Reports from Interest and Working Groups

Archivists Interest Group

# Ben Goldman

Digital Preservation Interest Group

Chris Awre

Hydra DPIG

- ## What is Digital Preservation?
  - *"Digital preservation is the active management of digital content over time to ensure ongoing access". (loc.gov)*
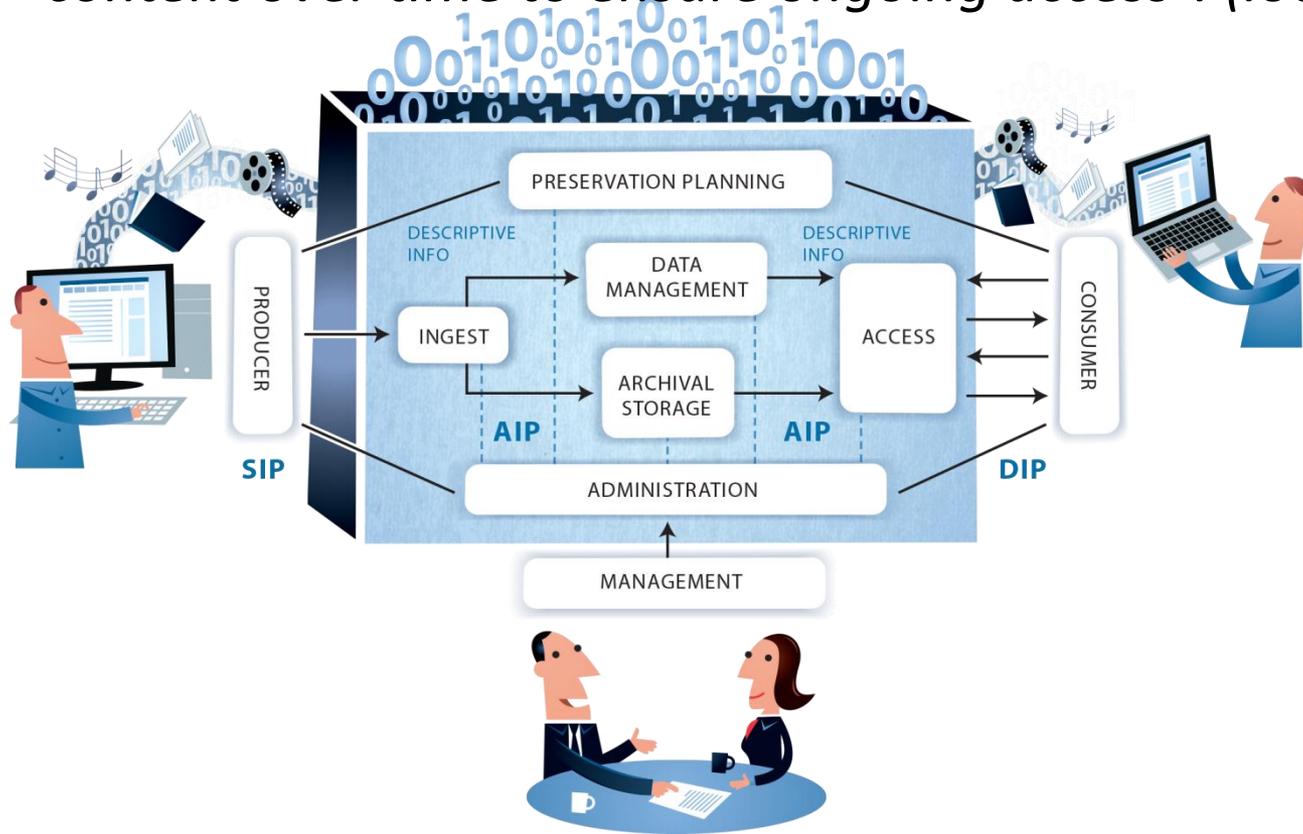
Hydra and Digital Preservation

- Outcome of discussions to date
  - It is **not** important that Hydra core support Digital Preservation
  - It is important that Hydra partners can make use of Hydra components in their Digital Preservation workflows
  - It is important that Hydra can interface smoothly to existing tools and application from Digital Preservation communities such as
    - *Long-term preservation (bit preservation and functional preservation)*
      - *Active management of preserved master preservation objects*
    - *Characterization / validation tools*
    - *Tools for preservation planning*
      - *Risk management*
    - *Migration / Normalization tools*
    - *Emulation tools*
    - *Audit functionality*

Highlights from DPIG Skype Conference (1)

- **December 2014**
  - **'Investing in Opportunity: Policy Practice and Planning for a Sustainable Digital Future'** (DPC/4C)
- **March 2015**
  - Presentation of Chronos, the Hydra based system from the Royal Library, Copenhagen, Denmark.
- **May 2015**
  - Presentation of the HydraDAM2 with focus on the digital preservation actions.

