Making metadata linked data

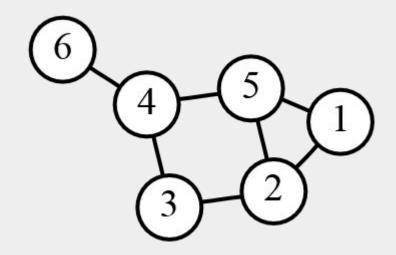
The PLODAB experience

Laura J. Smart

Digital Initiatives Symposium Panel:

Manipulation, Multiplicity, and Metadata:

Piloting Linked Open Data with Artists' Books
2016.04.28



What we did, what we learned

"If you can't be a good example, then you'll just have to be a horrible warning"

Catherine Aird





Month 1-3

Planning process

While we

- selected works
- digitized
- reviewed ILS metadata

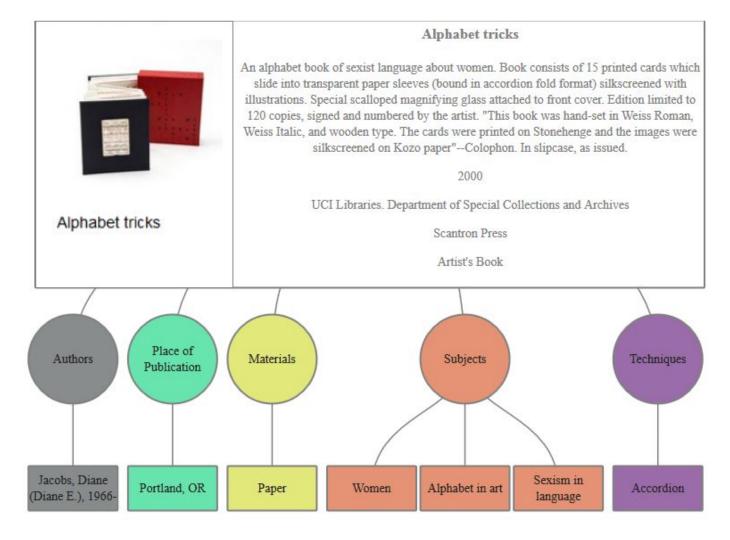
We also

- considered personas
- wrote use cases
- derived reqs
- evaluated systems

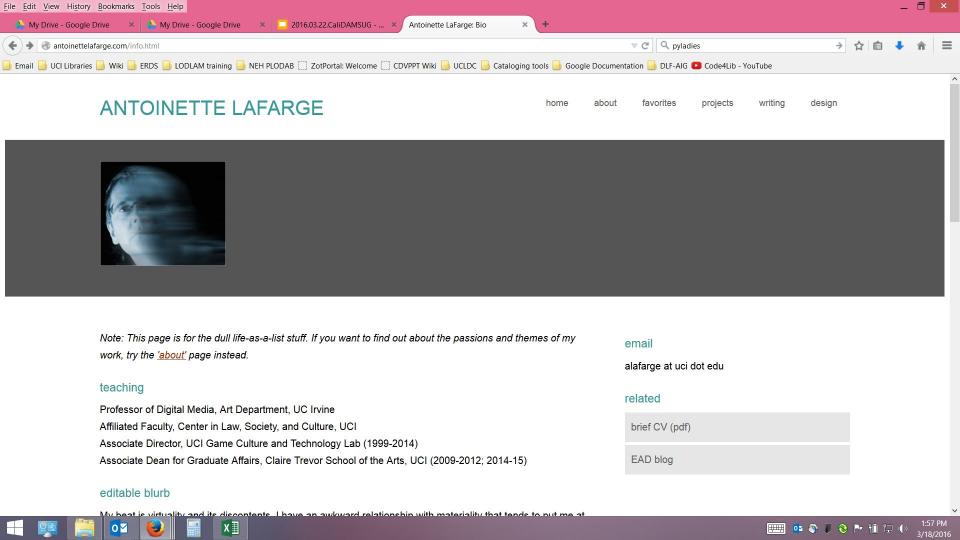
Record: Prev Next Author Jacobs, Diane (Diane E.), 1966-, artist More Resources: Title Alphabet tricks / by Diane Jacobs **Google Books** Find Exact Title in Melvyl **Published** [Portland, Oregon]: Scantron Press, 2000 Location Call # Status Langson Library Special Collections Gen N7433.4.J336 A43 2000 LIBRARY USE ONLY Send via Text Message Send via Email Permanent link to this http://antpac.lib.uci.edu:80/record=b3191751~S7 ANTPAC record Description 15 unnumbered leaves: illustrations; 11 cm Mobile: **Content Type** text still image Medium unmediated Format volume Subject - LCSH Women Alphabet in art Sexism in language Genre/Form Artists' books -- Oregon -- 20th century. Alphabet books -- 20th century. Accordion fold format (Binding) -- 20th century.

