions
- Internet bandwidth
  - Bandwidth intensive solutions
  - Cloud-centric solutions not feasible
- Existing solutions
  - Complexity

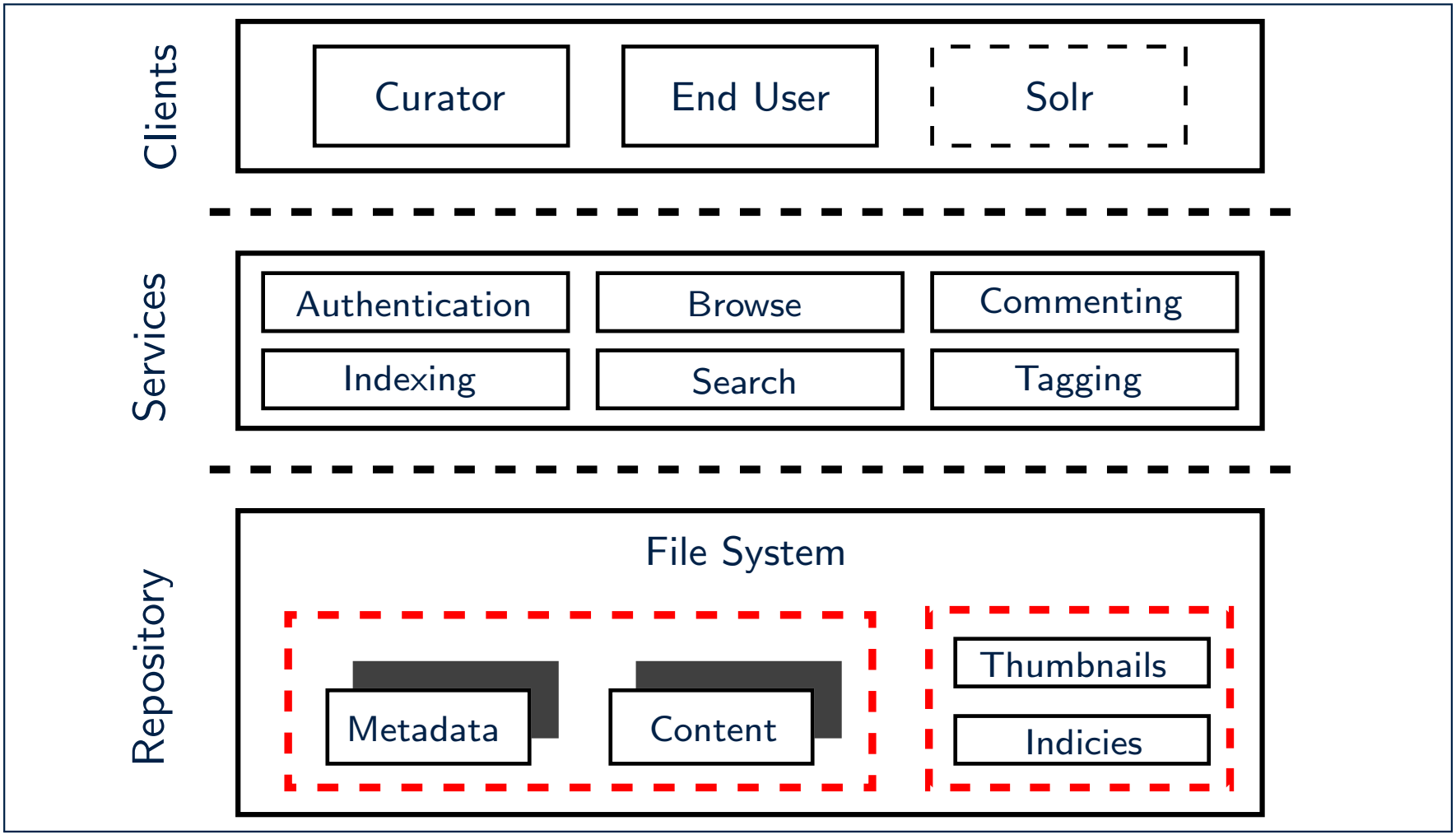
Previous Projects

- Largely driven by user requirements
- Computing resources
  - Minimal bandwidth use
  - Minimal use of computer resources
  - Distributable
- Minimalistic approach
- Lightweight tools and services
- Projects
  - Digital libraries without databases
  - In-browser digital library services
  - Hybrid online-offline collections