Highlights from DPIG Skype Conference (2)

- [June 2015](#)
  - Presentation of the [Academic Preservation Trust (APTrust)](#) and status of [The Digital Preservation Network (DPN)](#) where APTrust is one of the nodes.

- [August 2015](#)
  - Presentation of the final report from the JISC project ["Exploring Archivematica for research data preservation"](#)

- [September 2015](#)
  - Presentation and discussion of two different strategies for long-term preservation. Use case is e-mail accounts

Looking ahead

- Continue to act as a forum for discussions around digital preservation
  - Monthly calls, 2nd Thursday, 12noon EST
  - Further consideration of the role of Hydra in supporting/enabling digital presrevation
  - Sharing experiences and ongoing projects

Geospatial Interest Group

Darren Hardy

What are we doing?

- We are an interest group for managing geospatial data and metadata
- We started in Feb 2015 and have specific questions in our charter
  - What set of practices, standards, tools, and infrastructure do we want for geospatial objects, and non-geospatial objects with geo characteristics?
  - How can institutions move from simply providing access to GIS data to sustainable full-lifecycle repository based management of GIS data?
  - How can we create consensus and collaboration in answering these questions?
- We have a working group working with PCDM
- https://wiki.duraspace.org/display/hydra/Hydra+Geospatial+Interest+Group

Come to **Geo4LibCamp** in January 2016

- Where: Stanford University
- When: January 25-27, 2016
- Who: GIS developers, librarians, and managers
- What: hands-on meeting to bring together those building repository and associated services for geospatial data
- https://wiki.duraspace.org/display/hydra/Geo4LibCamp+2016

Hydra GIS Data Modeling Working Group

Darren Hardy

Hydra GIS Data Modeling Working Group

- ## Developing models for geospatial data in Hydra
  - Provide common GIS data models for shared tools to enhance interoperability among institutions
  - Drive shared approaches to metadata, workflows, and best-practices for managing GIS content
  - Provide a starting point, a pathway, for newcomers who want to manage GIS data in Hydra
  - Promote stewardship of GIS resources in a Hydra repository

- ## See our poster on Georeferenced Maps using PCDM and HydraWorks!

georectification

vectorization
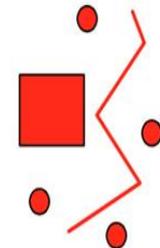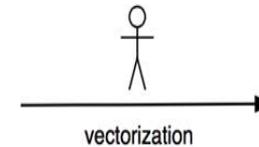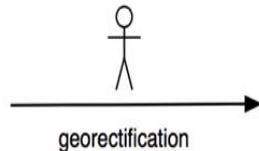
Metadata Working Group

Karen Estlund

Organizing for Action: Subgroups

- 38+ participants!

- Technical Subgroup - Aaron Coburn & Nick Ruest, leads
- Rights Subgroup - Esmé Cowles, lead
- Applied Linked Data Subgroup - Steven Anderson, lead
- Descriptive Subgroup- Carolyn Hansen, lead
  - MODS subsubgroup - Steven Anderson, lead
- Structural Subgroup - Julie Hardesty, lead

Metadata WG formed in April 2015.

