UC San Diego
Presentations and Posters

Title
Sharing Metadata: Building Collections and Communities

Permalink
https://escholarship.org/uc/item/5qf9h2k1

Authors
Chandler, Robin L.
Westbrook, Bradley D.
Rundblad, Kevin

Publication Date
2009-05-09

eScholarship.org Powered by the California Digital Library
University of California

Sharing Metadata: Building Collections and Communities

- Robin Chandler, UC San Diego Libraries
- Brad Westbook, UC San Diego Libraries
- Kevin Rundblad, UCLA

- Presented at Annual Meeting of the Society of California Archivists, Riverside, CA, May 9, 2009

Libraries, archives, and museums cannot afford to think about collections only in the context of the local community. Sharing benefits our users, and we can no longer assume our users will come in through the "front door." Cultural institutions must get their content "into the flow“ where working users will discover and reap the benefits of digital content delivery. The speakers in this session will describe the benefits of creating repurposable metadata discoverable in multiple environments; discuss some of the obstacles challenging institutions to create sharable metadata; compare two mechanisms for creating repurposable metadata; and share some success stories where leveraging metadata in social networking environments helped users discover the breadth and depth of an institution's collections.

Sharing Metadata:
Building Collections, Community & Computing Power

Robin L. Chandler
Society of California Archivists
May 9, 2009

- **Major Theme**:
  - Metadata Standards are the key to Sharing Metadata
- Repositories amass collection metadata in local databases *locked* away from the internet public in **Data Silos**
- Standardized metadata is flexible, shareable and promotes interoperability
- **Minor Theme** :
  - Tear Down your Data Silos & Share Your Metadata

“Shareable metadata is metadata which can be understood and used outside of its local environment by aggregators to provide more advanced services”

Sara Shreeves *Moving Towards Shareable Metadata*

What’s Wrong with Data Silos?

Data SILOS limit exposure of our collections to users…..

Data SILOS Freeze Innovation

- Our users are innovators

- **"Innovation comes only from readily & seamlessly sharing information rather than from hoarding it"** Tom Peters, author *In Search of Excellence*

Why Share Our Metadata?

- **Sharing Benefits our users**
  - Data must be where users are working -- getting “Into the Flow”
  - Federates distributed collections
  - Supports “one-stop” searching
  - Supports Web 2.0 culture and services
- **Sharing Benefits our institutions**
  - Increases exposure to collections
  - Broadens user base
  - Increases potential for collaboration & grant seeking opportunities
  - Enhances rather than limits your data

A History of Sharing:

- Knowledge
- Systems
- Standards
- Expertise

1877 – Standardization of size of catalog cards

1901 – LC catalog card distribution program

1968 – MARC

1971 – OCLC

1980 – RLIN

1987 – Z39.50

1994 – DCMI (Dublin Core Metadata Initiative)

1997 – MARC 21

1998 – EAD 1.0

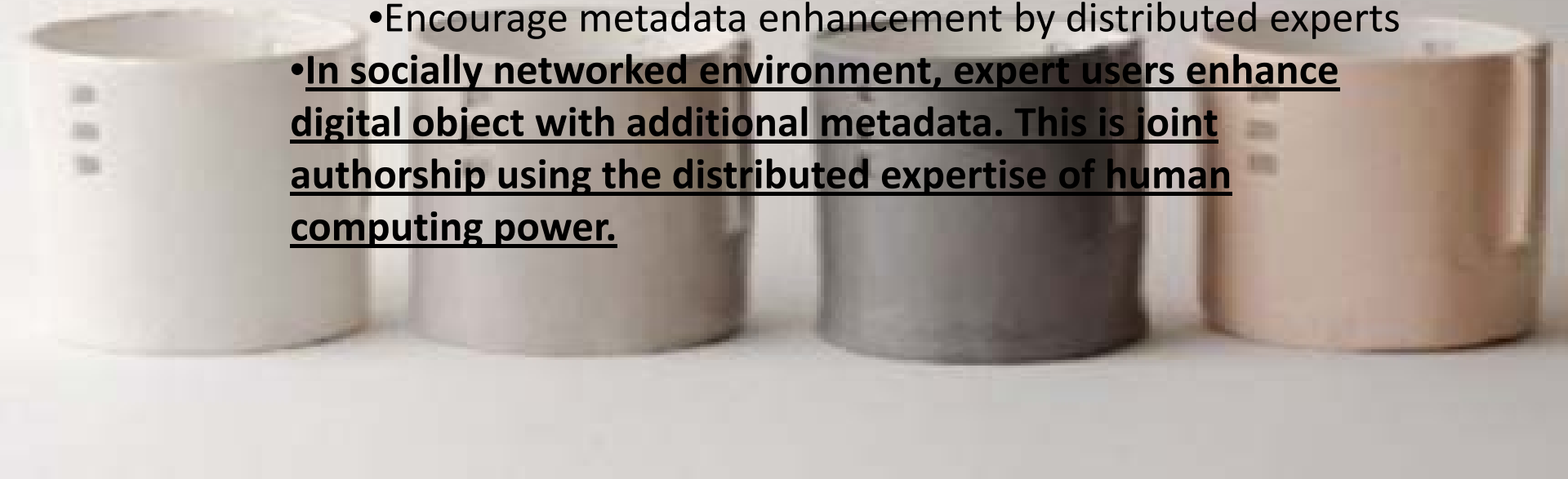
2001 – OAI-PMH ; METS ; MODS

2004 - Web 2.0

2007 – OAI-ORE

Evolution of Sharing: Analog to Web 2.0

- **Analog: Sharing the Creation of Authoritative Catalog Records**
  - Librarian expert creates a catalog record and shares the record with other Librarian experts
  - Motives:
    - Encourages shared cataloging & conserving staff resources
    - Encourages cooperative collection development
- **Web 2.0: Sharing the Production of Knowledge**
  - Archivist expert creates an interoperable digital object with standardized metadata and shares the object with users
  - Motives:
    - Increase exposure of collections
    - Encourage metadata enhancement by distributed experts
  - **In socially networked environment, expert users enhance digital object with additional metadata. This is joint authorship using the distributed expertise of human computing power.**



Key attributes to shareable metadata:

- Human understandable outside of its local context
- Provides contextual information at the object level
- Consistent across a collection: uses same data fields and same controlled vocabulary across a collection
- Easily packaged in multiple views for multiple audiences
- Machines can process
- Uses Standards: descriptive, technical, administrative

What is a standard?

"a description of a measurable item or process" which is agreed upon and published by recognizable standards-issuing organization such as national and international associations."

"Standards are used to assure that communication is clear. The use of a recognized industry standard assures everyone that processes and measurements are uniform."

Senkus, *Encyclopedia of Library and Information Science,* 2003, p. 2,737.

“Clear thinking about what we are doing with description is essential to creating useful tools and delivering them in a helpful and understandable format.”

--Michael Fox
2000 SAA Presentation

Types of Standards

Type	Definition	Example
Content	How descriptive data is expressed	DACS
Value	How descriptive content is controlled	LCNAF, LCSH, AAT
Structure	Specification of the elements of description, technical and administrative metadata and their relationship to one another	MARC, MODS, DC, VRA, MIX, PREMIS etc.
Communication	How description is formatted and transmitted	MARC, EAD, METS

A repository applying metadata standards will create effective metadata records for the

****discovery,***

****sharing,***

and,

****ongoing management***

of its digital objects.

Data Silos Can Happen Because:

- We lack staff & skills to implement standards
- We perceive our Users as Local Only
- We recognize that storage is cheap
- We believe we can't keep up with dynamic standards

Data SILOS are unstable

- Data standards are stable….they are the products of thorough intellectual process and are maintained by groups of experts.
- Data standards are also dynamic – evolving in response to community needs.
- Institutions customizing standards face migration issues
- Leads to local instability as institutions must resolve problems without community solutions

ALA Tiaga Forum Provocative Statements Discussion, 2006

“Within the next five years…….”

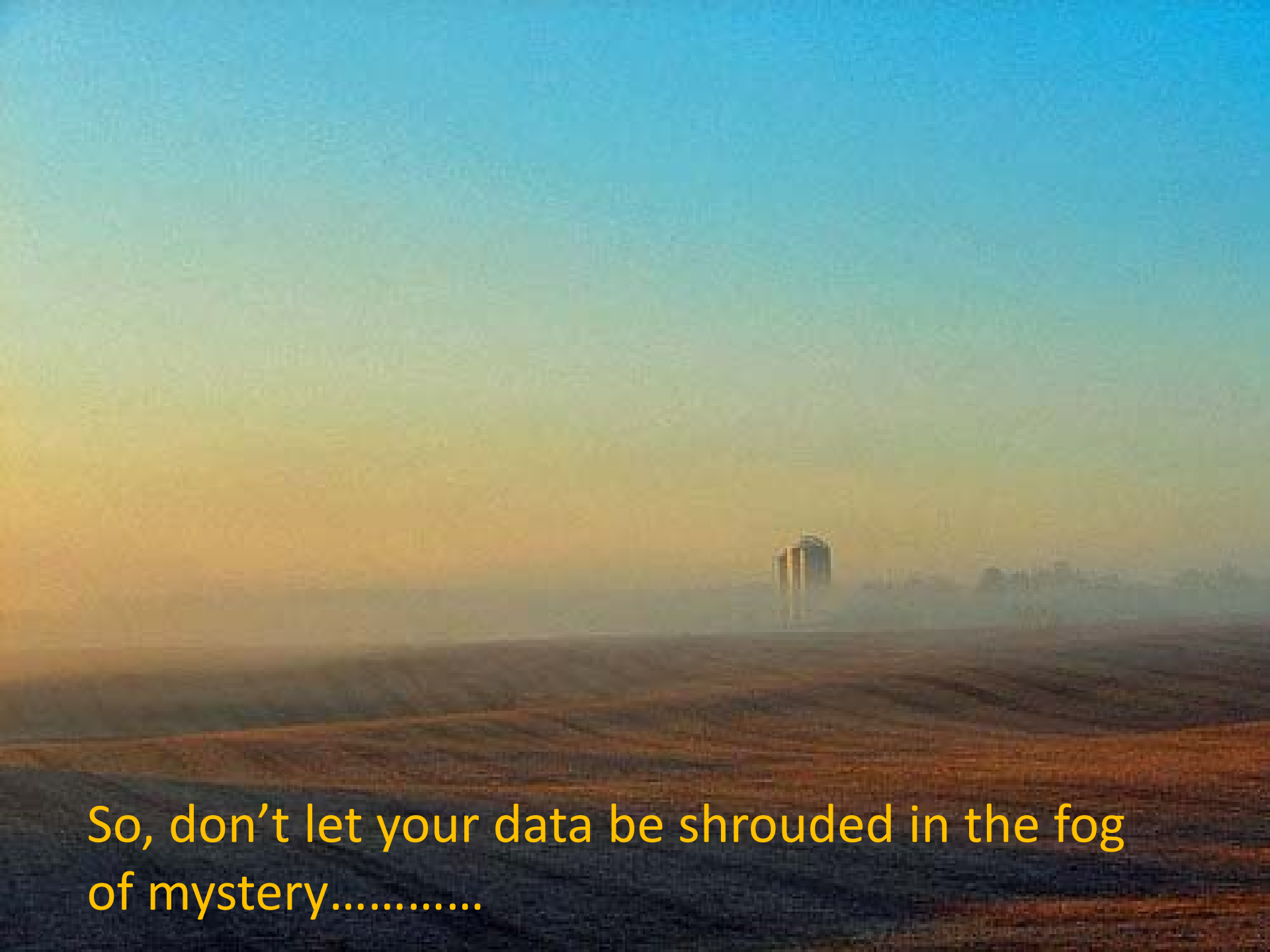
•The continuing disaggregation of content from original containers causes a revolution in resource discovery

•There will no longer be a monolithic “library” Web site….instead, data will be pushed out to many starting places on the Web and directly to users

•All information discovery will begin at Google, including discovery of “archival“ resources

Ithaka's 2006 Studies of Key Stakeholders in the Digital Transformation in Higher Education, 2008

- Since 2003, the number of scholars across disciplines who report starting their research at non-library discovery tools, has increased
- Over 90% of librarians list the gateway /portal role as very important
- There is a disconnect!


So, don’t let your data be shrouded in the fog of mystery……

Tear down SILOS!
Share your
Metadata!
Promote Discovery &
Use of your Data!

Library of Congress Working Group on the Future of Bibliographic Control, 2008

- Recommends making EAD finding aids accessible via online catalogs and internet
- Encourages inter-institutional collaboration for sharing metadata records

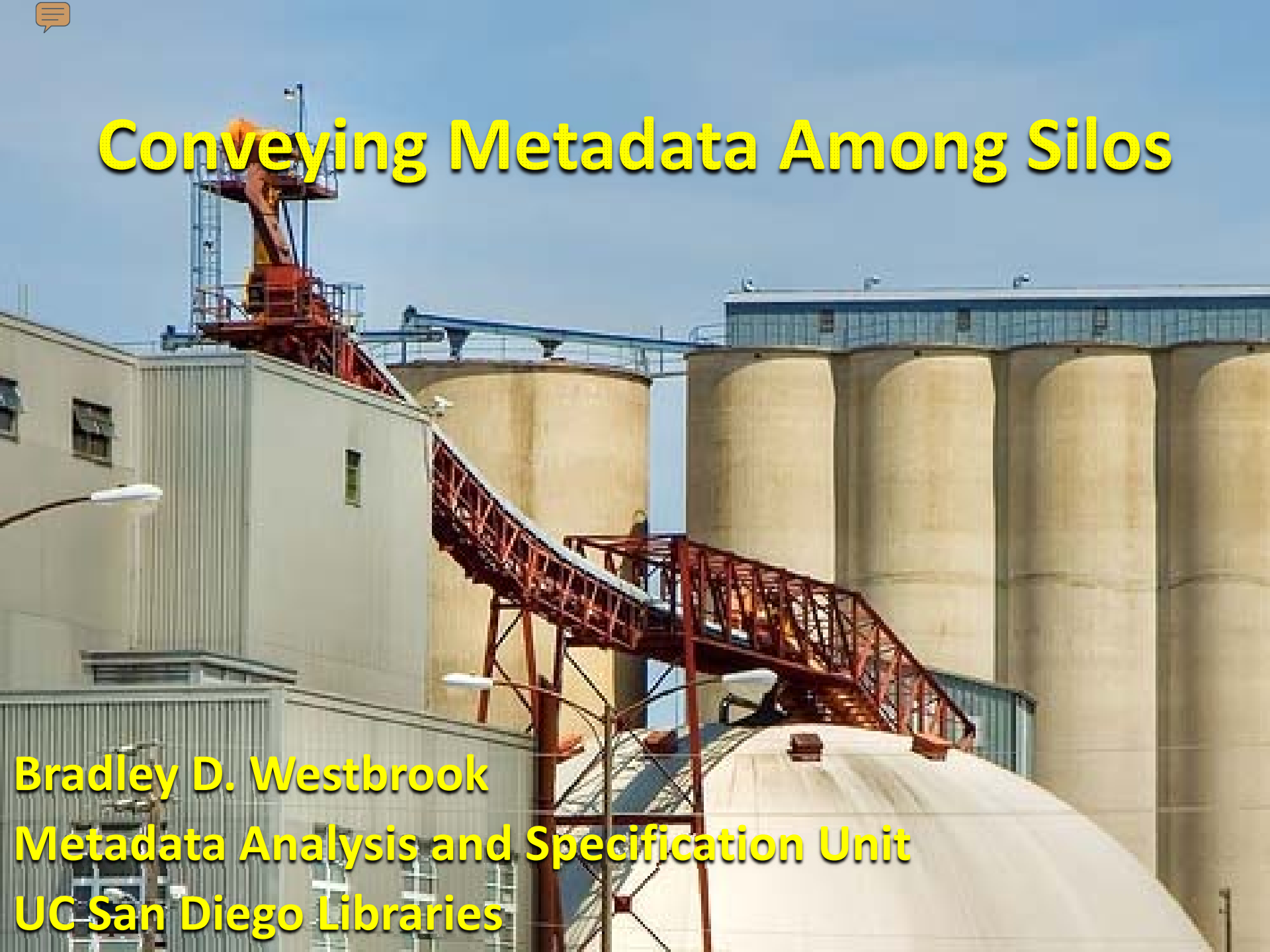
CLIR's Cataloging Hidden Special Collections and Archives Program, 2009

"Applicants will be expected to base their proposals on technical solutions that already exist for swift and efficient entry of data which can be translated into standard records formats such as EAD and MARC"

Thank You!

rochandler@ucsd.edu

Conveying Metadata Among Silos



Bradley D. Westbrook
Metadata Analysis and Specification Unit
UC San Diego Libraries

Overview

- Reframing the “silo” problem
- Challenges for metadata sharing
- Tools / methods for metadata sharing
- The problem of metadata synchronicity

Some facts about metadata silos

- Inevitable
- Ubiquitous
- Beneficial

What's good about metadata silos?