Design Principles

- Design for least possible resources
- Flexible design to facilitate extensibility
- Hardware and/or software platform independence
- Heterogeneous object, metadata and service integration
- Minimalist design approach
- Simplified preservation process
- Structured organisation of data
- Support for community and international standards

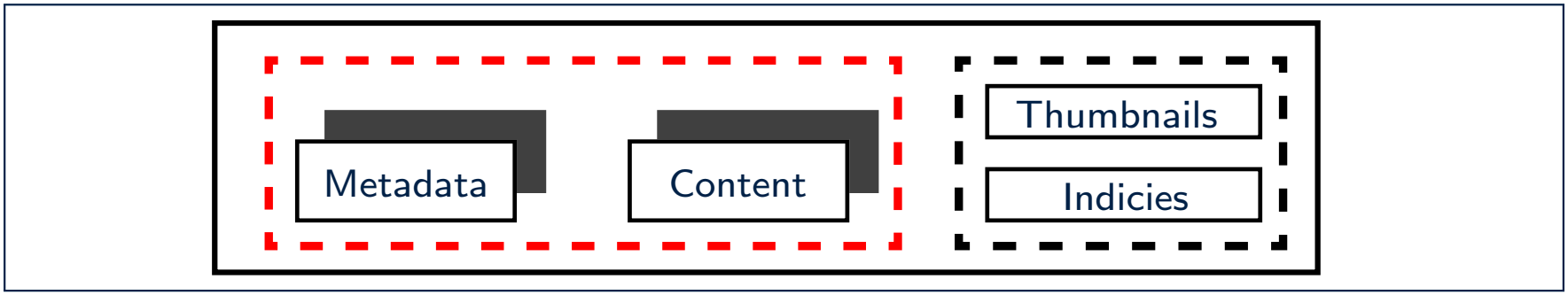
Prototype Implementation - Bonolo Architecture



Tools and Technologies

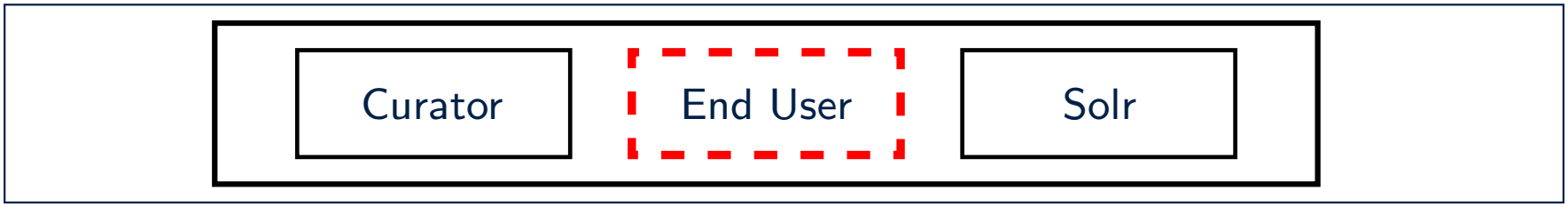
- Core technologies
  - Apache Tomcat
  - CSS
  - HTML
  - Java
  - JavaScript
  - XML
- Additional software components
  - Apache Solr
  - ImageMagick
  - SQLite

