# 圖書資訊系統的未來展望

報告者:陳亞寧

陳亞寧、黃鴻珠 淡江大學資訊與圖書館學系

21 Feb., 2014

### Outline

- Introduction
- Definition
- Trends
- Issues
- Conclusion

### Introduction



#### Definition-1

- Terminologies
  - Library Automation
  - Integrated Library System
  - Library Management System
- ILS + ERM(Dempsey, 2008)
- ODLIS
  - Acquisitions, cataloging and authority control, serials control, circulation and inventory, and interlibrary loan and document delivery.
  - Link resolver
  - Portal and metasearch interfaces
  - E-resource management module
  - WWW

#### Definition-2

- Breeding, 2012
  - Library services platforms
  - ILS
  - Discovery products and services
  - Federated search
  - Digital library management system
  - Electronic resource management support
  - Institutional repository
  - OpenURL linking application
  - Resource sharing systems
  - RFID support
  - Archives management

Source:

http://www.thedigitalshift.com/2012/03/ils/automation-marketplace-2012-the-complete-survey-data/

#### Definition-3

- Scope
  - Single v.s. multiple systems
  - Internal v.s. external
- Information carriers
  - Physical v.s. digital carriers such as monographs, books, journals, and so forth.
  - Genres: bibliographies, TOCs, full-texts and so forth.
- Object identifier
  - Traditional: Accession no, barcode, ISBN, ISSN, etc.
  - New: URL/URI, DOI, RFID, etc.

### Trends-1

- ICTs
  - Mobile
  - Visualization
  - Open source
  - Clouds: SaaS, PaaS & IaaS

#### Trends-2

#### Google

- Books
- Maps
- Scholar and Citation
- iGoogle
- Google Search
  - Simple query box
  - Unified search & representation
- Google Generation

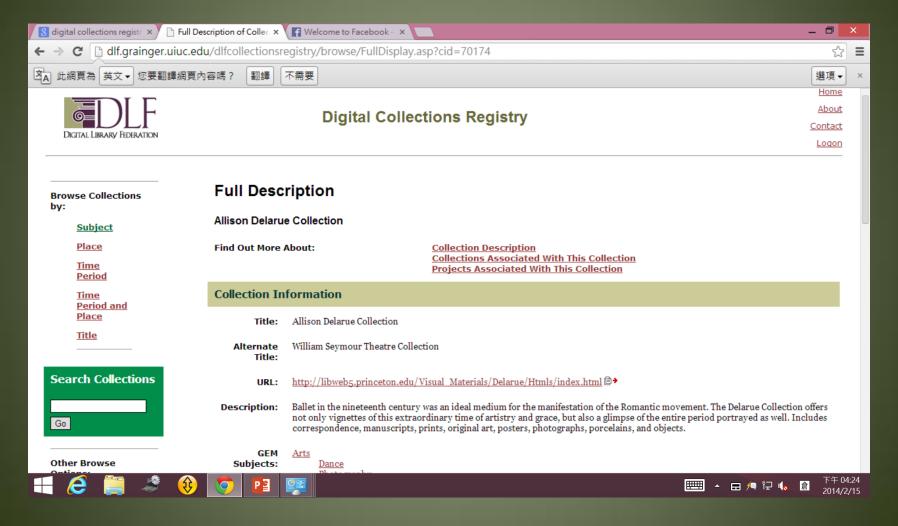
#### Amazon.com

- Bibliography
- TOCs
- Cover image
- Full-texts
- Comments
- Rating
- Recommendation
- Unified search & representation