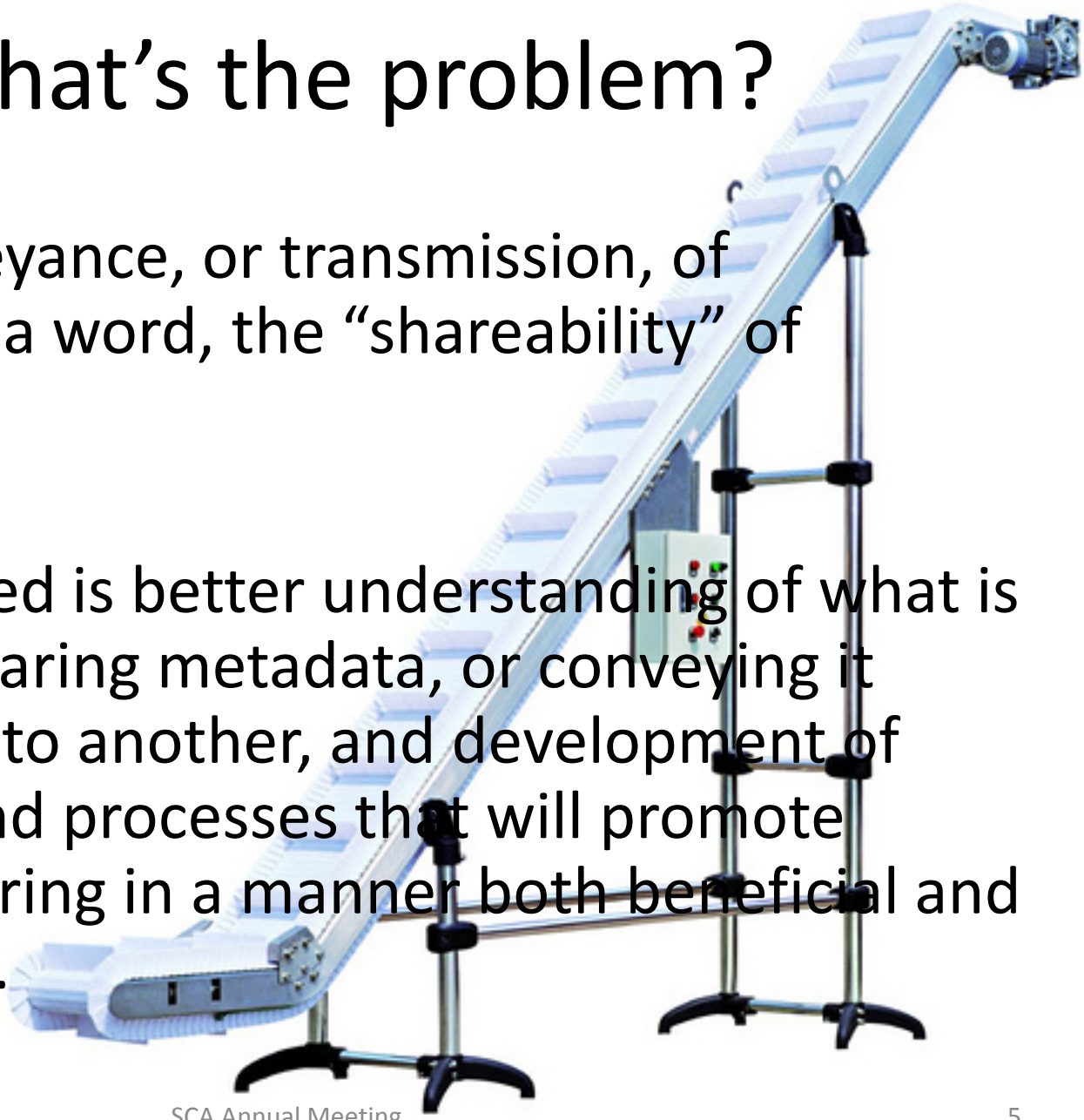
- Metadata silos share core purposes
  - Finding
  - Identifying
  - Selecting
  - Acquiring

IFLA Study Group on the Functional Requirements for Bibliographic Records (1997)

So what's the problem?

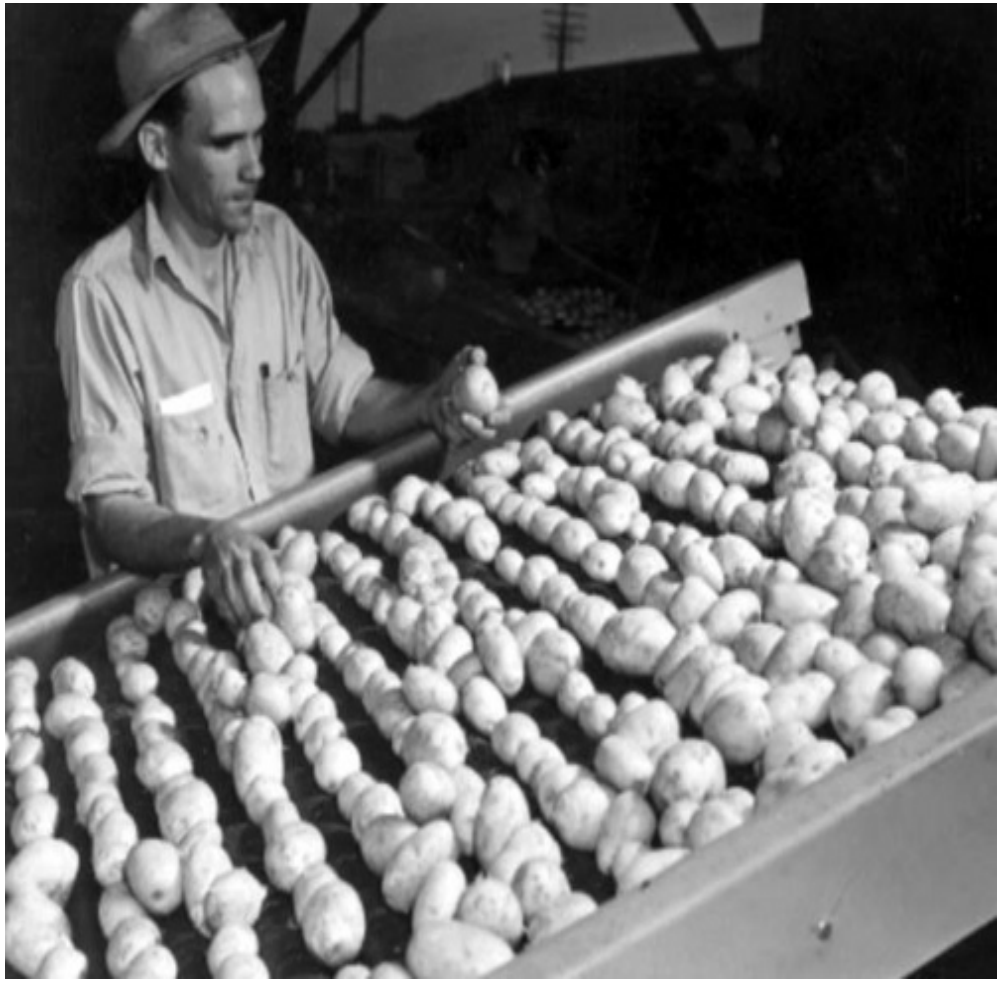
It is the conveyance, or transmission, of metadata. In a word, the "shareability" of metadata.

What is needed is better understanding of what is involved in sharing metadata, or conveying it from one silo to another, and development of techniques and processes that will promote metadata sharing in a manner both beneficial and cost-effective.



What's so important about sharing?

- Efficiency
- Increased access
- New data products
- Professionalism



Challenges for metadata sharing

- Increasingly complicated data environment
- Transformation barriers

Increasingly complicated metadata environment

- Book catalogs
- AACR and card catalogs
- Multiple content standards for different kinds of objects and functions
- Multiple format standards for communicating metadata
- Intersecting but bounded knowledge consuming / producing communities

Transformation barriers

Inequalities among standards create barriers when conveying metadata from one environment to another.

Kinds of Inequalities

- Specific to general
- Differences in data format / values
- One to many (splitting)
- Many to one (aggregation)
- Structural differences
- Missing data
- Superfluous data

Specific to general

Creator

source

- Author
- Illustrator
- Sculptor
- Etc.

target

Creator

Subject

source (marc)

```
650  0 Northern elephant seal|xGeographical distribution
       |zSouthern California Bight (Calif. and Mexico)
```

target (dc)

<dc:subject>Northern elephant seal—Geographical distribution –Southern California Bight (Calif. and Mexico)</dc:subject>

Format / value differences

Syntax

source	target
06-11-1958	1958-11-06
mm-dd-yyyy	yyyy-mm-dd

Vocabulary

file types:

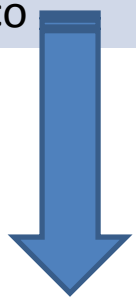
source	target
Image	still image
Text	text
Audio	sound recording
Video	moving image

One to many

The value of a single source element is divided into multiple elements in the target schema:

source (local database)

Subjects
Submarine valleys; Ocean bottom; Marine Sediments; Scientific Expeditions—Mexico



processing: **Multiple values are separated by a semicolon in the database, these should be separated and each value put into a separate <mods:genre> element.**

target (dc)

```
<dc:subject>Submarine valleys</dc:subject>
<dc:subject>Ocean bottom</dc:subject>
<dc:subject>Marine Sediments</dc:subject>
<dc:subject>Scientific Expeditions -- Mexico</dc:subject>
```

Many to one

The values of multiple source elements are combined into a single target element.

source (database)

Repositories/repositoryName
Repositories/address1
Repositories/address1
Repositories/address3
Repositories/city
Repositories/region
Repositories/mailCode
Repositories/country
Repositories/url

processing

Concatenate repositoryName, address1, address2, address3, city region mailCode, counry (url)

target (mods)

mods/note[@displayLabel='Digital object made available by']

<note displayLabel="Digital object made available by"> **Mandeville Special Collections Library, University of California, San Diego, La Jolla, CA 92093-0175 (http://orpheus.ucsd.edu/speccoll/)**</note>

Structural difference

Hierarchical

```
<mods xsi:schemaLocation="http://www.loc.gov/mods/v3 http://w
<relatedItem type="constituent">
   <titleInfo>
     <title> Report to the Commissioner of the Federal
       Public Housing Authority</title>
   </titleInfo>
   <identifier type="ARK">
     http://ark.cdlib.org/ark:/13030/hb50000652/FID28
   </identifier>
</relatedItem>
 <relatedItem type="constituent">
     <titleInfo>
       <title> Minutes from Housing Authority meeting</title>
     </titleInfo>
     <identifier type="ARK">
       http://ark.cdlib.org/ark:/13030/hb50000652/FID15
     </identifier>
  </relatedItem>
</mods>
```

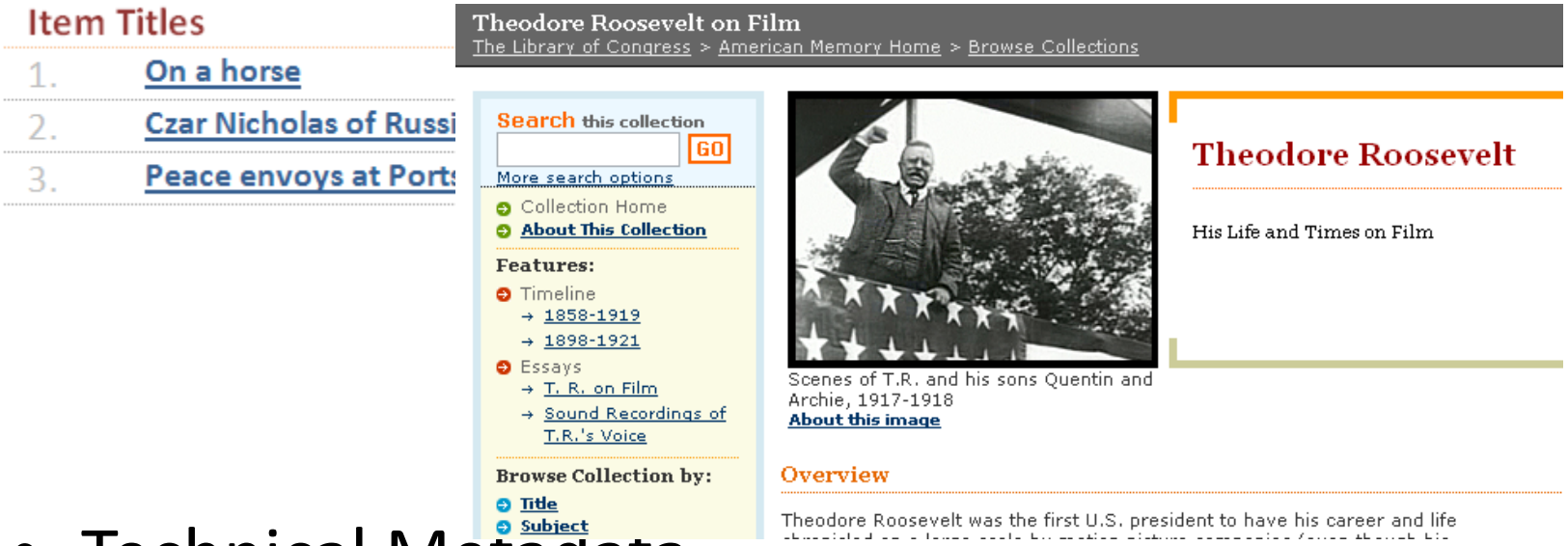
1

2

```
<dc xsi:schemaLocation="http://www.cdlib.org/schemas/xmldata http://ark.cdlib.or
   <relation>Minutes from Housing Authority meeting</relation>  1
   <relation>
      Report to the Commissioner of the Federal Public Housing Authority  2
   </relation>
   <identifier>http://ark.cdlib.org/ark:/13030/hb50000652/FID28</identifier>  1
   <identifier>http://ark.cdlib.org/ark:/13030/hb50000652/FID15 </identifier>  2
</dc>
```

Missing data

- Collection Description



- Technical Metadata

to support different system functions, like preservation

Replacing missing data

- Take from collection level metadata or context
- Draw on non-explicit metadata (e.g., organizational practices, workflow documentation, cataloging guidelines)
- Extract from files (mainly for technical metadata)
- Infer from existing metadata (e.g., if a file has an mp3 extension, assign a mime type of audio/mpeg)

Superfluous local metadata

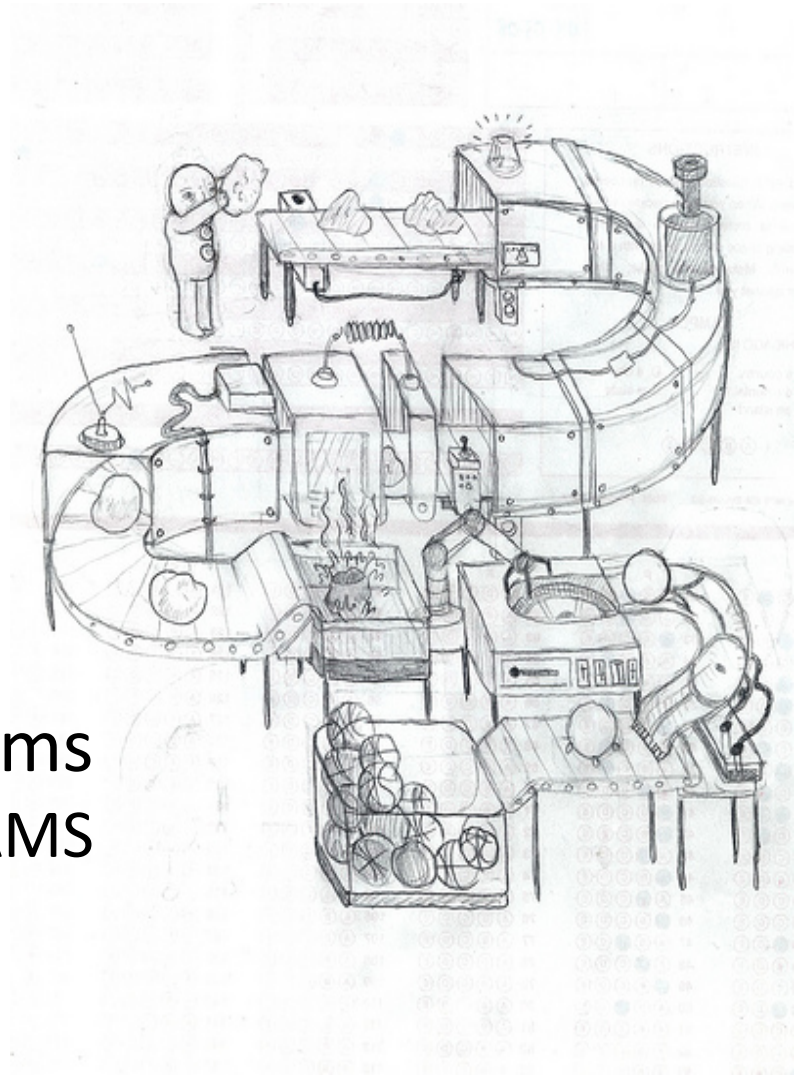
Relevant in the original context of an item but not in its new environment.