Repository Design



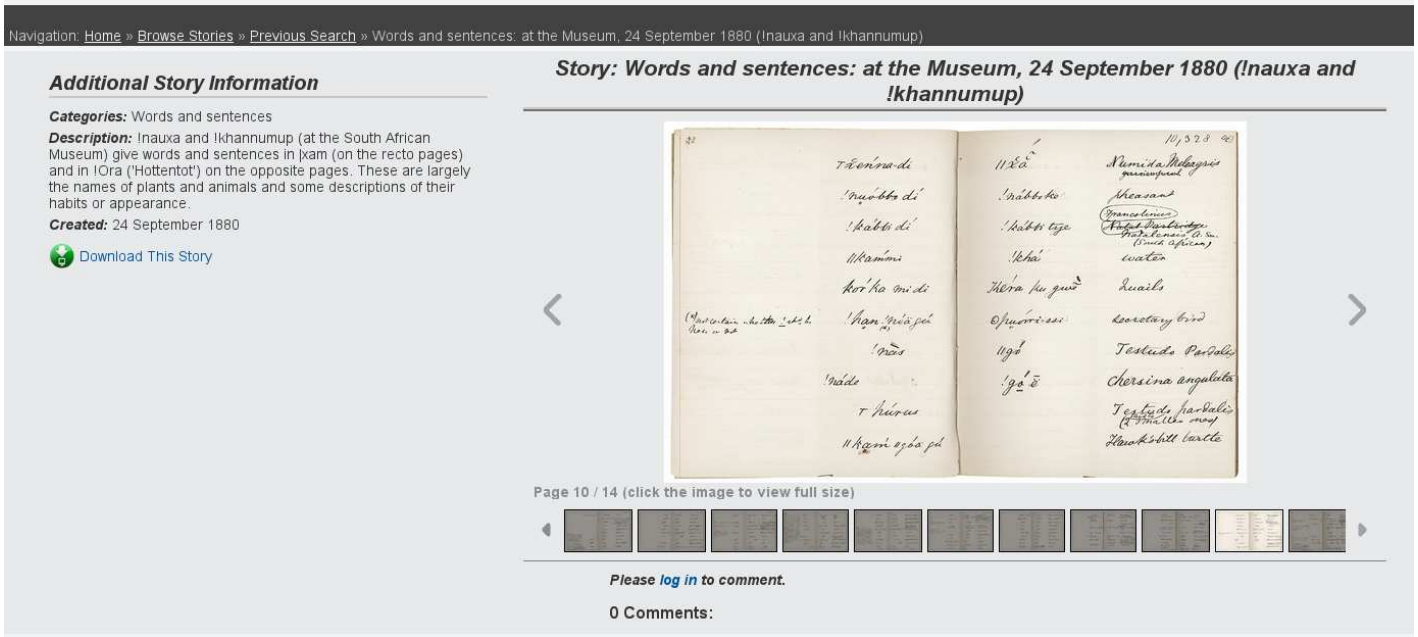
- Hierarchical structure
- Digital content and metadata stored on file system
- Objects stored with corresponding metadata record
- Metadata records encoded using XML