Technical Metadata Base Metadata Application Profile

- Core file metadata from EBUCore:
- Checksums from PREMIS
- Format identification
    - PCDM File Type Vocab
- Additional Descriptors (such as rdfs:label)

Technical SubGroup: including technical metadata file (such as FITS)

```
<> a pcdm:Object ;
    pcdm:hasFile </some/source/file> ;
    pcdm:hasRelatedFile </some/fits/xml/output> .
</some/source/file> a pcdm:File .
</some/fits/xml/output> a pcdm:File ;
    iana:describes </some/source/file> .
```

Technical SubGroup: Tool to Use w/Hydra Recommendations

File Identification
- file http://unixhelp.ed.ac.uk/CGI/man-cgi?file
- FIDO http://openpreservation.org/technology/products/fido/
- siegfried http://www.itforarchivists.com/siegfried
- DROID https://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/digital-continuity/file-profiling-tool-droid/

Characterization
- exiftool images http://www.sno.phy.queensu.ca/~phil/exiftool/
- mediainfo audio/video https://mediaarea.net/en/MediaInfo
- ffprobe (part of ffmpeg) audio/video https://www.ffmpeg.org/ffprobe.html
- fiwalk (part of sleuthkit) disk images http://www.forensicswiki.org/wiki/Fiwalk
- TIKA https://tika.apache.org/

Rights Metadata SubGroup - Minimal Recommendation

| Property	Description	Values
edm:rights	Rights statement	URIs from Creative Commons, RightsStatements.org, other other vocabularies, minting new URIs as needed
dcterms:rightsHolder	Rights holder	URIs from common vocabularies

Rights Metadata - Extended

Property	Description	Values
dc:rights	Rights description	Free text description of rights, such as administrative contact information, a description of the specific license terms, or other rights-related restrictions.
marcrel:cpc	Copyright claimant (rights holder before rights expired)	URIs from common vocabularies (e.g., LCNAF, VIAF, etc.)
premis:hasCopyrightStatus	Copyright status	URIs from PREMIS CopyrightStatus
premis:hasCopyrightJurisdiction	Copyright jurisdiction	Country/jurisdiction codes from common vocabularies (e.g., ISO 3166)
pcdmrts:rightsOverride	Rights statement override (temporary rights statement that overrides edm:rights, such as a temporary embargo or license)	URIs from Creative Commons, RightsStatements.org, other other vocabularies, minting new URIs as needed
pcdmrts:rightsOverrideExpiration	Rights statement override expiration (date when status reverts to base edm:rights statement)	Date in a standardized format (e.g., W3C DTF profile of ISO 8601)