Workflow and management metadata, e.g.

- who digitized an item and other work actions
- accession and stack location information
- notes about the physical condition of a source item
- terms used to drive local access mechanisms

Tools / methods sharing metadata

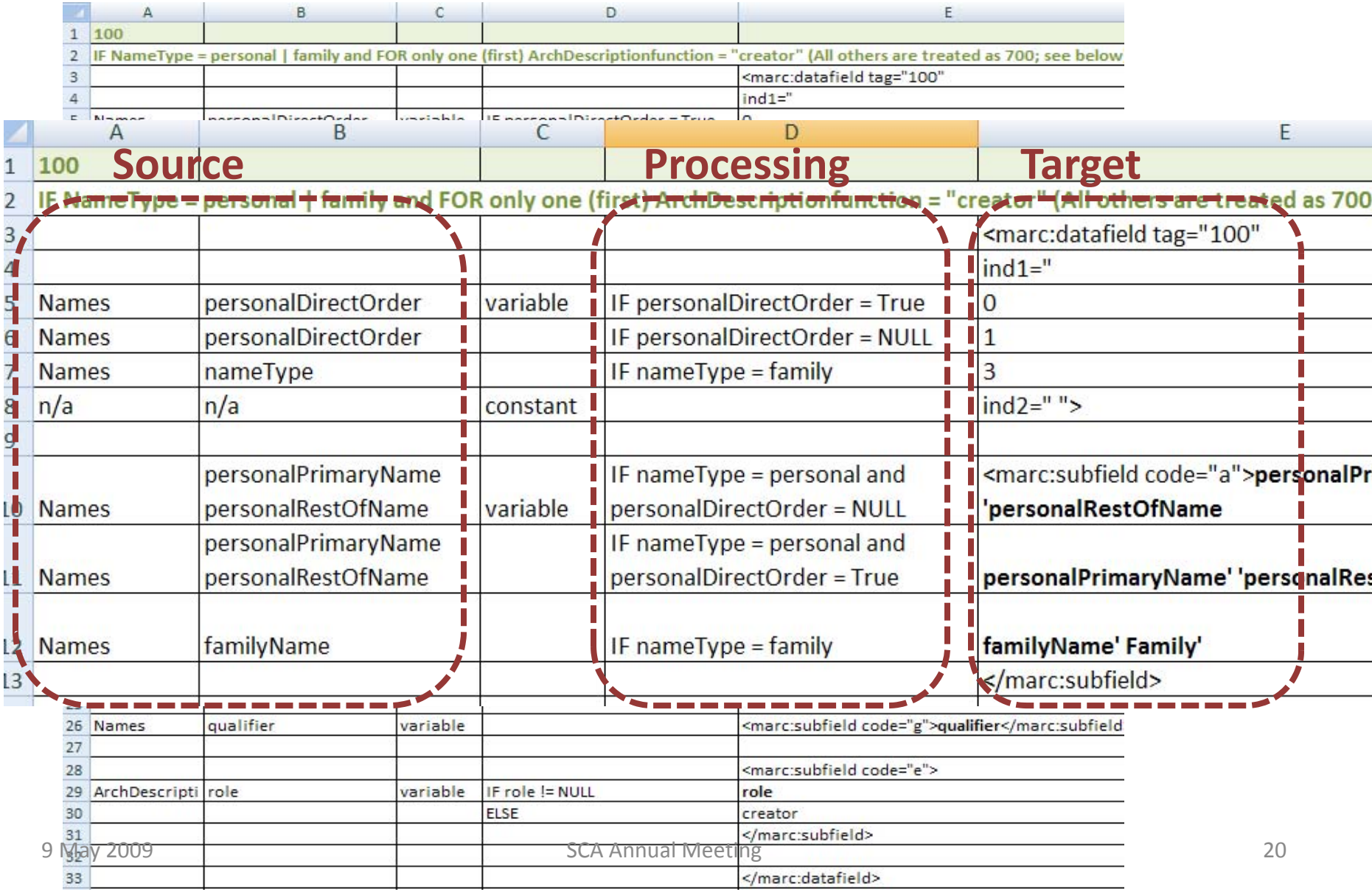
- Basic software programs
  - Excel
  - XML editors
  - XSLT
  - Altova Map Force
- Collection management / metadata authoring systems
  - UC San Diego Libraries' DAMS
  - Archivists' Toolkit



Excel

- Good for visual scanning of data for patterns or disruptions of patterns
- Easy to add data
- Supports a variety of data manipulation
  - Find and Replace
  - Concatenation
  - Date formatting
  - A lot more…

Archivists' Toolkit export metadata map



	A	B	C	D	E
1	100				
2	IF NameType = personal | family and FOR only one (first) ArchDescriptionfunction = "creator" (All others are treated as 700; see below				
3					<marc:datafield tag="100"
4					ind1="
5	Names	personalDirectOrder	variable	IF personalDirectOrder = True	0
6	Names	personalDirectOrder		IF personalDirectOrder = NULL	1
7	Names	nameType		IF nameType = family	3
8	n/a	n/a	constant		ind2=" ">
9					
10	Names	personalPrimaryName personalRestOfName	variable	IF nameType = personal and personalDirectOrder = NULL	<marc:subfield code="a">**personalPr** **'personalRestOfName**
11	Names	personalPrimaryName personalRestOfName		IF nameType = personal and personalDirectOrder = True	**personalPrimaryName' 'personalRes**
12	Names	familyName		IF nameType = family	**familyName' Family'**
13					</marc:subfield>
26	Names	qualifier	variable		<marc:subfield code="g">**qualifier**</marc:subfield
27					
28					<marc:subfield code="e">
29	ArchDescripti	role	variable	IF role != NULL	**role**
30				ELSE	creator
31					</marc:subfield>
32					
33					</marc:datafield>

Source

Processing

Target

XML Editors

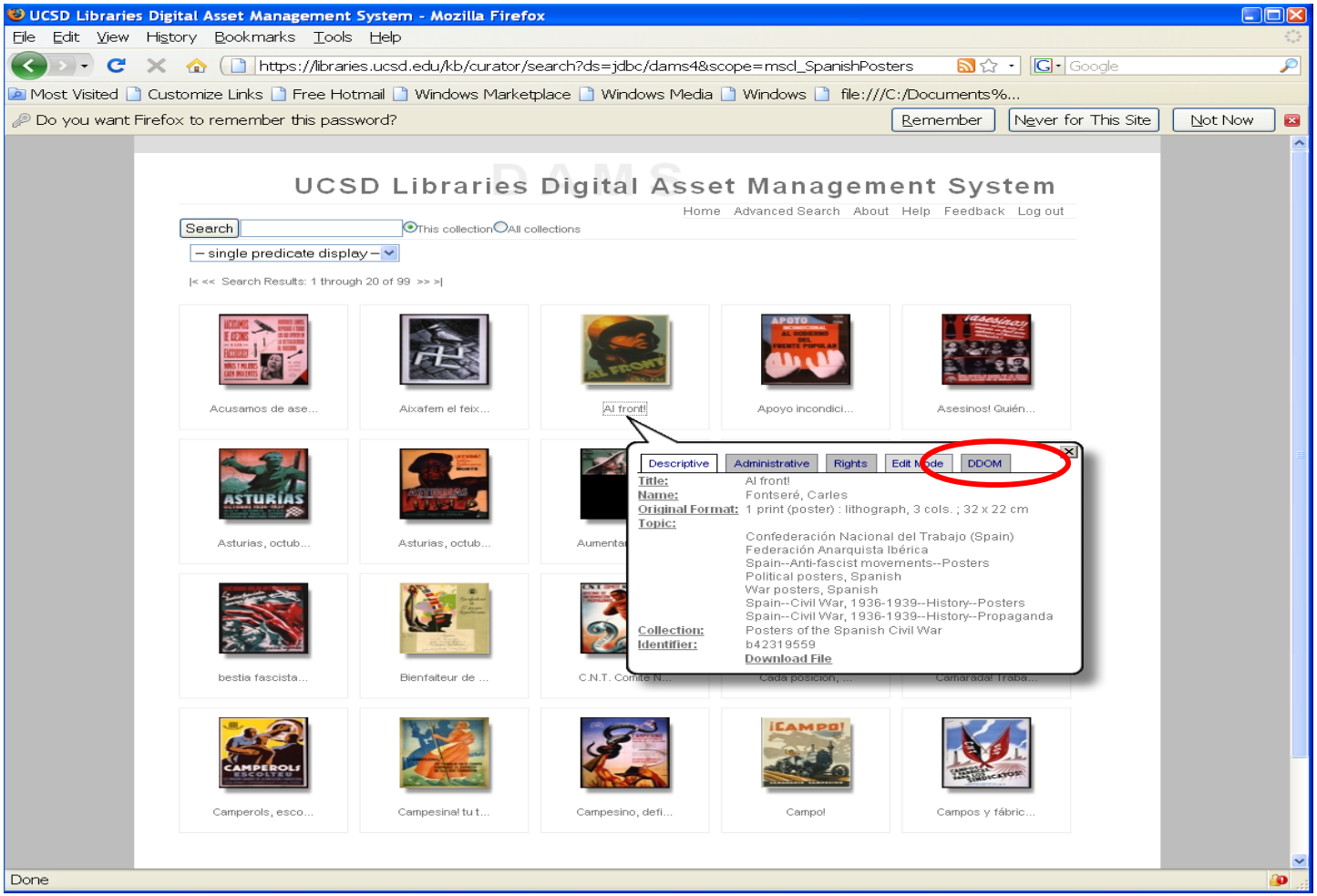
- For working with XML encoded data
  - Can import excel spreadsheets into xml format
  - Can be used to process xml data with an xslt style sheet
  - Contextual find and replace
  - Record / model validation

XSLT

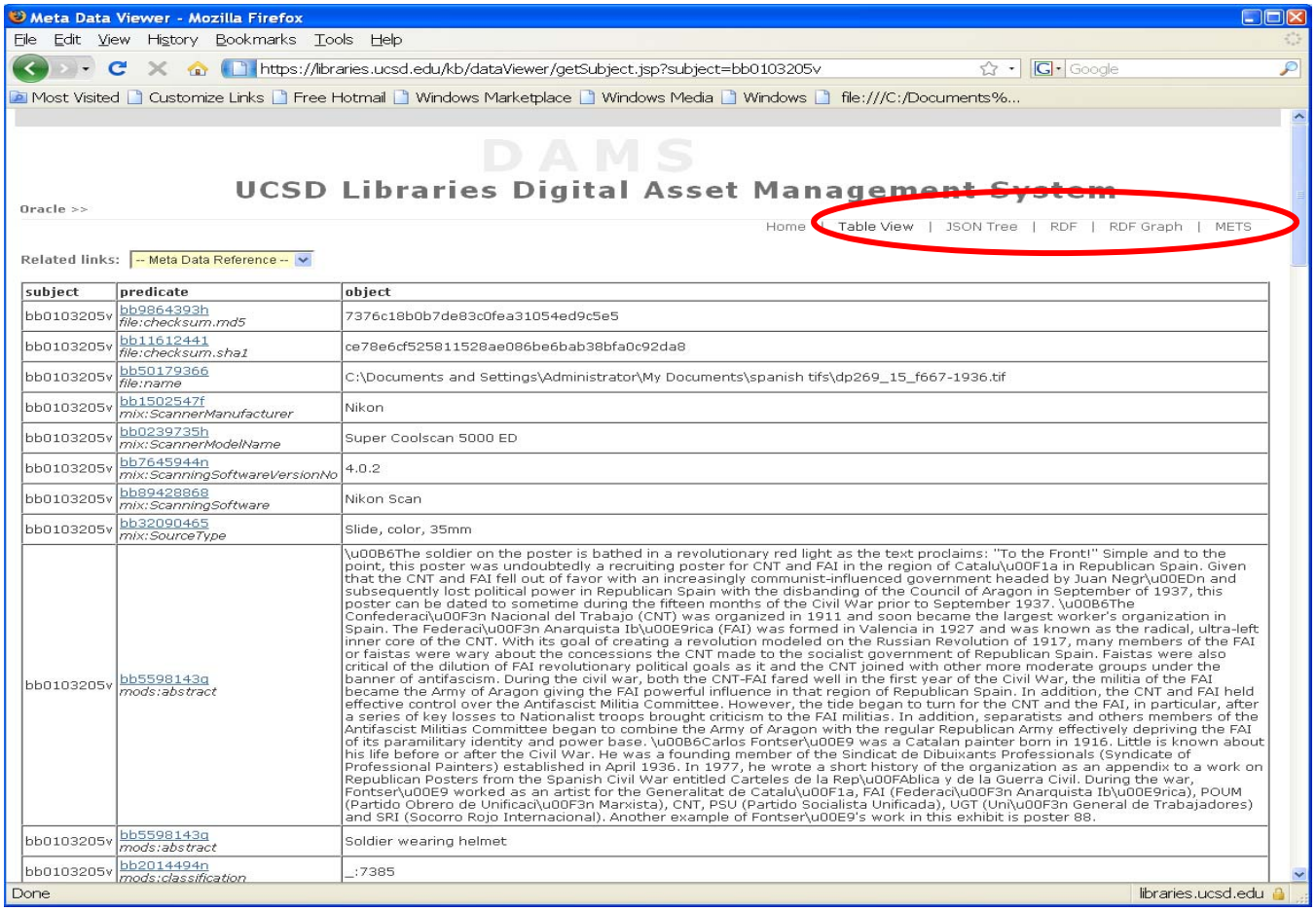
Extensible Stylesheet Language Transformations

- For transformation of XML encoded data
- High technical requirement to create, modify and troubleshoot
- Can be reused & shared
- Supports more advanced data transformations