End User Interface - Functionality

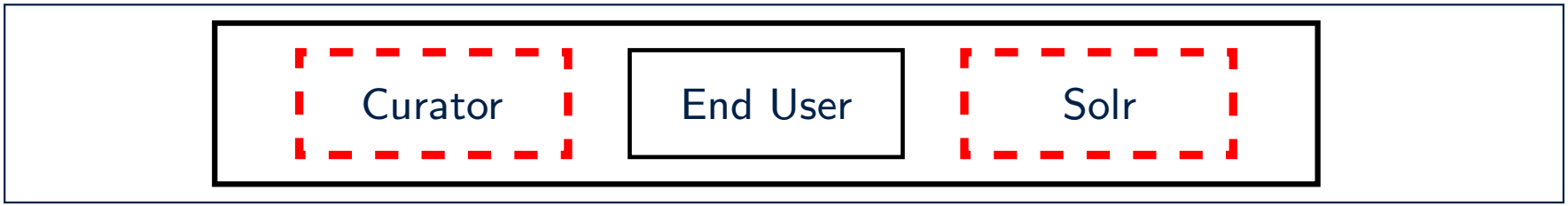


- Authentication
- Browse
- Commenting
- Search
- Tagging

End User Interface - Item View

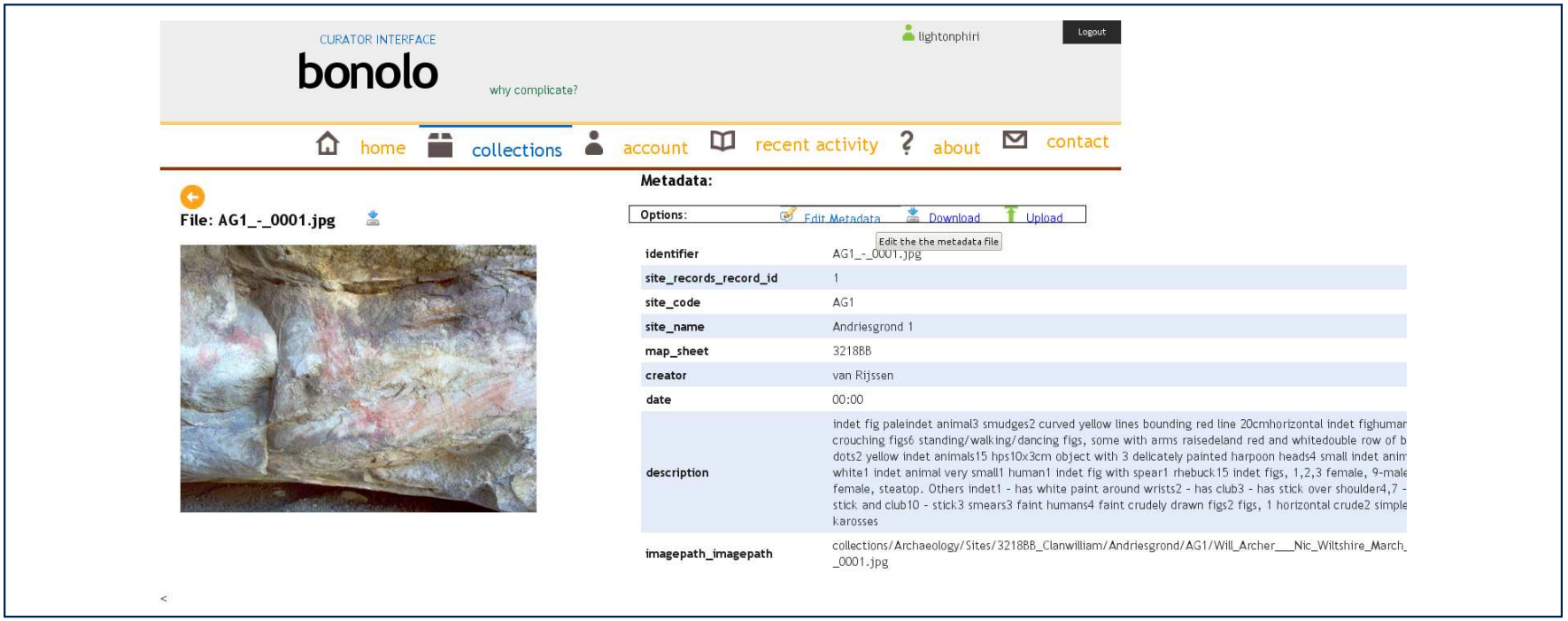


Curator Interface - Functionality



- Authentication
- Batch import and export
- Browse
- Collection and Object management (add, edit, delete)
- Indexing
- Search