Applied Linked Data - Linked Data Fragments

http://j.mp/hy-ldf
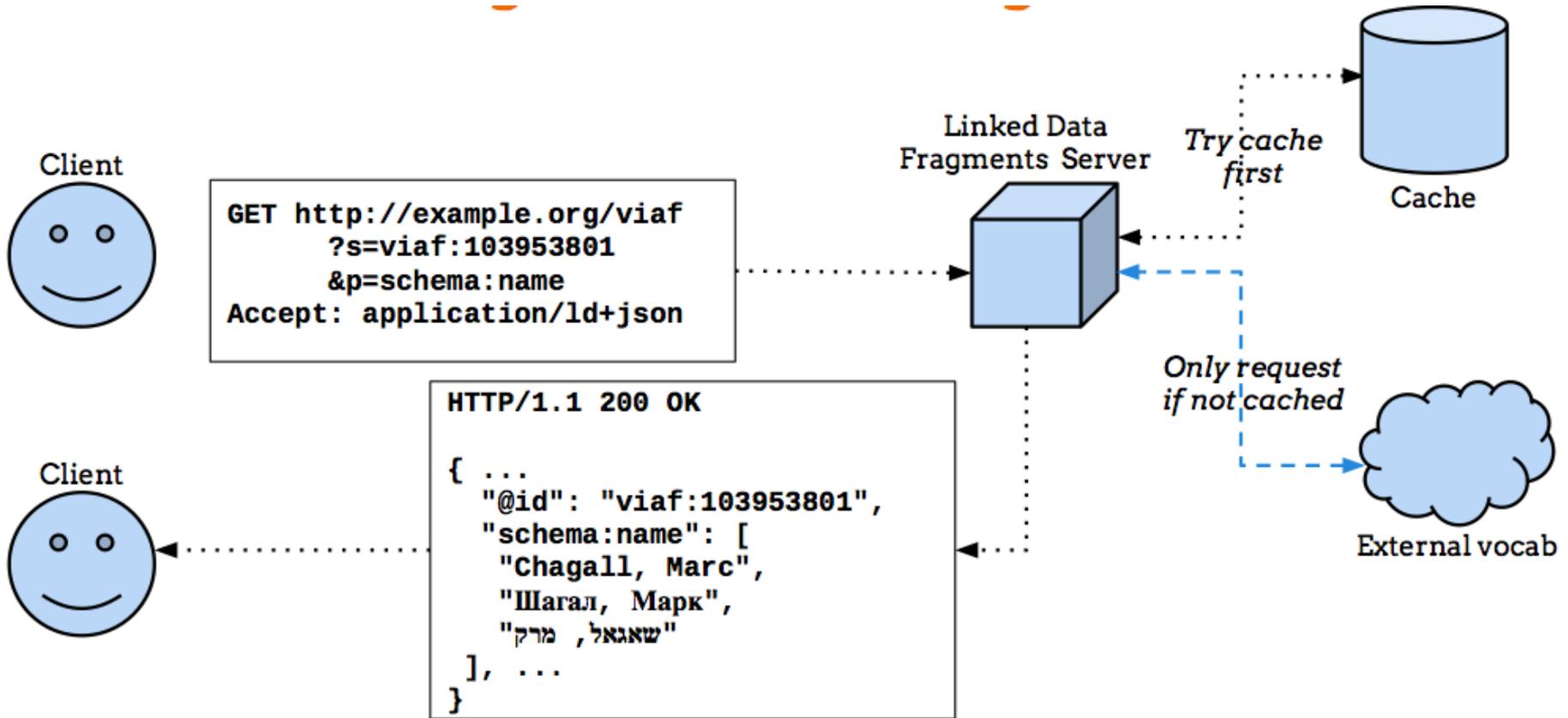
Descriptive Metadata SubGroup - Collect User Stories
What Encoding Standards Are you Using?
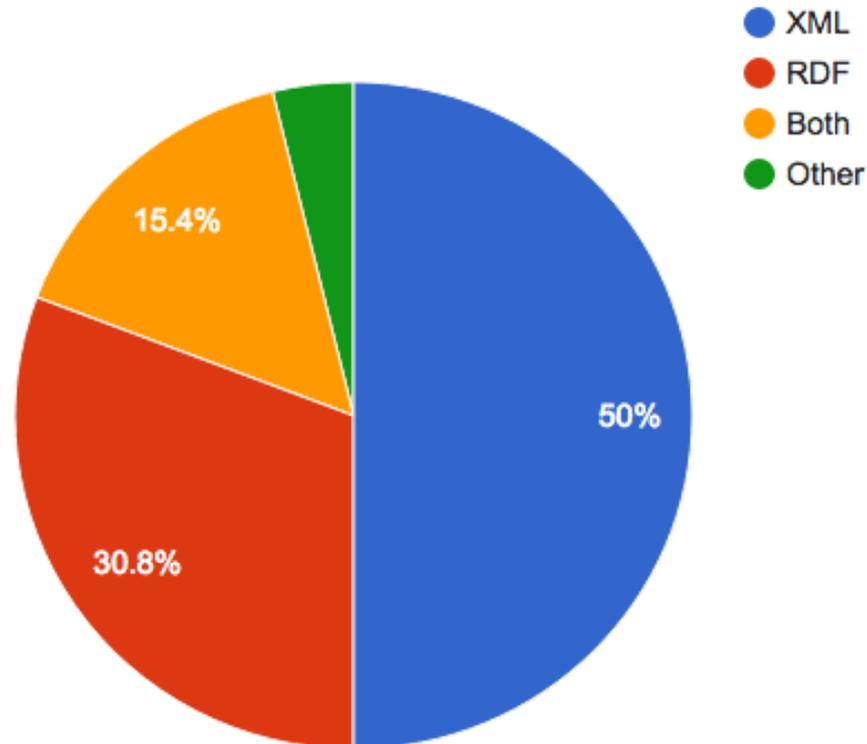
Descriptive subsubgroup - MODS

Who did what?

https://goo.gl/i1XiDq

On Deck

Structural SubGroup & Time Based Media

Join us!