UC San Diego Libraries' DAMS

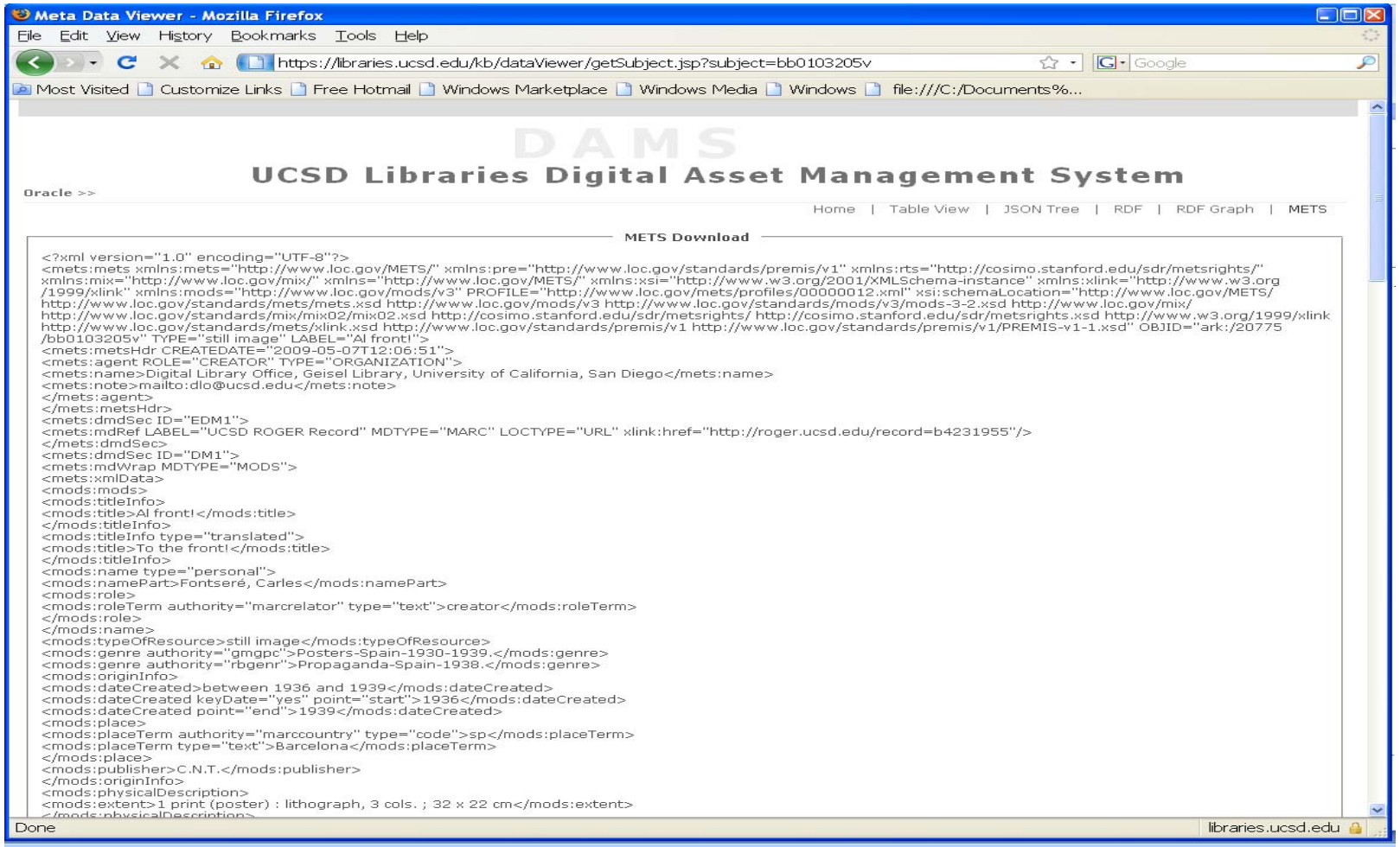


Multiple metadata views via XSLT

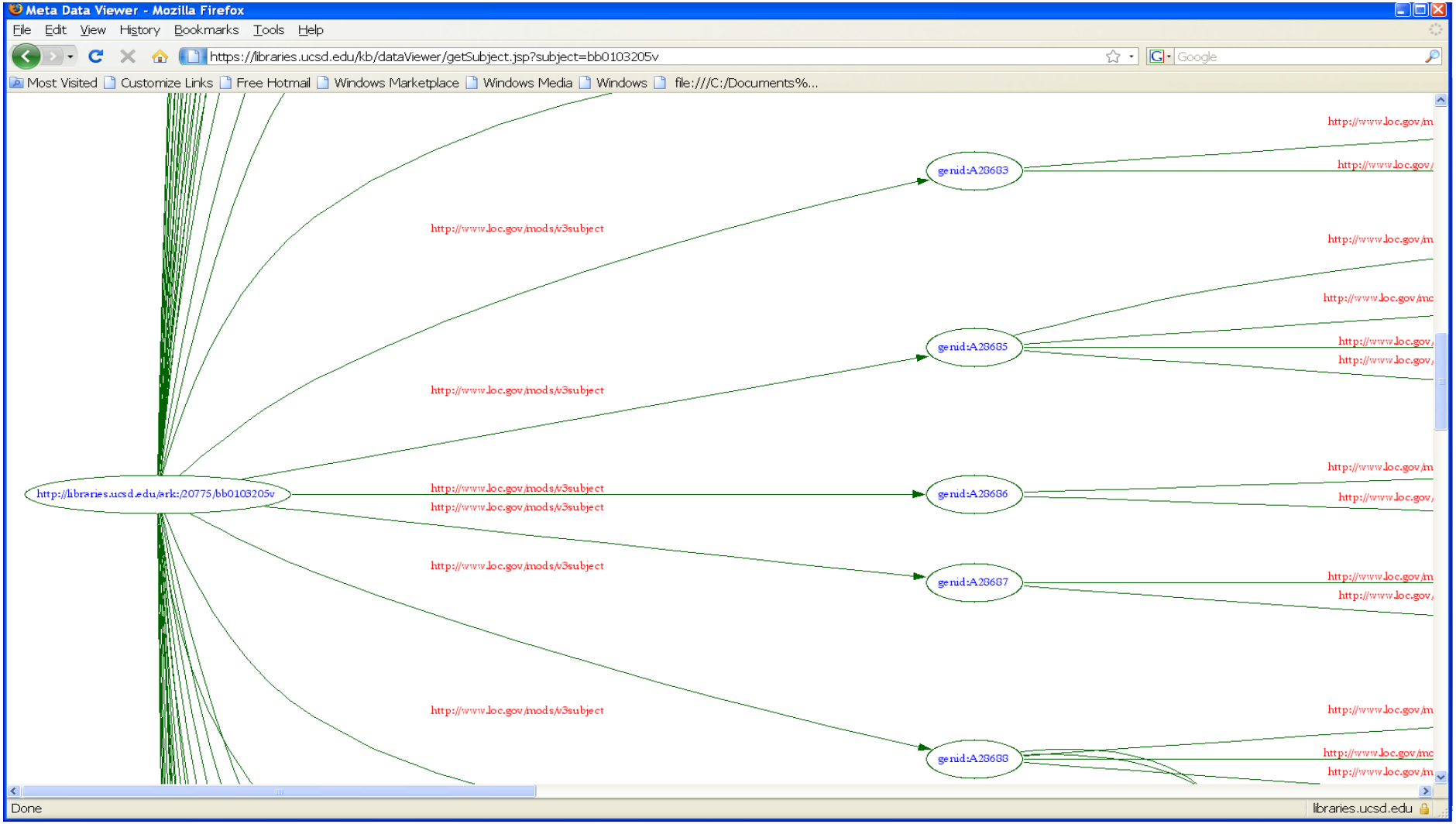


subject	predicate	object
bb0103205v	bb9864393h *file:checksum.md5*	7376c18b0b7de83c0fea31054ed9c5e5
bb0103205v	bb11612441 *file:checksum.sha1*	ce78e6cf525811528ae086be6bab38bfa0c92da8
bb0103205v	bb50179366 *file:name*	C:\Documents and Settings\Administrator\My Documents\spanish tifs\dp269_15_f667-1936.tif
bb0103205v	bb1502547f *mix:ScannerManufacturer*	Nikon
bb0103205v	bb0239735h *mix:ScannerModelName*	Super Coolscan 5000 ED
bb0103205v	bb7645944n *mix:ScanningSoftwareVersionNo*	4.0.2
bb0103205v	bb89428868 *mix:ScanningSoftware*	Nikon Scan
bb0103205v	bb32090465 *mix:SourceType*	Slide, color, 35mm
bb0103205v	bb5598143g *mods:abstract*	\u00B6The soldier on the poster is bathed in a revolutionary red light as the text proclaims: "To the Front!" Simple and to the point, this poster was undoubtedly a recruiting poster for CNT and FAI in the region of Catalu\u00F1a in Republican Spain. Given that the CNT and FAI fell out of favor with an increasingly communist-influenced government headed by Juan Negr\u00EDn and subsequently lost political power in Republican Spain with the disbanding of the Council of Aragon in September of 1937, this poster can be dated to sometime during the fifteen months of the Civil War prior to September 1937. \u00B6The Confederaci\u00F3n Nacional del Trabajo (CNT) was organized in 1911 and soon became the largest worker's organization in Spain. The Federaci\u00F3n Anarquista Ib\u00E9rica (FAI) was formed in Valencia in 1927 and was known as the radical, ultra-left inner core of the CNT. With its goal of creating a revolution modeled on the Russian Revolution of 1917, many members of the FAI or faistas were wary about the concessions the CNT made to the socialist government of Republican Spain. Faistas were also critical of the dilution of FAI revolutionary political goals as it and the CNT joined with other more moderate groups under the banner of antifascism. During the civil war, both the CNT-FAI fared well in the first year of the Civil War, the militia of the FAI became the Army of Aragon giving the FAI powerful influence in that region of Republican Spain. In addition, the CNT and FAI held effective control over the Antifascist Militia Committee. However, the tide began to turn for the CNT and the FAI, in particular, after a series of key losses to Nationalist troops brought criticism to the FAI militias. In addition, separatists and others members of the Antifascist Militias Committee began to combine the Army of Aragon with the regular Republican Army effectively depriving the FAI of its paramilitary identity and power base. \u00B6Carlos Fontser\u00E9 was a Catalan painter born in 1916. Little is known about his life before or after the Civil War. He was a founding member of the Sindicat de Dibuixants Professionals (Syndicate of Professional Painters) established in April 1936. In 1977, he wrote a short history of the organization as an appendix to a work on Republican Posters from the Spanish Civil War entitled Carteles de la Rep\u00FAblica y de la Guerra Civil. During the war, Fontser\u00E9 worked as an artist for the Generalitat de Catalu\u00F1a, FAI (Federaci\u00F3n Anarquista Ib\u00E9rica), POUM (Partido Obrero de Unificaci\u00F3n Marxista), CNT, PSU (Partido Socialista Unificada), UGT (Uni\u00F3n General de Trabajadores) and SRI (Socorro Rojo Internacional). Another example of Fontser\u00E9's work in this exhibit is poster 88.
bb0103205v	bb5598143g *mods:abstract*	Soldier wearing helmet
bb0103205v	bb2014494n *mods:classification*	_:7385

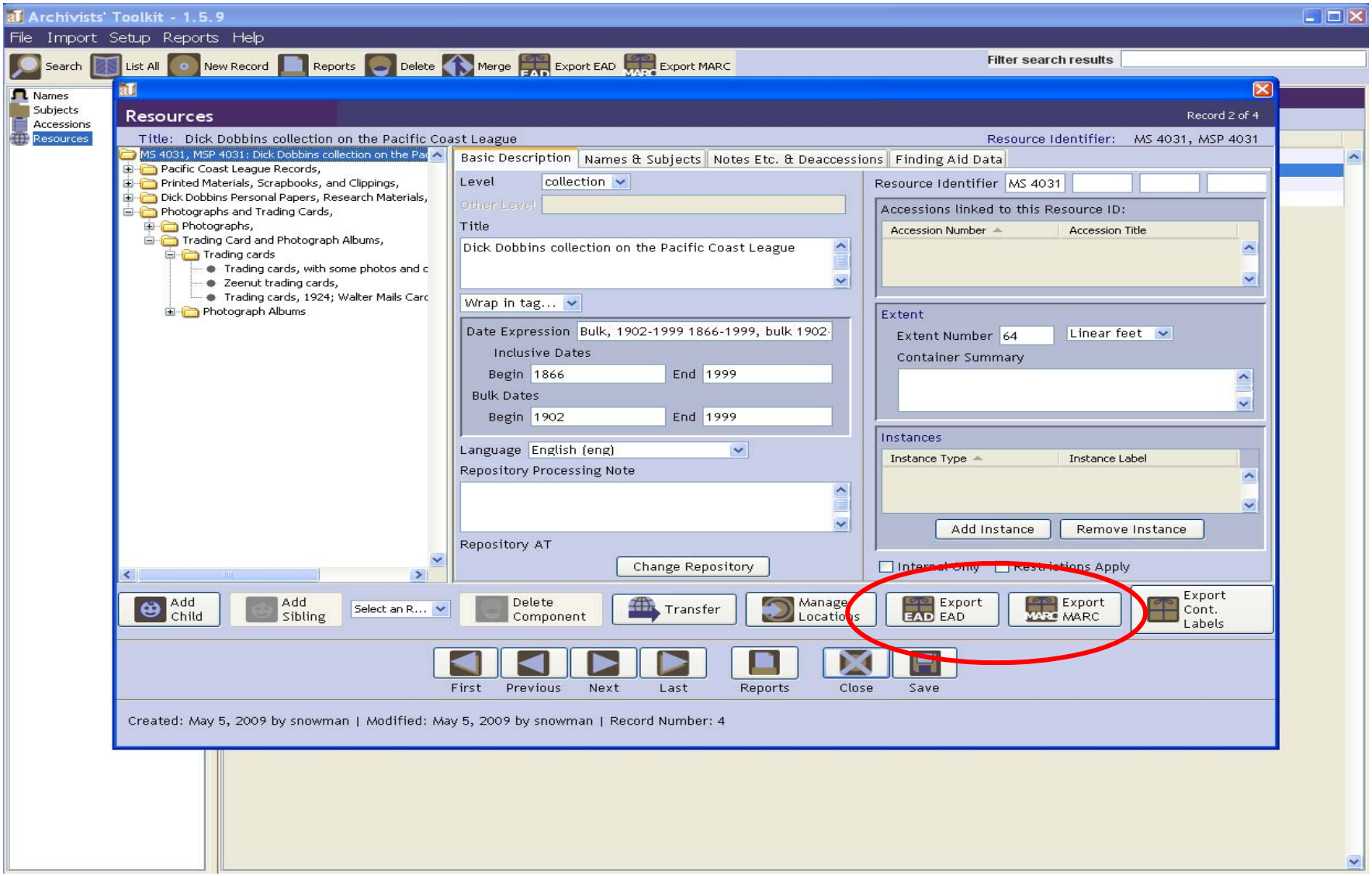
METS view


Meta Data Viewer - Mozilla Firefox
File Edit View History Bookmarks Tools Help
https://libraries.ucsd.edu/kb/dataViewer/getSubject.jsp?subject=bb0103205v
Most Visited Customize Links Free Hotmail Windows Marketplace Windows Media Windows file:///C:/Documents%...
DAMS
UCSD Libraries Digital Asset Management System
Oracle >>
Home | Table View | JSON Tree | RDF | RDF Graph | METS
METS Download
<?xml version="1.0" encoding="UTF-8"?>
<mets:mets xmlns:mets="http://www.loc.gov/METS/" xmlns:pre="http://www.loc.gov/standards/premis/v1" xmlns:rts="http://cosimo.stanford.edu/sdr/metsrights/"
xmlns:mix="http://www.loc.gov/mix/" xmlns="http://www.loc.gov/METS/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:xlink="http://www.w3.org
/1999/xlink" xmlns:mods="http://www.loc.gov/mods/v3" PROFILE="http://www.loc.gov/mets/profiles/00000012.xml" xsi:schemaLocation="http://www.loc.gov/METS/
http://www.loc.gov/standards/mets/mets.xsd http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-2.xsd http://www.loc.gov/mix/
http://www.loc.gov/standards/mix/mix02/mix02.xsd http://cosimo.stanford.edu/sdr/metsrights/ http://cosimo.stanford.edu/sdr/metsrights.xsd http://www.w3.org/1999/xlink
http://www.loc.gov/standards/mets/xlink.xsd http://www.loc.gov/standards/premis/v1 http://www.loc.gov/standards/premis/v1/PREMIS-v1-1.xsd" OBJID="ark:/20775
/bb0103205v" TYPE="still image" LABEL="Al front!">
<mets:metsHdr CREATEDATE="2009-05-07T12:06:51">
<mets:agent ROLE="CREATOR" TYPE="ORGANIZATION">
<mets:name>Digital Library Office, Geisel Library, University of California, San Diego</mets:name>
<mets:note>mailto:dlo@ucsd.edu</mets:note>
</mets:agent>
</mets:metsHdr>
<mets:dmdSec ID="EDM1">
<mets:mdRef LABEL="UCSD ROGER Record" MDTYPE="MARC" LOCTYPE="URL" xlink:href="http://roger.ucsd.edu/record=b4231955"/>
</mets:dmdSec>
<mets:dmdSec ID="DM1">
<mets:mdWrap MDTYPE="MODS">
<mets:xmlData>
<mods:mods>
<mods:titleInfo>
<mods:title>Al front!</mods:title>
</mods:titleInfo>
<mods:titleInfo type="translated">
<mods:title>To the front!</mods:title>
</mods:titleInfo>
<mods:name type="personal">
<mods:namePart>Fontseré, Carles</mods:namePart>
<mods:role>
<mods:roleTerm authority="marcrelator" type="text">creator</mods:roleTerm>
</mods:role>
</mods:name>
<mods:typeOfResource>still image</mods:typeOfResource>
<mods:genre authority="gmgpc">Posters-Spain-1930-1939.</mods:genre>
<mods:genre authority="rbgenr">Propaganda-Spain-1938.</mods:genre>
<mods:originInfo>
<mods:dateCreated>between 1936 and 1939</mods:dateCreated>
<mods:dateCreated keyDate="yes" point="start">1936</mods:dateCreated>
<mods:dateCreated point="end">1939</mods:dateCreated>
<mods:place>
<mods:placeTerm authority="marccountry" type="code">sp</mods:placeTerm>
<mods:placeTerm type="text">Barcelona</mods:placeTerm>
</mods:place>
<mods:publisher>C.N.T.</mods:publisher>
</mods:originInfo>
<mods:physicalDescription>
<mods:extent>1 print (poster) : lithograph, 3 cols. ; 32 x 22 cm</mods:extent>
Done
libraries.ucsd.edu