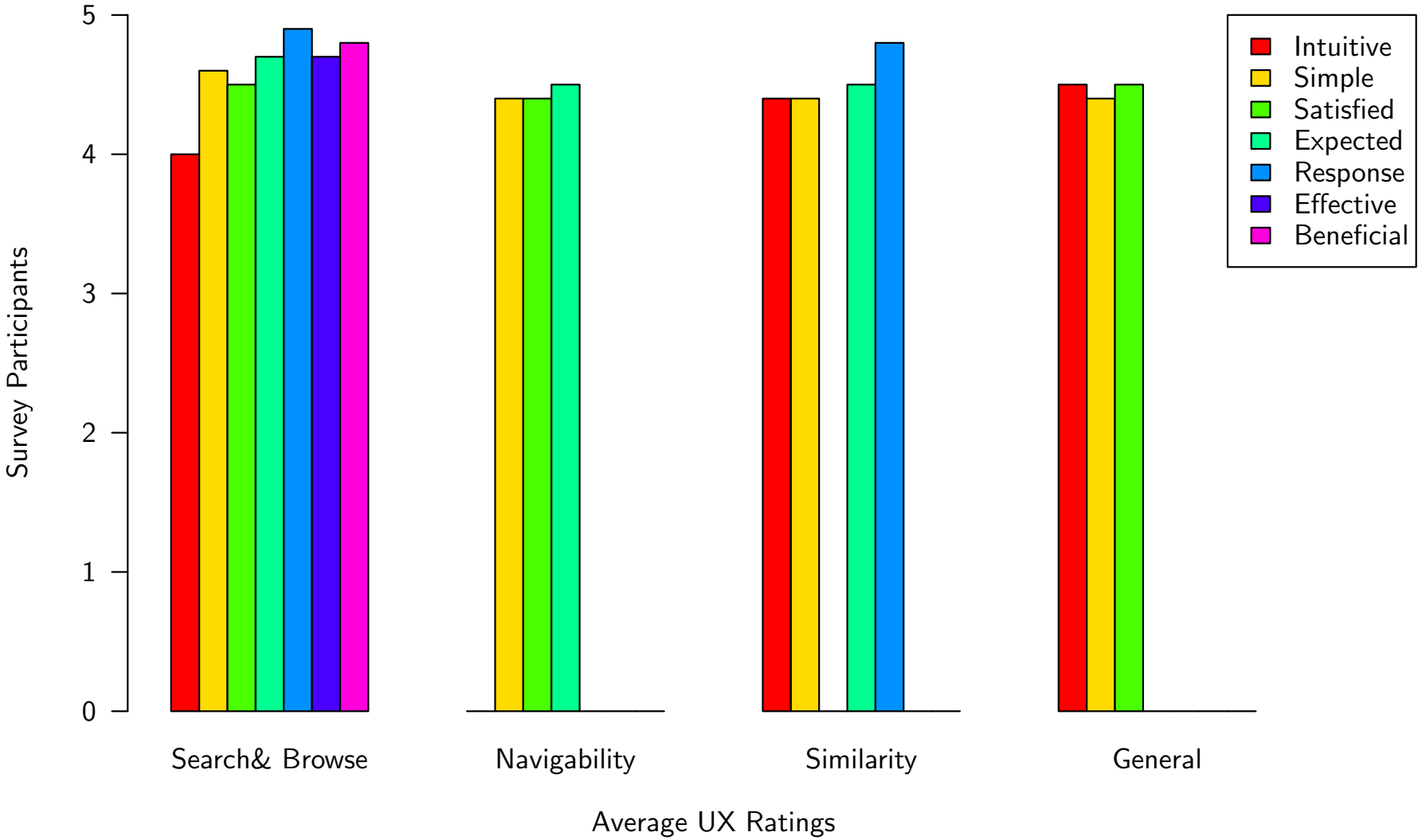
Curator Interface - Item View



End User Interface UX - Lab Experiment

- Objective
  - Feasibility of building user interface using file store
  - Assess impact of file store on user experience
- Target Group
  - Individuals with basic Web experience
  - 17 participants (80% students)
- Approach
  - Informed consent
  - Pre-experiment briefing
  - Typical DL end user tasks (search, browse)
  - Questionnaire