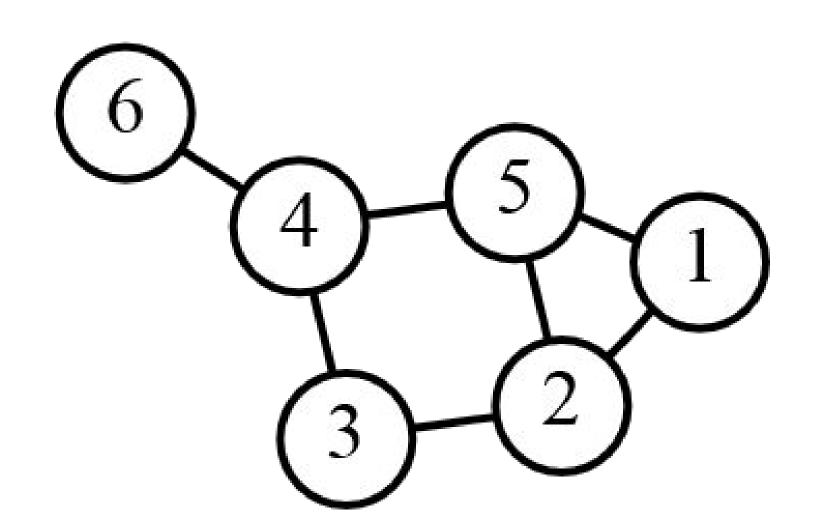
Other author

Scantron Press, publisher Women's Studio Workshop

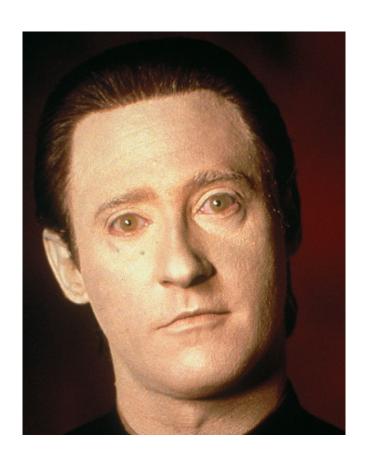


Input \rightarrow BLACK BOX \rightarrow Output





I never metadata I didn't like

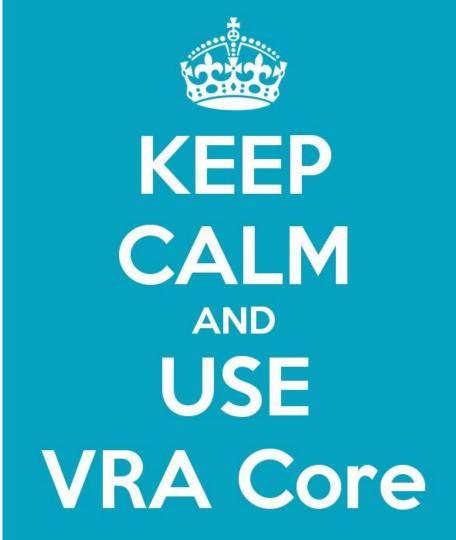


Picking descriptive metadata

- Context of use
- Format of content
- Descriptive richness

Options under consideration

- AACR2/RDA in MARCXML
- MODS
- CCO/VRA



MarcEdit -- Your complete free MARC editing utility



MarcEdit -- Your complete free MARC editing utility





MarcEdit -- Your complete free MARC editing utility







Must store URIs Must be customizable Must support complex objects

Must have bulk ingest Functional requirements specification

Must export metadata as Linked Data

Must support visualization

Must use DC, VRA

We reviewed

Omeka

Scalar

Collective Access

Shared Shelf

Nuxeo/CDL

We went with

Nuxeo/CDL

We were satisfied it met our need

We didn't just go with it because it was what we had

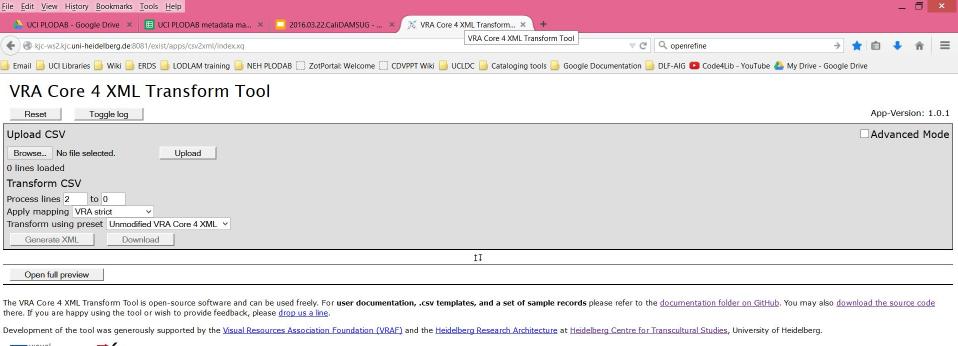
Month 4-8

Prepping and testing

- metadata mapping
- MARC -> tab delimited
- DAMS test ingest
- start hiring process
- test transform tools
 - XSLT
 - o csv2VRA

-							
	A	F	G	Н	1	J	К
1	Nuxeo metadata scheme			Nuxeo import QA review notes	MARC21 analog element for bulk importing / default data v alue notes	VRA Core 4.0 analog element	VRA Core 4.0 XML express
2	Basic Information		(2015-12)	11 11 11 11 11 11 11 11 11 11 11 11 11	unary cromont		
3	Field Label	Repeatable	Required?				
4	Title	No	No		245 \$a \$b omit \$c	Title	titleSet/display
5	Alternative Title	Yes	Nb		246		