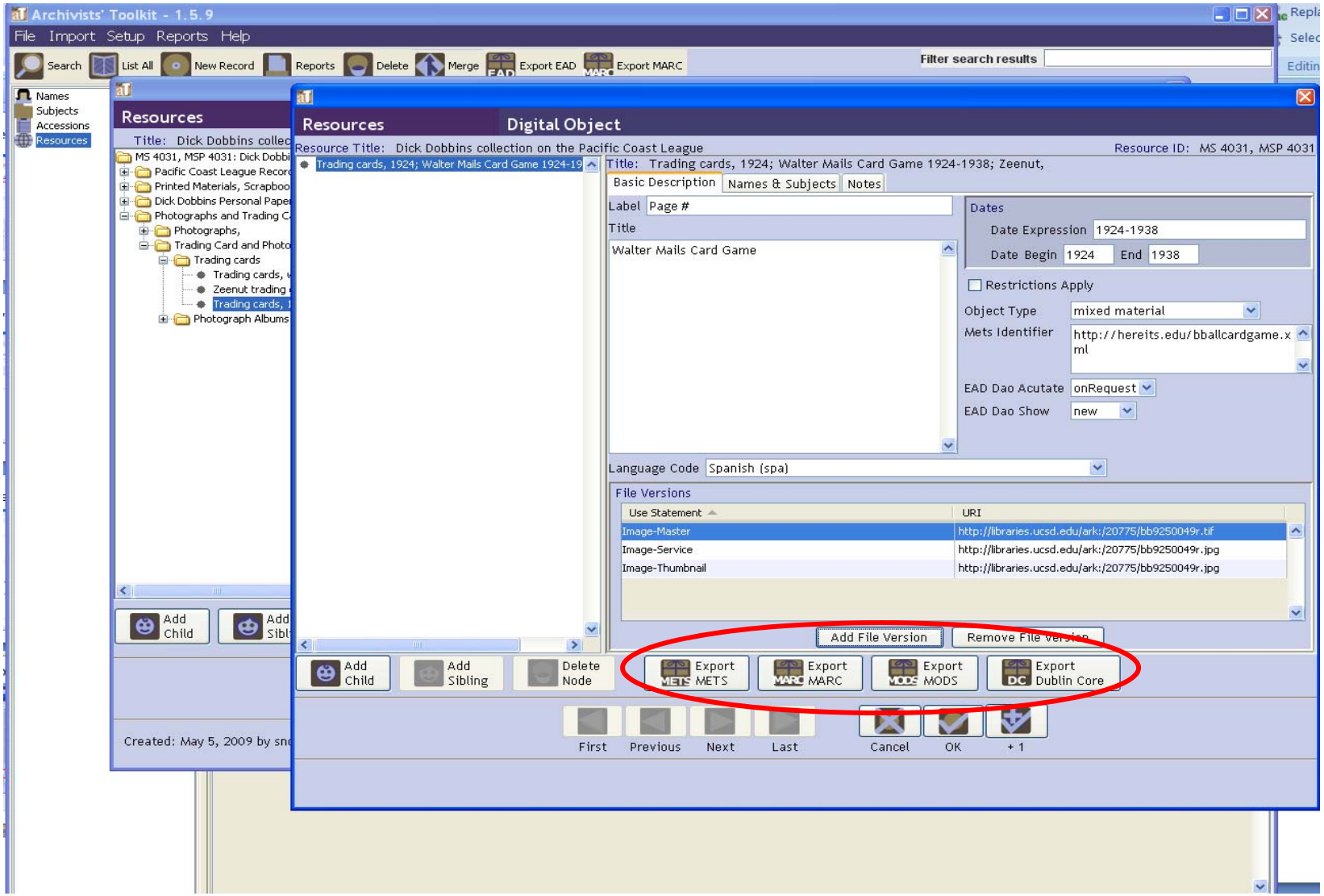
RDF graph view



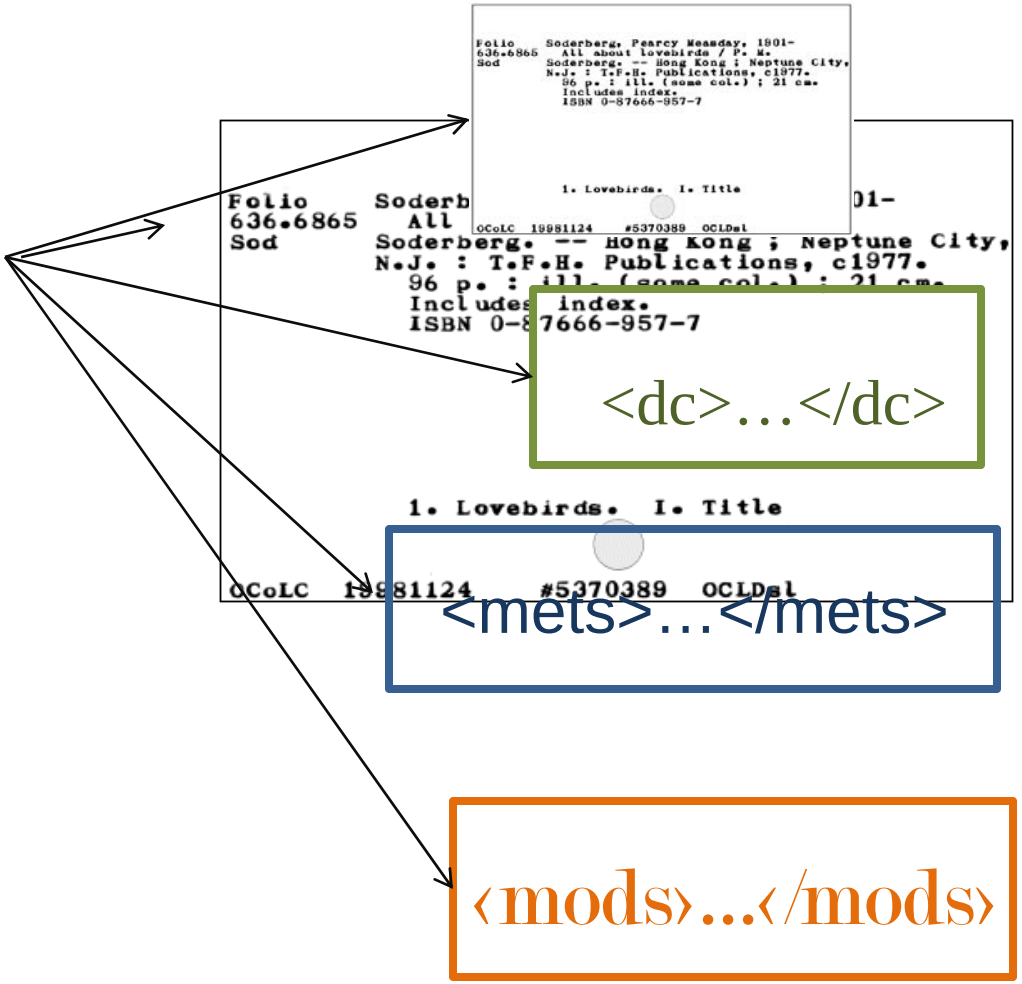
Archivists' Toolkit



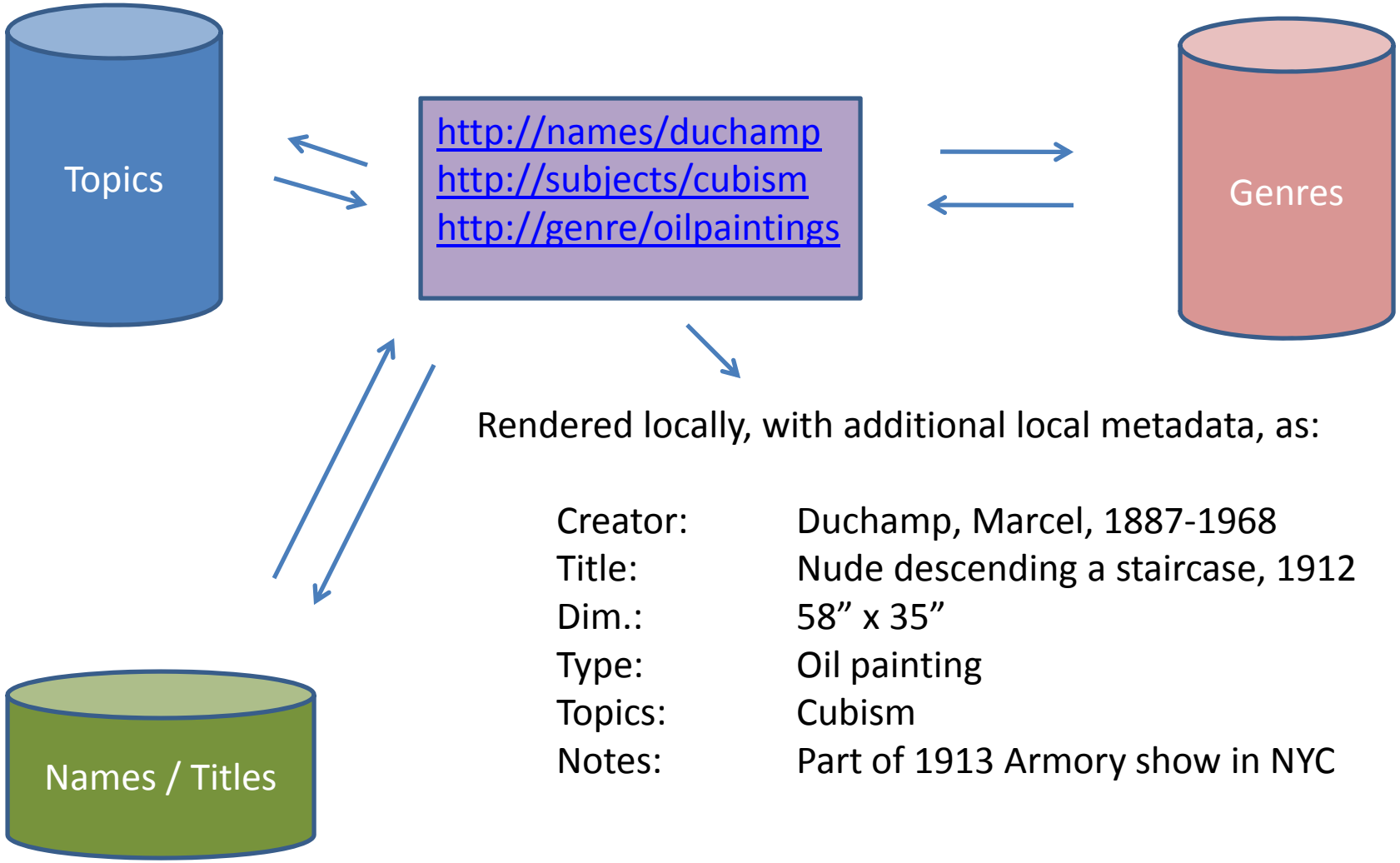
AT DO exports



The problem of metadata synchronicity

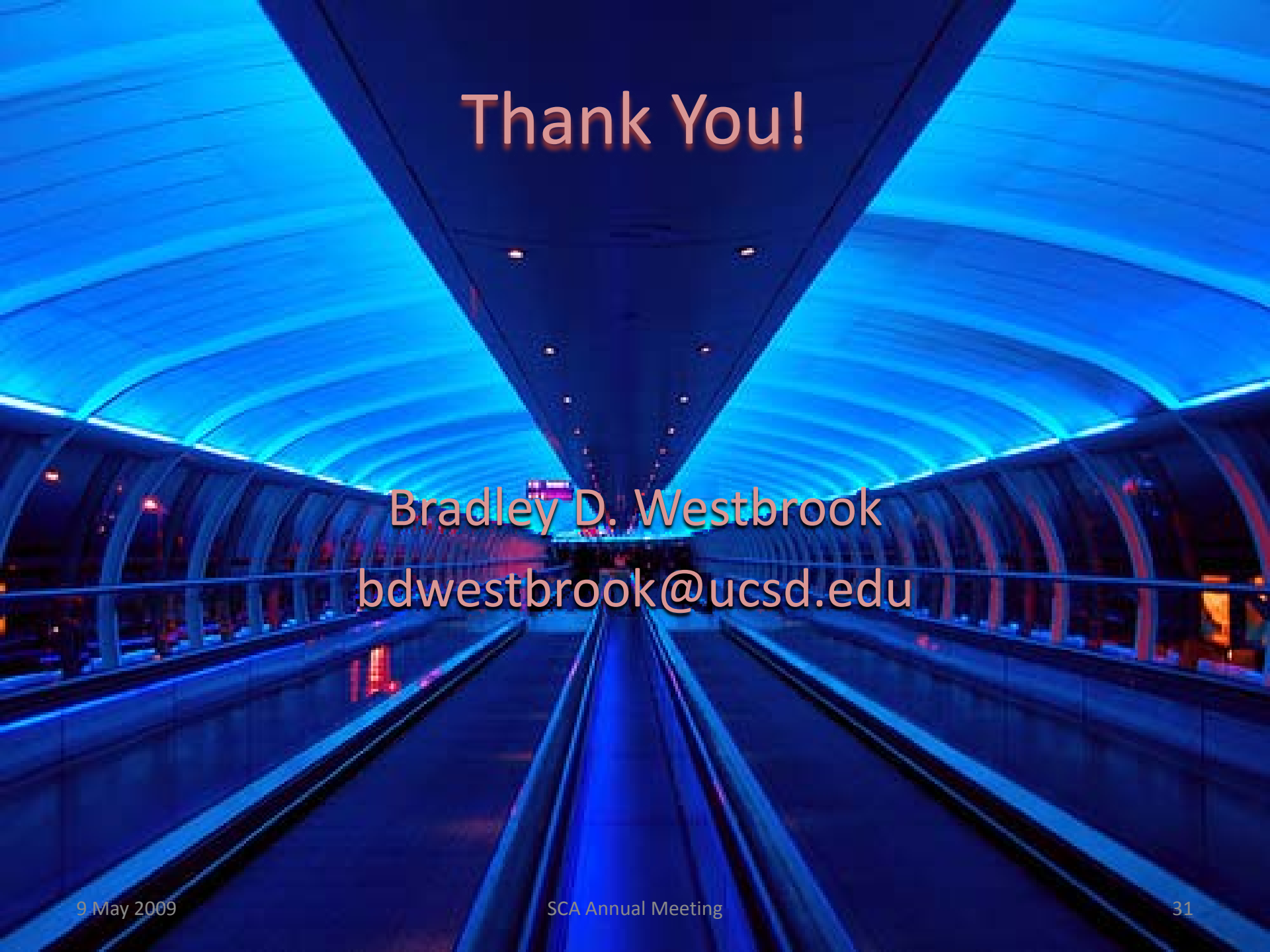


Linked resources



Rendered locally, with additional local metadata, as:

Creator:	Duchamp, Marcel, 1887-1968
Title:	Nude descending a staircase, 1912
Dim.:	58” x 35”
Type:	Oil painting
Topics:	Cubism
Notes:	Part of 1913 Armory show in NYC

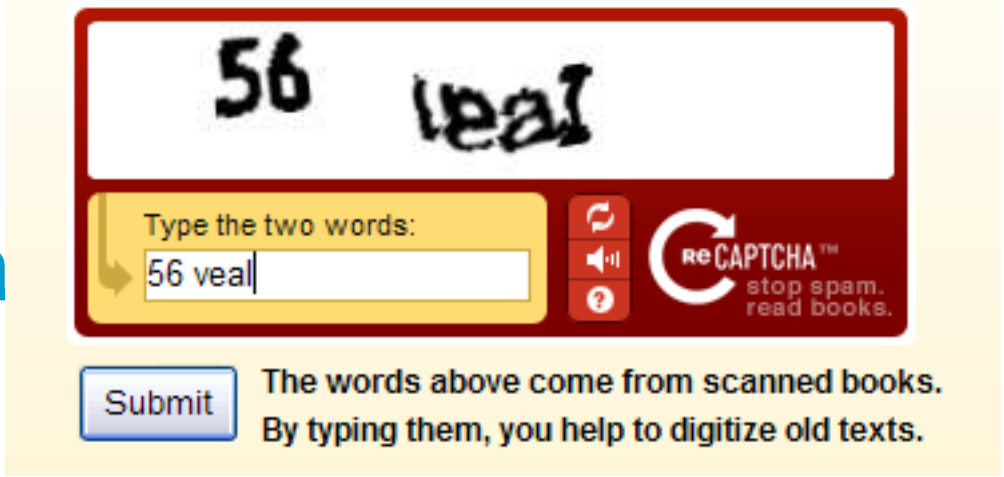


Thank You!