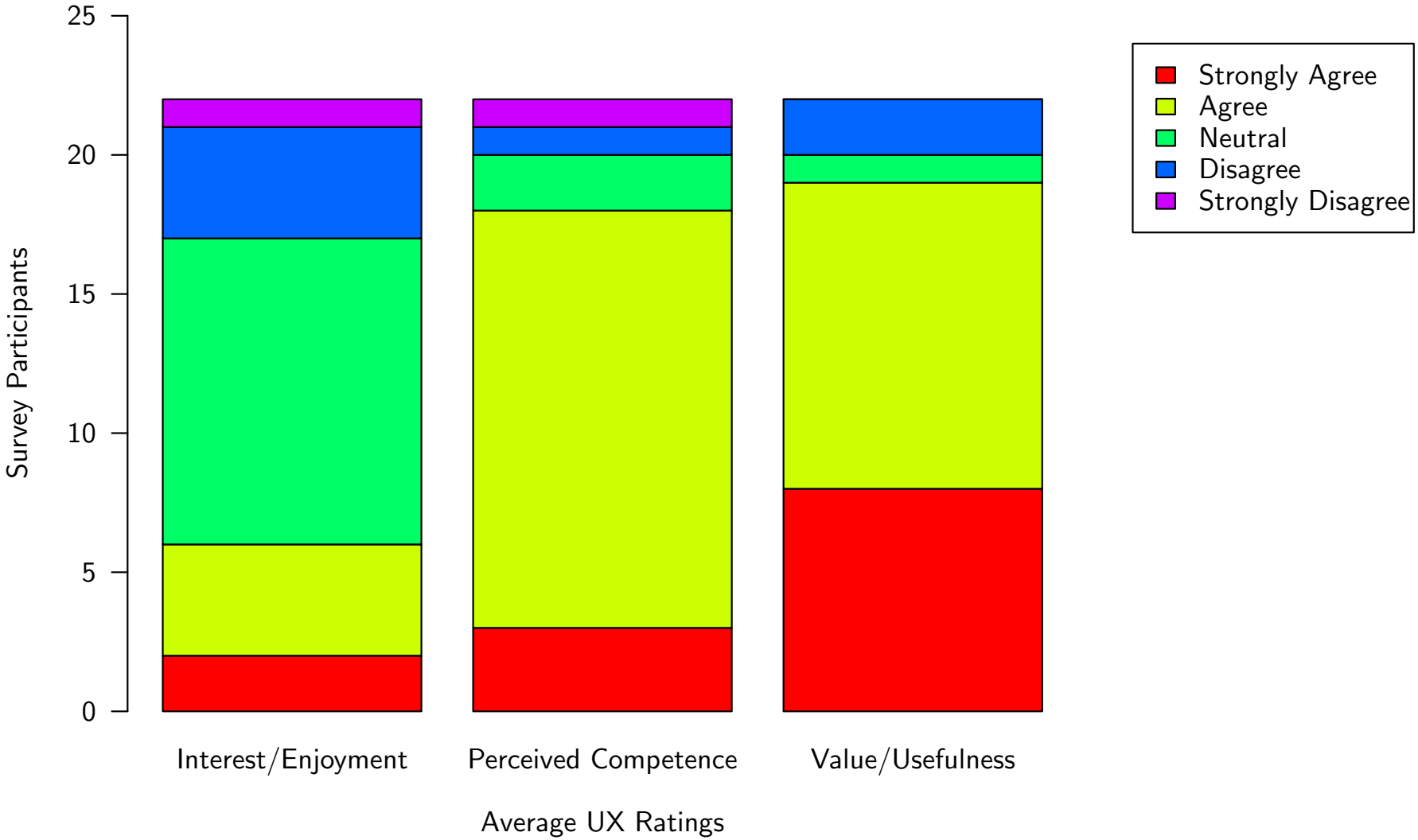
End User Interface UX - Experiment Results



Curator Interface UX - Online Experiment

- Objective
  - Assess user experience when performing curation tasks
- Target Group
  - Individuals with no experience working with DL tools
  - Social networking site recruitment
  - 23 participants
- Approach
  - Intrinsic Motivation Inventory
  - Five (5) minute 'HOWTO' screencast
  - Curation tasks with two datasets
  - Online questionnaire