6	Identifier	No	No		856 \$u - ADDED 02/02/16		
7	Local Identifier	Yes	No		946	Textref	textrefSet/display
	Туре	No	No		Globally supply "physical object"		

Reference open source, powerful tool for working with messy data



there. If you are happy using the tool or wish to provide feedback, please drop us a line.



























Month 9-10

Finalize metadata & prototype

- enrichment Getty vocabs
- hand coding VRA-XML
- creating RDF-XML
- hire & trainprogrammer
- prototyping visualizations

Month 11-12

Mad dash to the finish

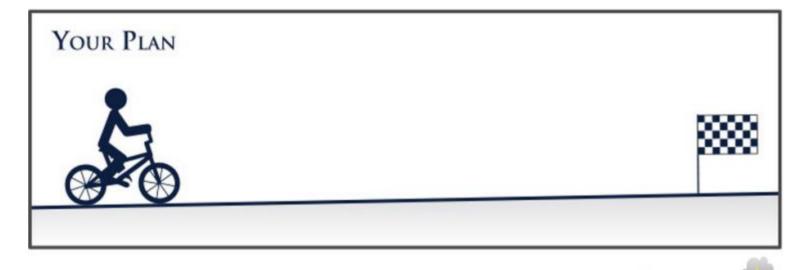
- final NUXEO load
- connect viz to DAMS
- UCI look & feel
- Ux review
- write, present
- plan phase 2

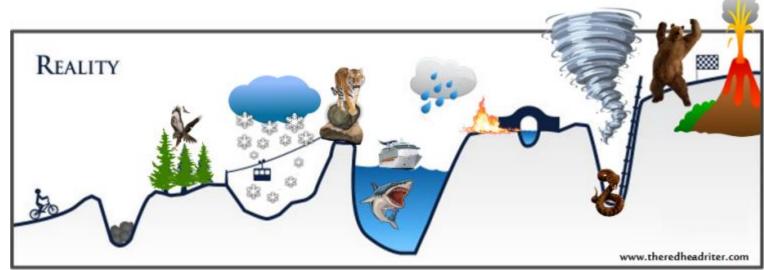
Process broken down

- Community audit & development
- Use cases & requirements
- Collection development, i.e. choose the stuff
- DAMS decisions
- Metadata standards decisions
- Output decisions
- Standardize metadata inputs and asset formats
- Admin interface design
- Ingest content and descriptions
- Linked data/API output
- Basic user interface design
- Advanced user interface design Enhancements

Lessons

- Break down work plan in more detail during grant proposal phase
- Team members should understand general systems dev life cycle:
 - articulate the end product: use cases & functional requirements are your friends
- Heterogeneity is here to stay
- Metadata maxims
 - analyse, clean, enhance FIRST
 - version control
 - estimate time then triple
 - "easy automated tools" may take as long as making records from scratch
- Centralized service priorities ≠ your priorities







NATIONAL ENDOWMENT FOR THE

Humanities