Bradley D. Westbrook
bdwestbrook@ucsd.edu

Metadata in a [redacted]

Metadata in a



Metadata in a

Metadata in a crowd

Metadata in a crowd

shared knowledge

Hi, I’m Kevin

Hi, I'm Kevin Rundblad

What is Metadata in a Crowd?

INTRO: Real 2.0 vs. Simulated 2.0

MAIN: Simon Says “Tag This Photo”
Human Computation Models

Thinking in 2.0 – Engaging the User

2.0 is a Culture, Not

2.0 is a Culture, Not Tech

If the 2.0 idea is not about the user
…it is artificial 2.0

Real 2.0 comes directly from users

Traditional development will struggle with 2.0 development, since it is focused on the process, not the user.

Technology only expresses 2.0 ideas

2.0 is about Creators / Participants

Individuals post valuable info

Many consume it

Expert Seeks Objective Articulation

Structured and Precise

Narrowly Defined Expertise

Social

Very Wide Knowledge

Fragments of Expertise

Tends to be Locally Derived

Subjective Perspective

Emotional Metadata

Social knowledge: Many Perspectives

Is the Crowd
a Viable Source of Metadata?

Aren’t Users Lazy?

Um, Yes and Yes

But they do like to play games…

…spending many hours earning points that don’t mean anything

…motivation is key for participation

…and participation often means completing an intelligent task

(ok, mindless too)

Human Computation

Human Computation

Finding tasks that humans
can do better than computers

Human Computation

Finding tasks that humans
can do better than computers

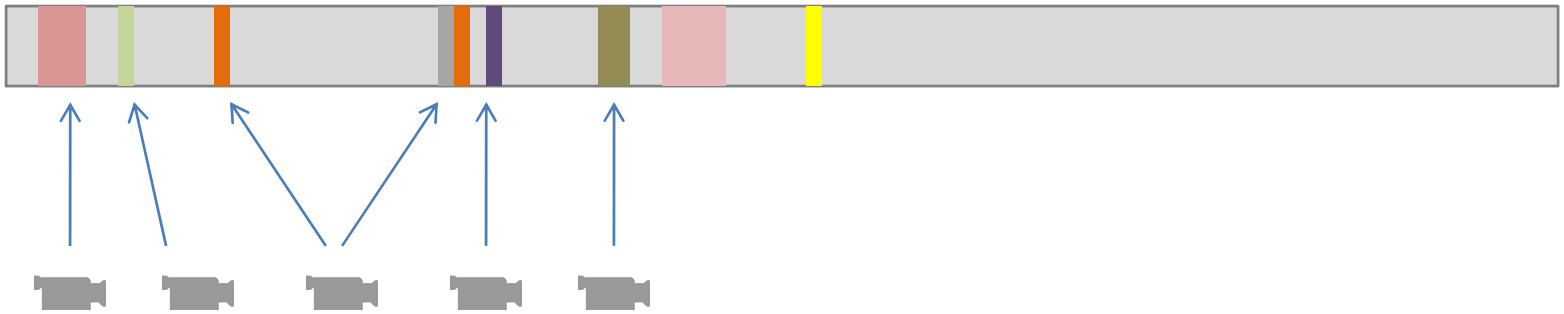
And creating motivation to perform them

Human Computation and

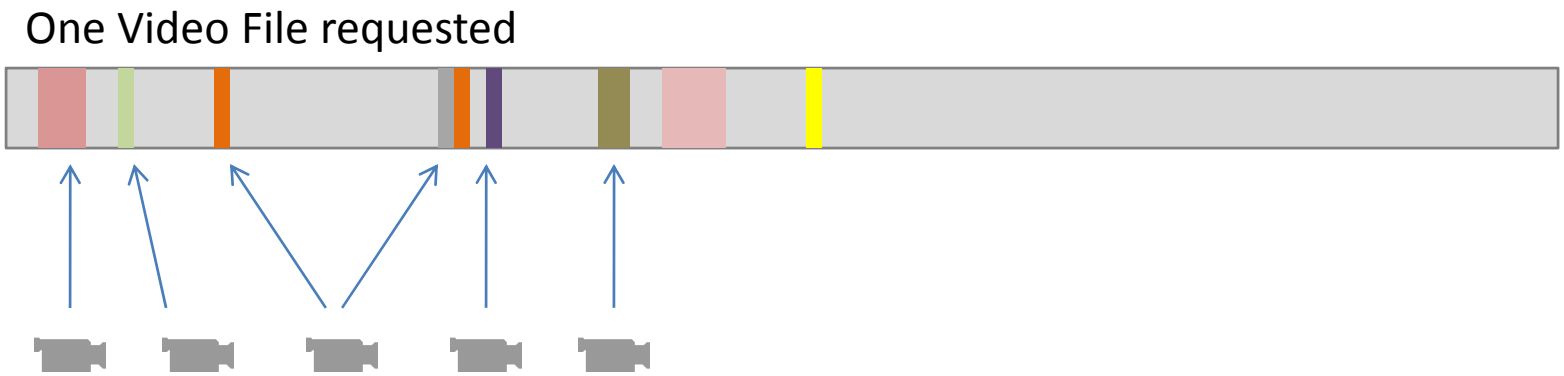
Human Computation and P2P Model

Human Computation and P2P Model

Human Computation and P2P Model

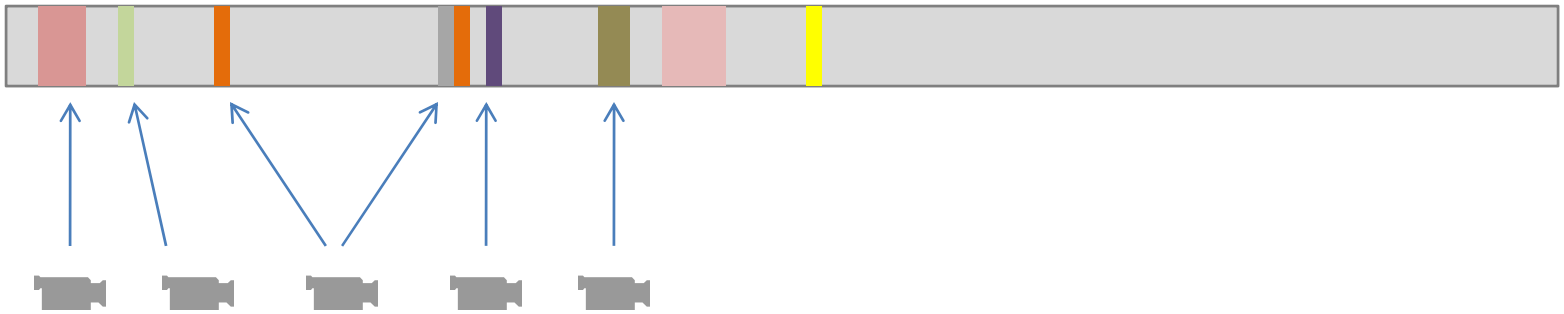


Human Computation and P2P Model



Human Computation and P2P Model

One Video File requested



Many systems each delivering parts of file, until entire file is complete

Socially formed knowledge
is like the P2P distribution architecture

Social Knowledge - Human Computation

Content formed from many individuals

Many individuals contributing parts of the content

But “…unlike computer processors, humans require some incentive to become part of a collective computation”.

Louis Von Ahn

Source: invited talk at CMU on Human Computation

Models of **Human Computation**

AKA Human Intelligence Tasks (HITs)

(Coined by Amazon)

Human Intelligence Tasks – 3 Types

- Socially motivated
- Economically motivated
- Tacit (user may be unaware of task itself they perform)

Human Intelligence Tasks (social)

(or social knowledge systems)

Tagging (Flickr) and Comments (Amazon)
Ratings (Yelp)
Problem Solving (StackOverflow)

Tagging @ flickr

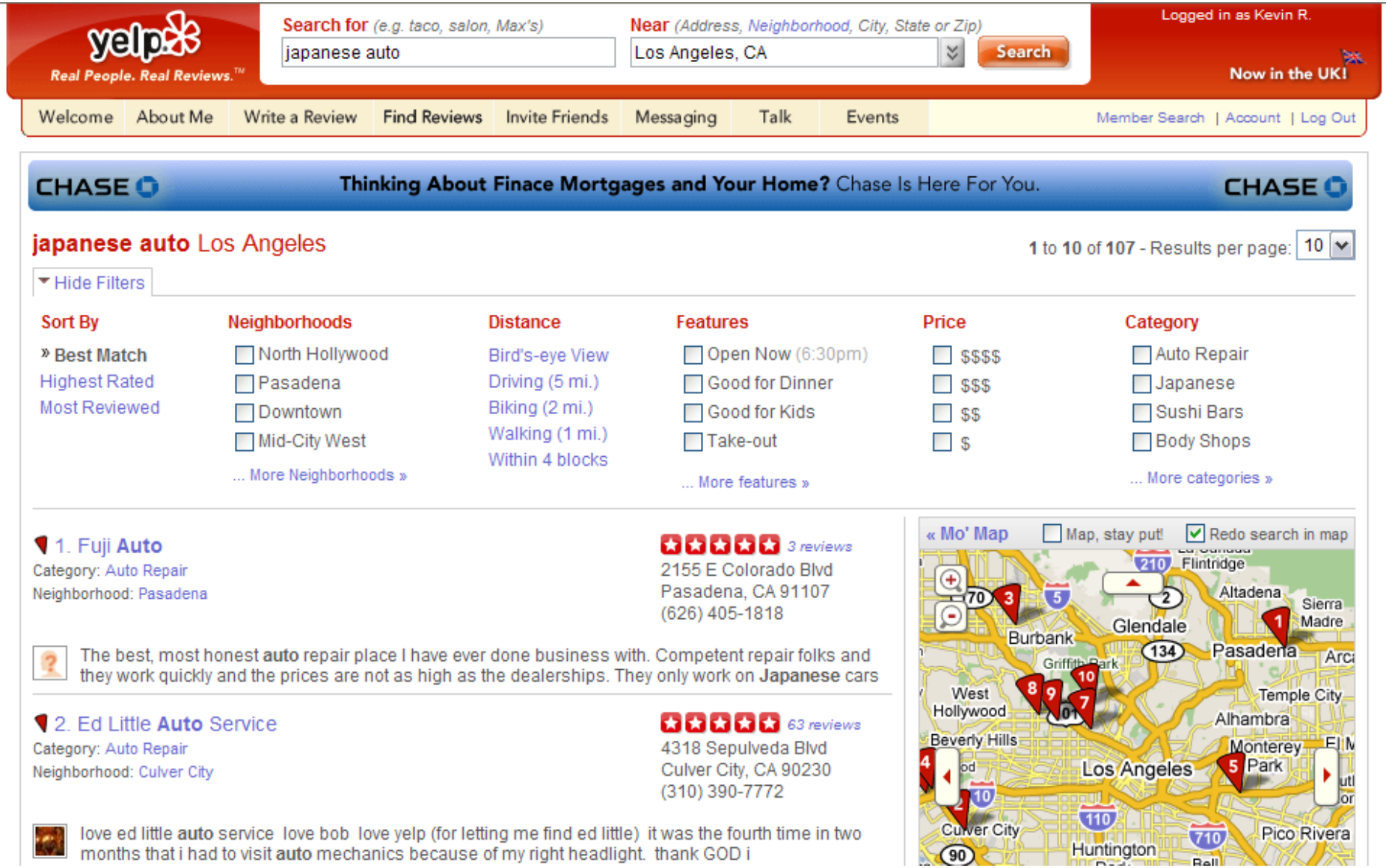
Tagging @ flickr

Wider Discovery

Public Tagging

Ratings @ Yelp

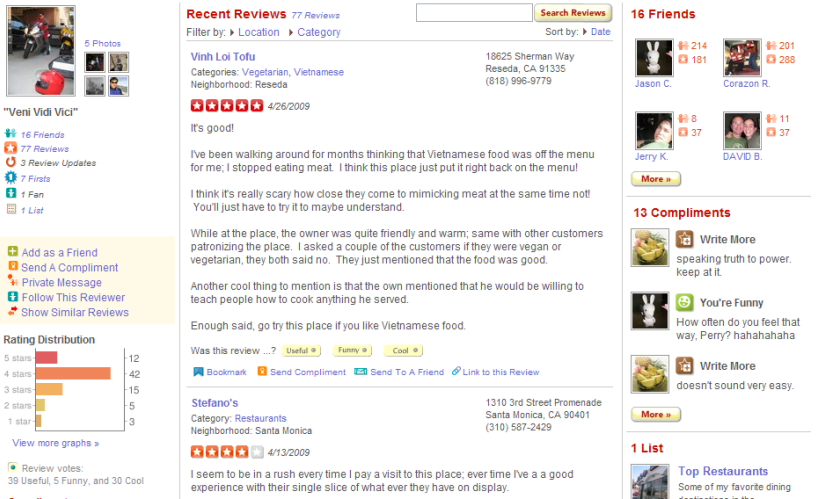
Social is local



Ratings @ Yelp

Social is local

Efficient Discovery

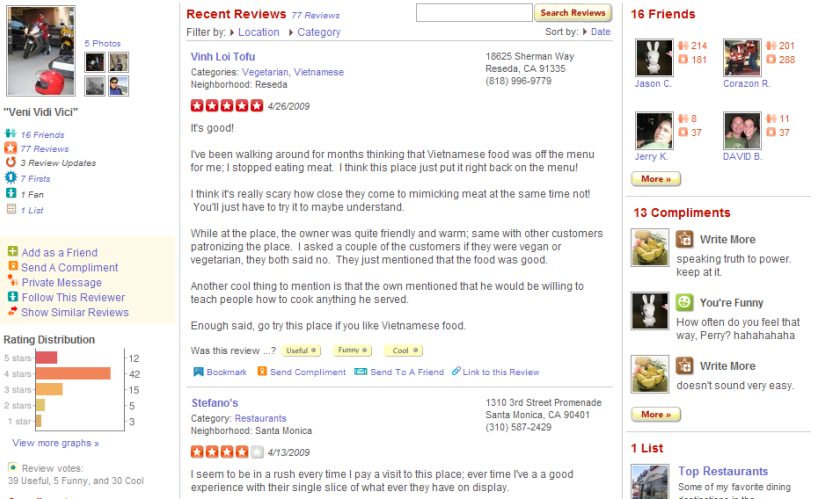


Ratings @ Yelp

Social is local

Efficient Discovery

“Hearsay” becomes Reliable



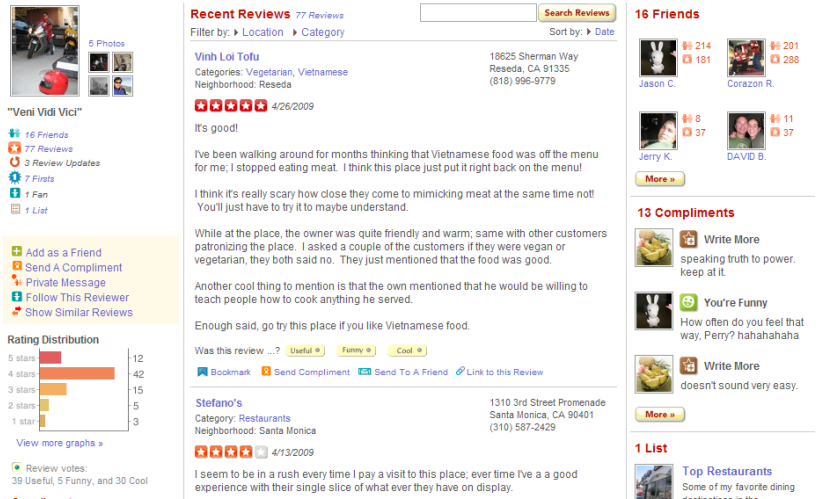
Ratings @ Yelp

Social is local

Efficient Discovery

“Hearsay” becomes Reliable

Marketing becoming less credible



Ratings @ Yelp

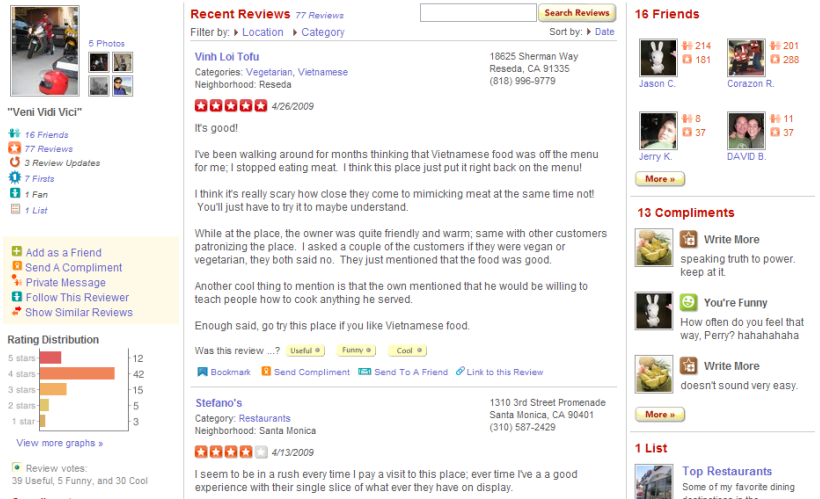
Social is local

Efficient Discovery

“Hearsay” becomes Reliable

Marketing becoming less credible

Many perspectives creates trustworthiness



Ratings @ Yelp

Social is local

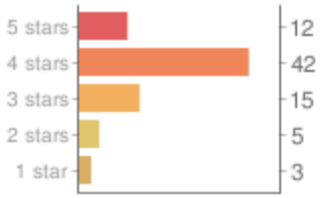
5 Photos

"Veni Vidi Vici"

- 16 Friends
- 77 Reviews
- 3 Review Updates
- 7 Firsts
- 1 Fan
- 1 List

- Add as a Friend
- Send A Compliment
- Private Message
- Follow This Reviewer
- Show Similar Reviews

Rating Distribution



View more graphs »

Review votes:
39 Useful, 5 Funny, and 30 Cool

Recent Reviews *77 Reviews*

Search Reviews

Filter by: Location Category

Sort by: Date

Vinh Loi Tofu
Categories: Vegetarian, Vietnamese
Neighborhood: Reseda

18625 Sherman Way
Reseda, CA 91335
(818) 996-9779

4/26/2009

It's good!