Curator Interface UX - Experiment Results



Curator Interface UX - Experiment Results

- Participants general comments

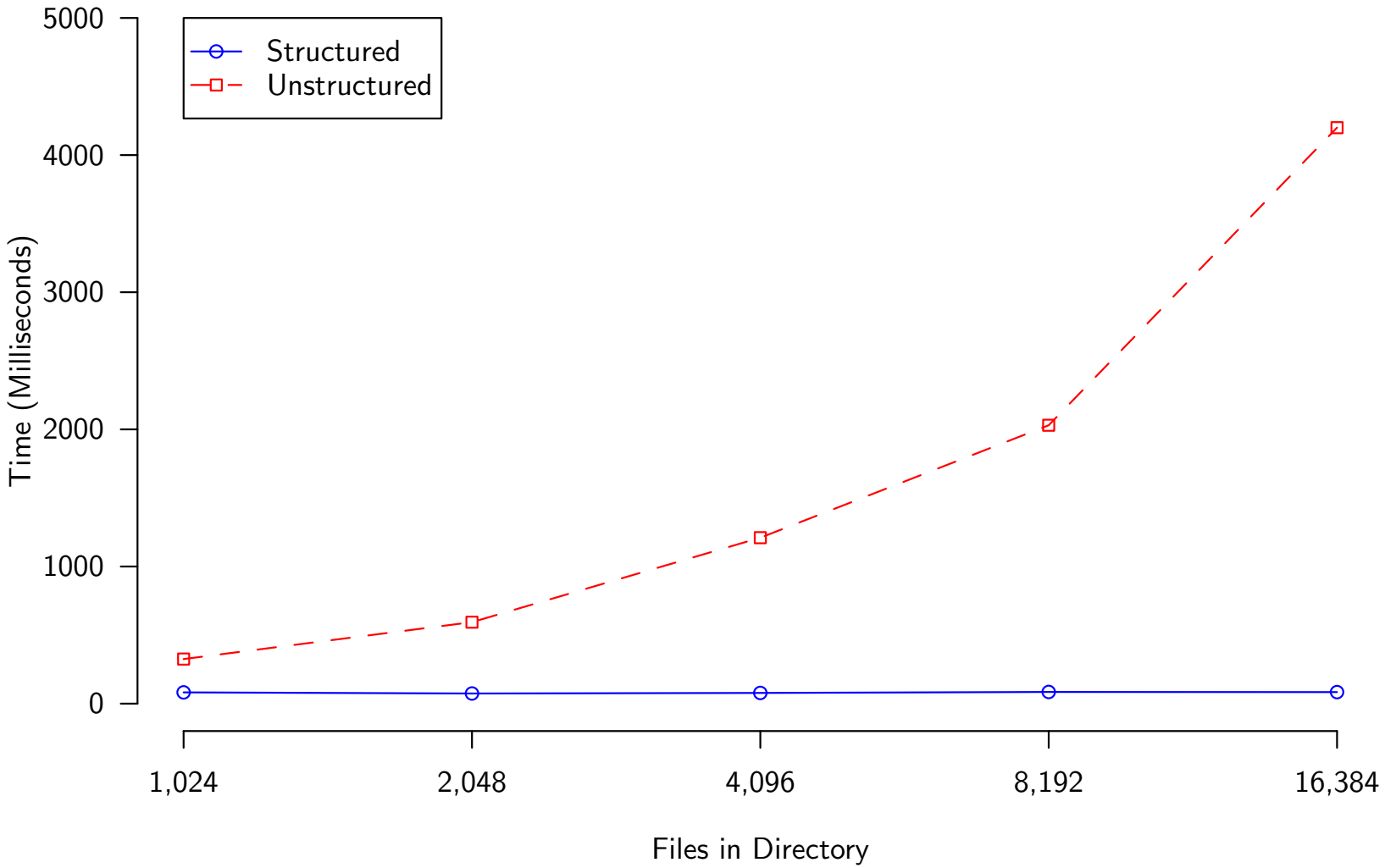
"· · · **Also, I fail to see how Bonolo differentiates itself from something like Dropbox. I can create a folder structure on my PC and upload it to Dropbox very easily. I can then browse my files and folders in Dropbox's web interface.**· · ·"

"· · · **I have to say though that I managed to complete the tasks without watching the video (which is a great sign I think). I'm impatient with manuals but even worse with instructional videos**· · ·"

Repository Performance - Lab Experiment

- Objective
  - Impact of file store structure on performance
  - Performance metrics: response time
- Test Environment
  - Intel Core 2 Duo CPU E7400@ 2.80GHz
  - 2 GB RAM
  - 32-bit Windows 7 Ultimate edition
- Approach
  - Structured and unstructured collections
  - Exponential increase of files in collections
  - Load time and corresponding data transfer during navigation

Repository Performance - Experiment Results



Conclusion

- Experimental results look promising
  - Effectiveness
  - Usability
  - Medium-sized collections
- Work in progress
  - Evaluation
    - Flexibility
    - Scalability
- Future work
  - Reference implementation
    - Design principles
    - Extensibility

Thank You

Questions?

Additional Information

http://dl.cs.uct.ac.za