Meetings

Time: Wednesday at 10 am (PDT). Every two weeks. Call duration: 30 min. **Next call October 7, 2015.**

Call-In Info: 1-530-881-1400, access code 651025

Metadata Requests & Priorities

https://wiki.duraspace.org/display/hydra/Metadata+WG+Requests+and+Priorities

Pause pressed
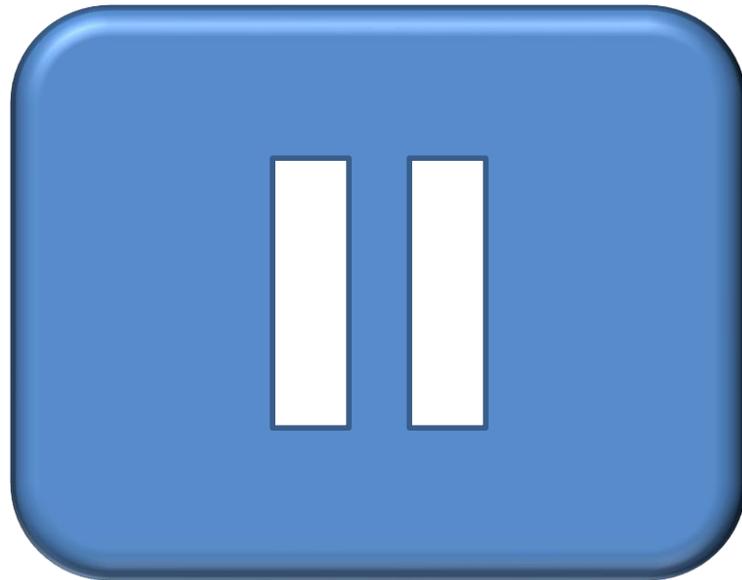
Page Turner Interest/Working Group

Richard Green

Page Turner Interest Group

- The Group was established to look at recommending or developing a common Hydra approach to delivering 'page turning' of digital books and manuscripts in Hydra repositories.

- It was clear before assembling the group that the International Image Interoperability Framework (IIIF) was the obvious foundation for this work.

PTIG research

- The group acted as a gathering place for information about IIIF, assessed different image viewing solutions and investigated how the two could be implemented in a Hydra application.

- The group also pulled together the following:

  - A set of common requirements.
  - Examples of local practices.

PTIG report

- The full report of the Group's work, and its recommendations, can be found on its wiki page:

  https://wiki.duraspace.org/display/hydra/Hydra+Page+Turner+Interest+Group

- Thanks to the 15 or so people who contributed along the way and, especially, Julie Allinson from York who wrote the summary report for us.

Service Management Interest Group

Ellen Ramsey

User Experience Interest Group

Patricia Hwse

Web Presence Interest Group

Tom Cramer

HWPWG

Hydra Web Presence Working Group

Hydra's website & wiki have grown organically. What should go where?

How can essential information be presented in a concise and professional manner?

HWPWG Charter

- A list of primary users of the website and wiki

- A proposal for an overarching information architecture for the wiki and website

- Strategies for generating & maintaining content

HWPWG Charter

- List primary users of the website and wiki

- Propose an overarching information architecture for the wiki and website

- Generate strategies for generating & maintaining content

Benchmarking

	PrHy	HyWk	Aper	Fed	DSp	Apa	Isl	Fig	BPr	Arc	Opc
Responsive design	❌	❌	❌	✅	✅	❌	✅	✅	✅	✅	✅
Multilingual	❌		Few pages in Spanish and Chinese			❌	❌	❌	❌	✅❓ (sort of)	❌
Search tool	✅	✅	✅	✅	✅	✅	✅	✅❓ (repo search not site search)	✅	❌	✅
News channel / blog	✅		✅	✅	✅	✅	✅	✅	✅	✅	✅
Philosophy behind our product	✅		✅	✅	❌	✅	❌❓ (not really)	❌❓ (not really)	✅	❌❓ (not really) ✅ (buried on wiki)	❌❓ (not really)

Primary Audiences

Developer (new)	Web
Developer (experienced)	Wiki
Functional / content specialists	Web
Manager	Web
Partner	Wiki
Senior leader / administrator	Web

Next Steps

- Redefine the information architecture

- Tidy up the wiki

- Redo the website

Poster session this afternoon

- The session is in two parts; refreshments will be available half way through.

- If you are presenting a poster alone, please be there for one half and look around in the other.

- If there are two or more of you, please try to split your time so that someone is there all afternoon.

And now…

- Please will you spare us 10 minutes to take a group photo?

- Lunch!!!

- If you have a poster, come along at 2.15pm to set up.

- Else, please be back prompt at 2.30pm because the afternoon session kicks off with some introductions to Connect newcomers