I've been walking around for months thinking that Vietnamese food was off the menu for me; I stopped eating meat. I think this place just put it right back on the menu!

I think it's really scary how close they come to mimicking meat at the same time not! You'll just have to try it to maybe understand.

While at the place, the owner was quite friendly and warm; same with other customers patronizing the place. I asked a couple of the customers if they were vegan or vegetarian, they both said no. They just mentioned that the food was good.

Another cool thing to mention is that the own mentioned that he would be willing to teach people how to cook anything he served.

Enough said, go try this place if you like Vietnamese food.

Was this review ...? Useful Funny Cool

Bookmark Send Compliment Send To A Friend Link to this Review

Stefano's
Category: Restaurants
Neighborhood: Santa Monica

1310 3rd Street Promenade
Santa Monica, CA 90401
(310) 587-2429

4/13/2009

I seem to be in a rush every time I pay a visit to this place; ever time I've a a good experience with their single slice of what ever they have on display.

16 Friends

214
181
Jason C.

201
288
Corazon R.

8
37
Jerry K.

11
37
DAVID B.

More »

13 Compliments

Write More
speaking truth to power. keep at it.

You're Funny
How often do you feel that way, Perry? hahahahaha

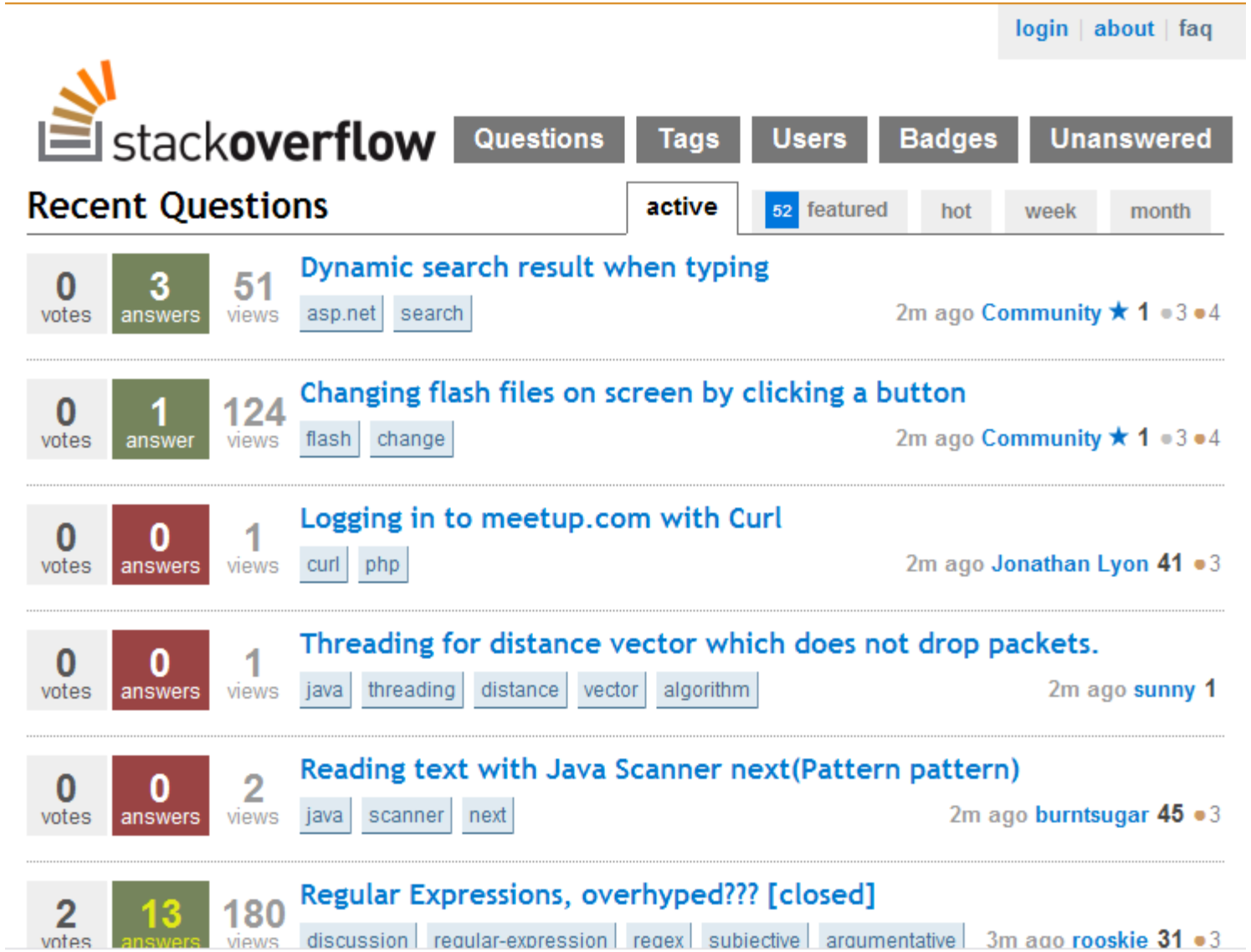
Write More
doesn't sound very easy.

More »

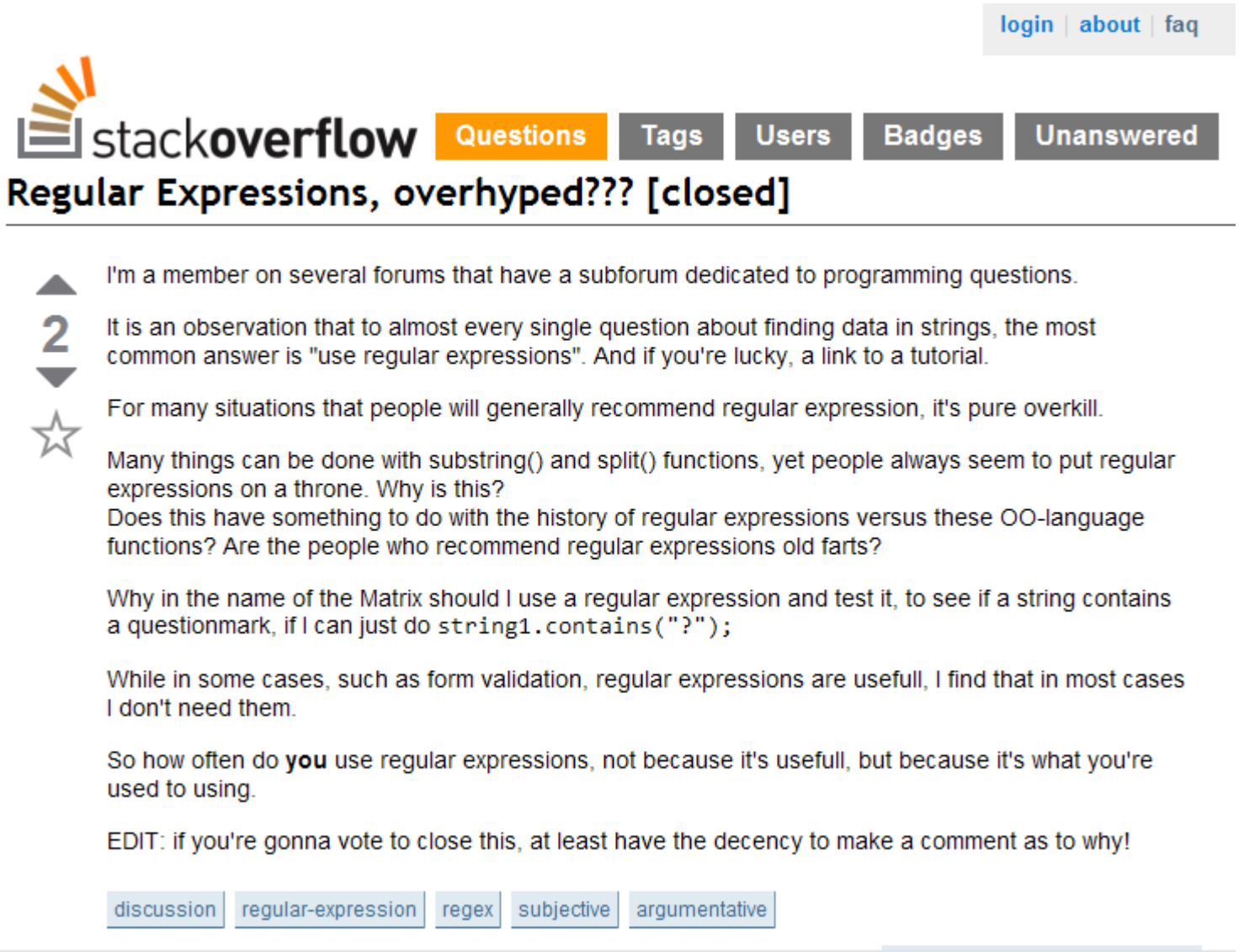
1 List

Top Restaurants
Some of my favorite dining

Problem Solving @ stackoverflow



Problem Solving @ stackoverflow



login | about | faq

stackoverflow Questions Tags Users Badges Unanswered

Regular Expressions, overhyped??? [closed]

2

I'm a member on several forums that have a subforum dedicated to programming questions.

It is an observation that to almost every single question about finding data in strings, the most common answer is "use regular expressions". And if you're lucky, a link to a tutorial.

For many situations that people will generally recommend regular expression, it's pure overkill.

Many things can be done with substring() and split() functions, yet people always seem to put regular expressions on a throne. Why is this?
Does this have something to do with the history of regular expressions versus these OO-language functions? Are the people who recommend regular expressions old farts?

Why in the name of the Matrix should I use a regular expression and test it, to see if a string contains a questionmark, if I can just do `string1.contains("?");`

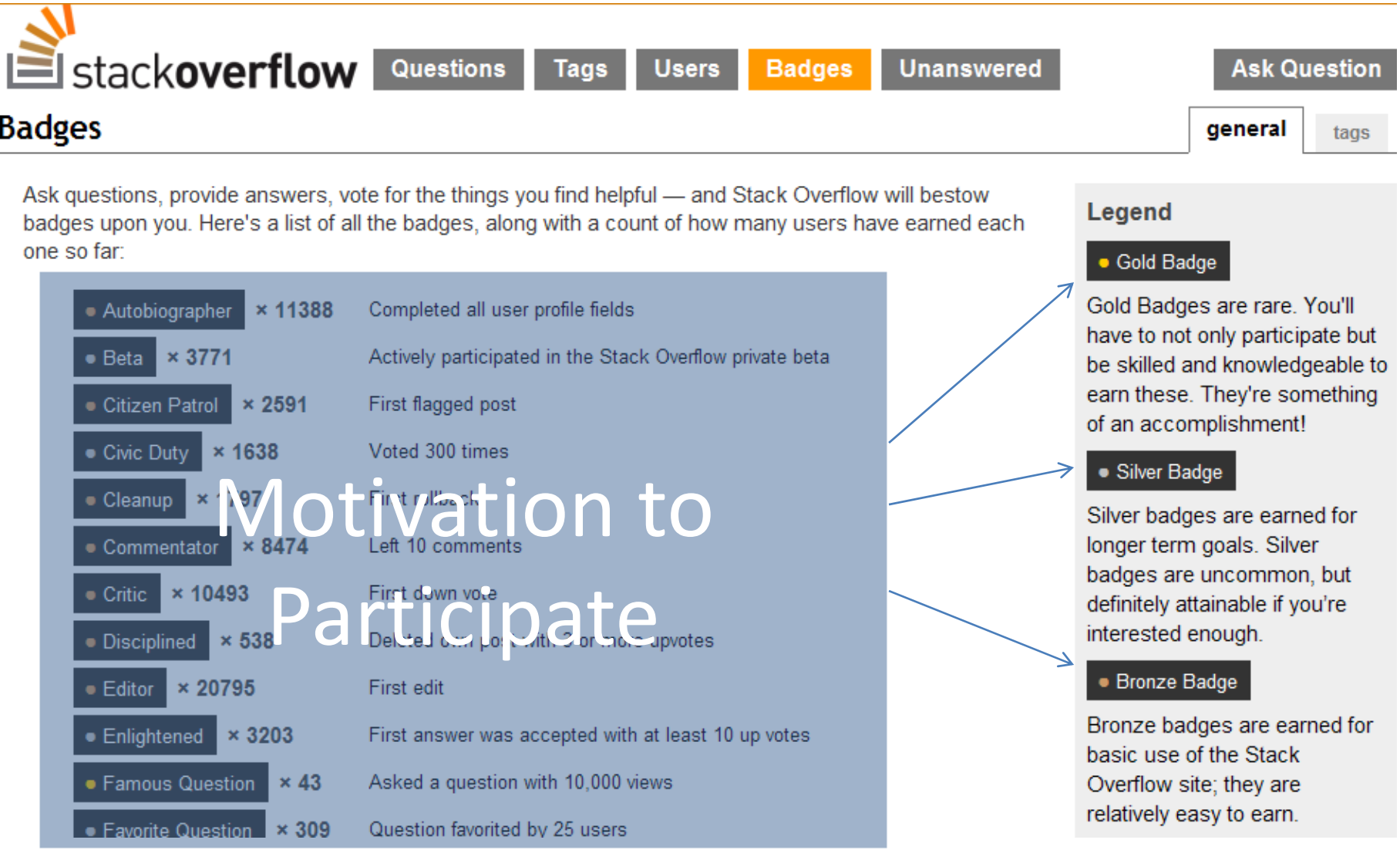
While in some cases, such as form validation, regular expressions are usefull, I find that in most cases I don't need them.

So how often do **you** use regular expressions, not because it's usefull, but because it's what you're used to using.

EDIT: if you're gonna vote to close this, at least have the decency to make a comment as to why!

discussion | regular-expression | regex | subjective | argumentative

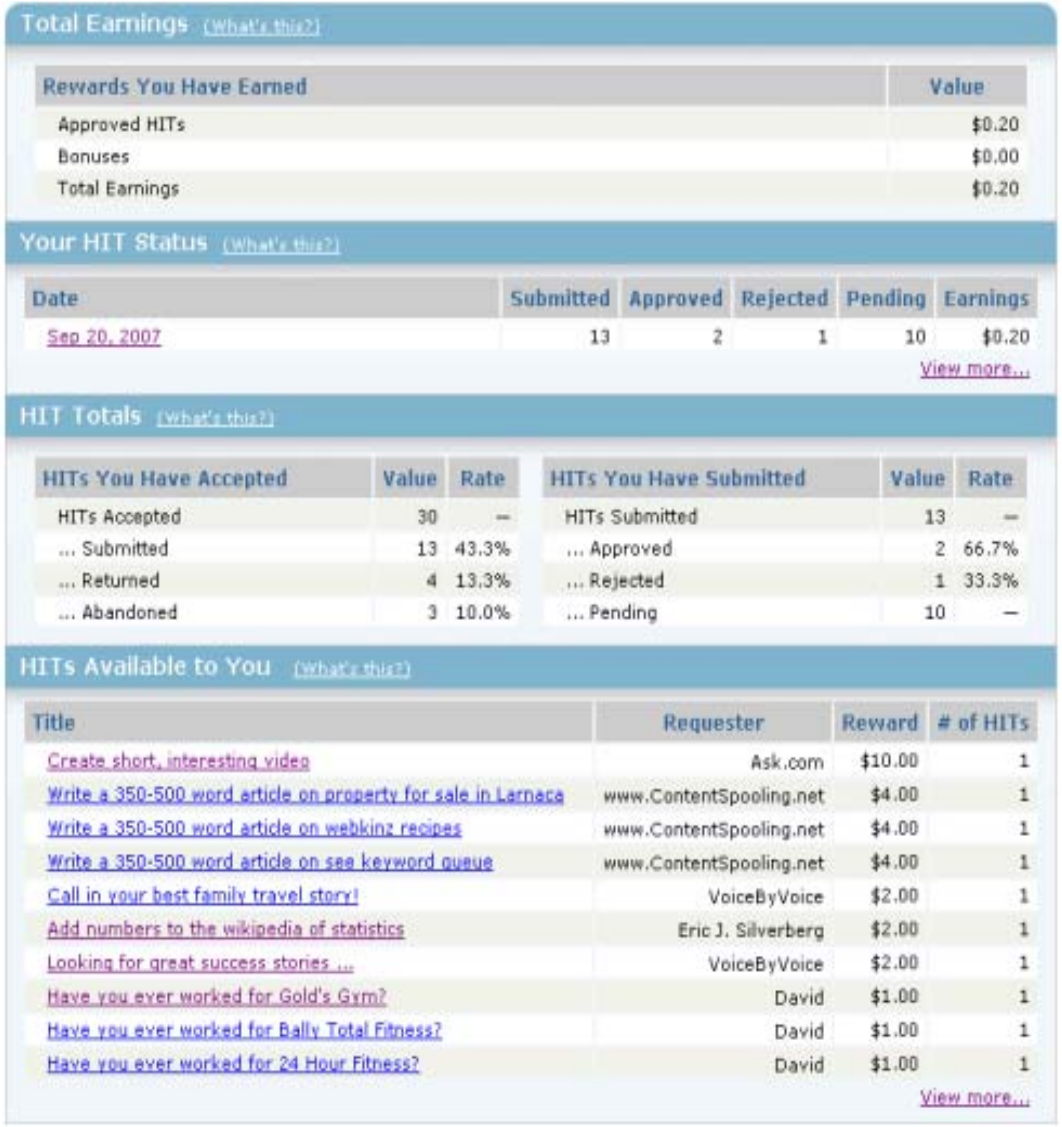
Problem Solving @ stackoverflow



Human Intelligence Tasks (economic)

Mechanical Turk (Amazon) - small tasks $0 - $10
TextEagle (Nathan Eagle) – tasks for mobile phone

Small Tasks @ Mechanical Turk



Total Earnings (What's this?)

Rewards You Have Earned	Value
Approved HITs	$0.20
Bonuses	$0.00
Total Earnings	$0.20

Your HIT Status (What's this?)

Date	Submitted	Approved	Rejected	Pending	Earnings
Sep 20, 2007	13	2	1	10	$0.20

View more...

HIT Totals (What's this?)

HITs You Have Accepted	Value	Rate	HITs You Have Submitted	Value	Rate
HITs Accepted	30	–	HITs Submitted	13	–
... Submitted	13	43.3%	... Approved	2	66.7%
... Returned	4	13.3%	... Rejected	1	33.3%
... Abandoned	3	10.0%	... Pending	10	–

HITs Available to You (What's this?)

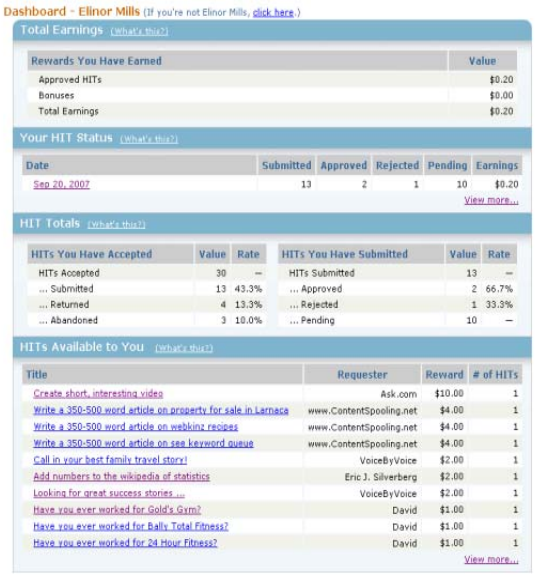
Title	Requester	Reward	# of HITs
Create short, interesting video	Ask.com	$10.00	1
Write a 350-500 word article on property for sale in Larnaca	www.ContentSpooling.net	$4.00	1
Write a 350-500 word article on webkinz recipes	www.ContentSpooling.net	$4.00	1
Write a 350-500 word article on see keyword queue	www.ContentSpooling.net	$4.00	1
Call in your best family travel story!	VoiceByVoice	$2.00	1
Add numbers to the wikipedia of statistics	Eric J. Silverberg	$2.00	1
Looking for great success stories ...	VoiceByVoice	$2.00	1
Have you ever worked for Gold's Gym?	David	$1.00	1
Have you ever worked for Bally Total Fitness?	David	$1.00	1
Have you ever worked for 24 Hour Fitness?	David	$1.00	1

View more...

Small Tasks @ Mechanical Turk

Small Tasks

Time vs. Economics
(does not add up)



Dashboard - Elinor Mills (If you're not Elinor Mills, click here.)

Total Earnings (What's this?)

Rewards You Have Earned	Value
Approved HITs	$0.20
Bonuses	$0.00
Total Earnings	$0.20

Your HIT Status (What's this?)

Date	Submitted	Approved	Rejected	Pending	Earnings
Sep 20, 2007	13	2	1	10	$0.20

View more...

HIT Totals (What's this?)

HITs You Have Accepted	Value	Rate
HITs Accepted	30	–
... Submitted	13	43.3%
... Returned	4	13.3%
... Abandoned	3	10.0%

HITs You Have Submitted	Value	Rate
HITs Submitted	13	–
... Approved	2	66.7%
... Rejected	1	33.3%
... Pending	10	–

HITs Available to You (What's this?)

Title	Requester	Reward	# of HITs
Create short, interesting video	Ask.com	$10.00	1
Write a 350-500 word article on property for sale in Larnaca	www.ContentSpooling.net	$4.00	1
Write a 350-500 word article on webkinz recipes	www.ContentSpooling.net	$4.00	1
Write a 350-500 word article on sea keyword queue	www.ContentSpooling.net	$4.00	1
Call in your best family travel story!	VoiceByVoice	$2.00	1
Add numbers to the wikipedia of statistics	Eric J. Silverberg	$2.00	1
Looking for great success stories ...	VoiceByVoice	$2.00	1
Have you ever worked for Gold's Gym?	David	$1.00	1
Have you ever worked for Bally Total Fitness?	David	$1.00	1
Have you ever worked for 24 Hour Fitness?	David	$1.00	1

View more...

Small Tasks @ TextEagle

Mobile Human Computation in Africa

TextEagle is a…"system enabling the 3 billion mobile phone subscribers living in the developing world to earn small amounts of money by completing short, SMS-based tasks."

Nathan Eagle
Research Scientist, MIT

Source: http://web.media.mit.edu/~nathan/

Small Tasks @ TextEagle

Translation/Transcription Services

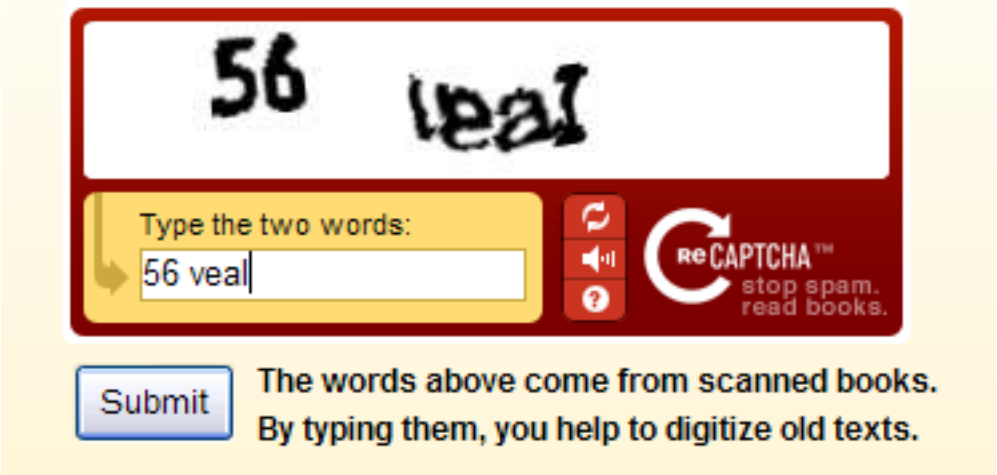
Question: Translate the phrase "Address Book" into Giriama.
Question: Transcribe the following audio clip from a New York hospital.

Also Citizen Journalism

Human Intelligence Tasks (tacit)

User may not be aware of how tasks are utilized.

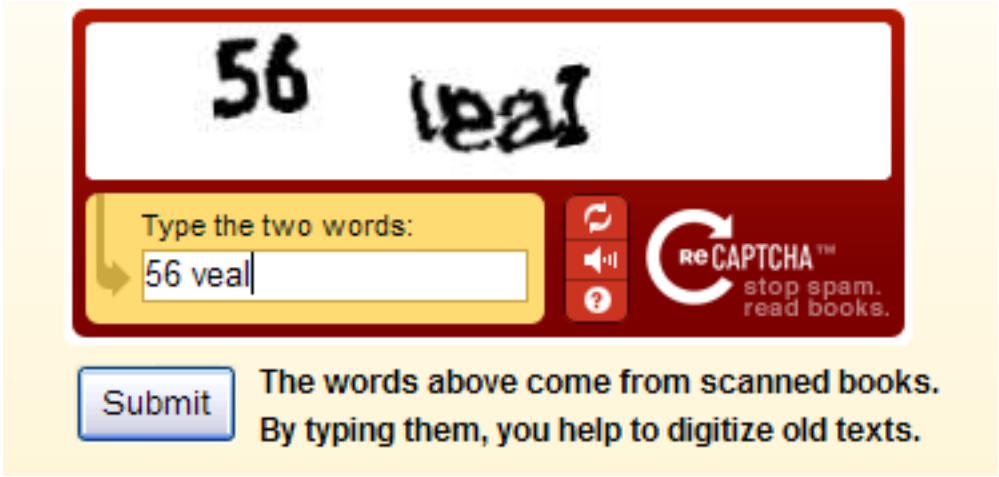
reCAPTCHA (based on CAPTCHA)
ESP Game (now Google Image Labeler)



CAPTCHA:

“Are you human?”

(CAPTCHA: Developed by Louis Von Ahn)

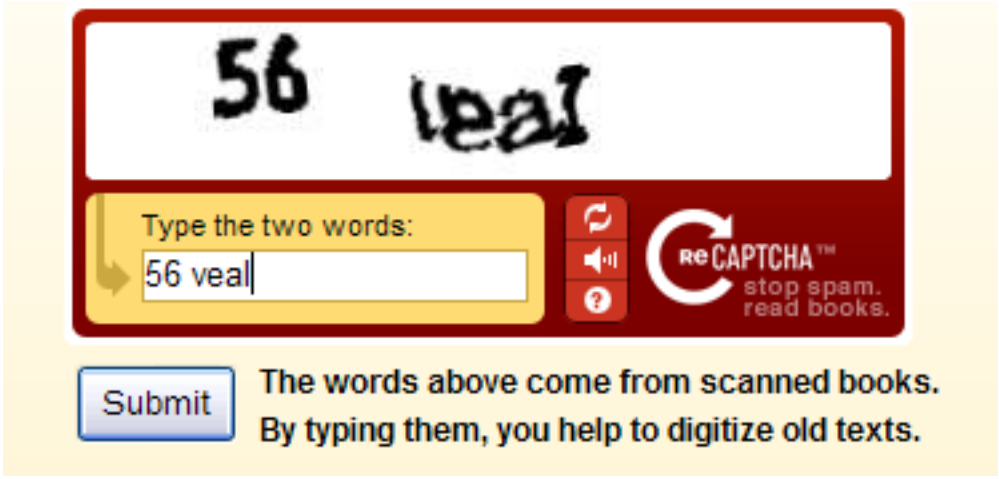


CAPTCHA:

(Completely Automated Public Turing Test
To Tell Computers and Humans Apart)

Security based on human perception
Turing Test administered by AI

(CAPTCHA: Developed by Louis Von Ahn)

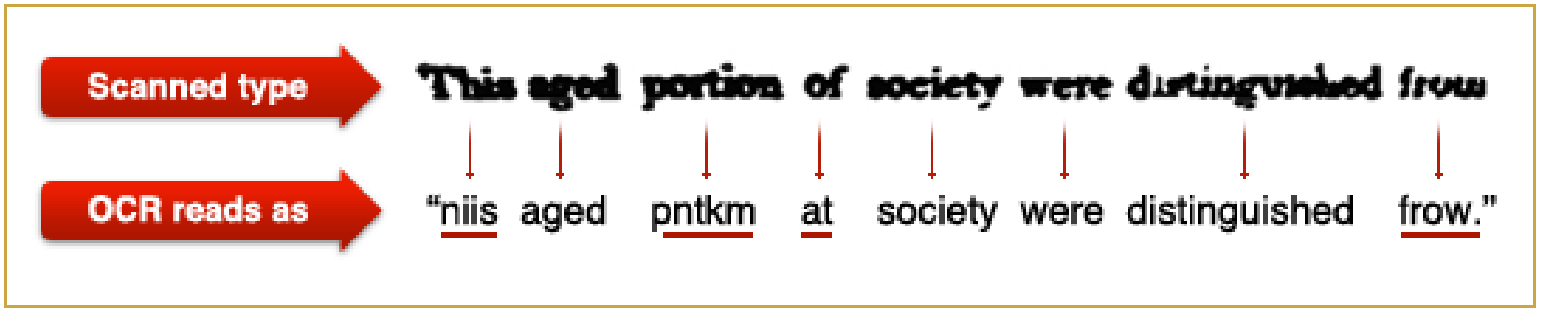


reCAPTCHA

Security + Failed OCR = Opportunity

(CAPTCHA: Developed by Louis Von Ahn)

reCAPTCHA



Currently being used to help scan books for **Internet Archive** and old editions of the **New York Times**.

(graphic - http://recaptcha.net/learnmore.html)

reCAPTCHA

Bots getting better at deciphering the CAPTCHAs

(graphic - http://recaptcha.net/learnmore.html)

reCAPTCHA

Bots getting better at deciphering the CAPTCHAs

Creates feedback loop - Good thing, since it means OCR gets more precise at same time

(graphic - http://recaptcha.net/learnmore.html)

ESP Game (now Google Image Labeler)

2 individuals match = high probability of reliable result

With Human Computation…

…computing becomes the coordinating force between many individuals and intelligent or perceptual tasks.

Human Computation Models

Social

Tagging (Flickr) and Comments (Amazon)
Ratings (Yelp)
Problem Solving (StackOverflow)

Economic

Mechanical Turk (Amazon) - small tasks $0 - $10
TextEagle (Nathan Eagle) – tasks for mobile phone

Tacit

reCAPTCHA (based on CAPTCHA)
ESP Game (now Google Image Labeler)

How can you create a platform for capturing social knowledge?

Shared Knowledge Production

Public platforms – ex. Flickr

Develop platform – ex. Simul8 Model

UCLA Library **Simul8** Group Model

Listen to users – experience paradigm
See user mode, lifestyle, aesthetic
Engage in their mode – playful and experimental dev
Fast prototyping with high level programming
Student developers and designers
Creative Commons License

One Question:

Kevin Rundblad
kevinrundblad@library.ucla.edu

One Question:

Are You Human?

Kevin Rundblad
kevinrundblad@library.ucla.edu