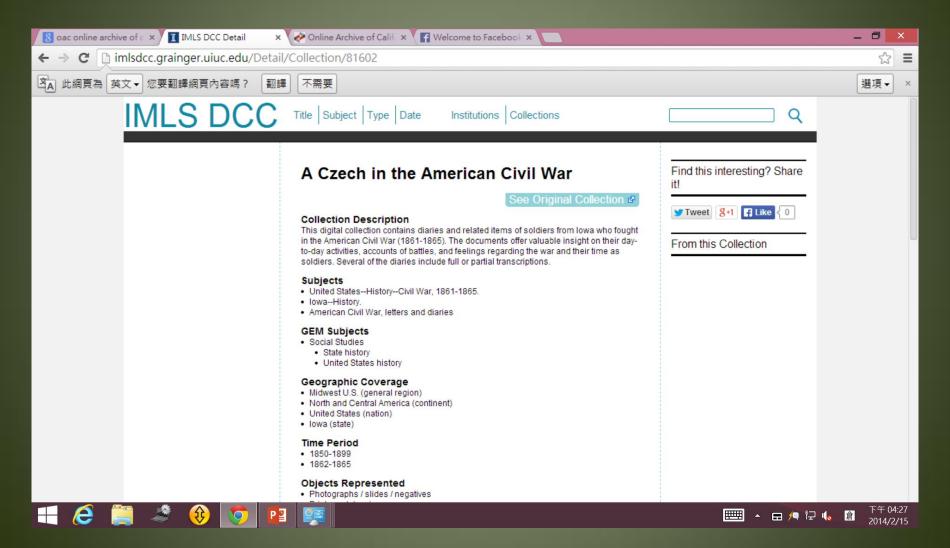
#### Trends-3.1

- Digitization and Open
  - Digitization and born digital (digital libraries)
    - Journals, monographs, learning objects, etc.
  - Raw data and datasets (eScience, eResearch)
- Structure and Organization-DLibraries
  - Items
  - Series
  - Collections

# DLF's Digital Collections Registry



#### IMLS's DCC



#### Trends-3.2

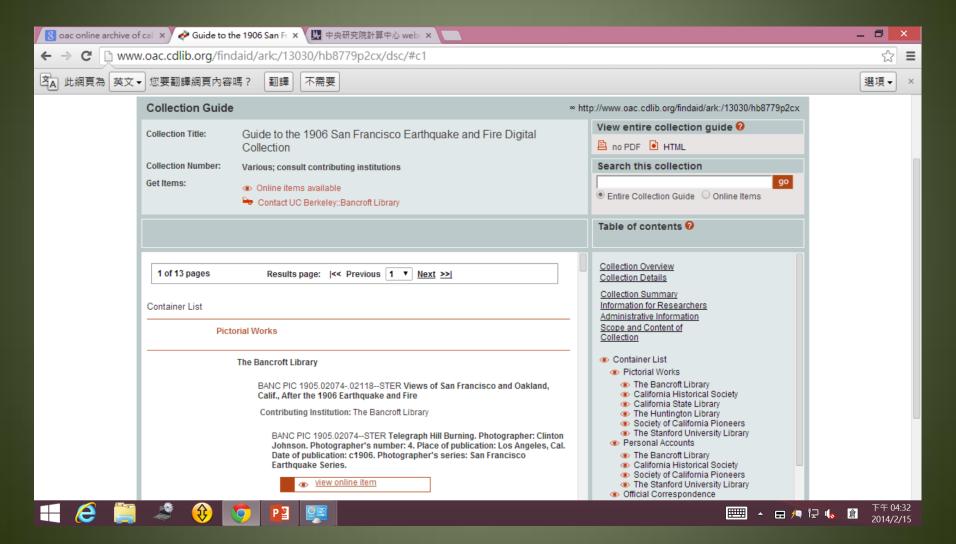
## **Structure and Organization- DArchives**

- Fonds
- Series
- Files
- Items

#### **Structure and OrganizationeLearning**

- Content Aggregation
- Sharable Content Object
- Asset

#### OAC



# eLearning-1



Source: 教育部數位教學資源入口網, https://isp.moe.edu.tw/

## eLearning-2



Source: 教育部數位教學資源入口網, https://isp.moe.edu.tw/

## eLearning-3

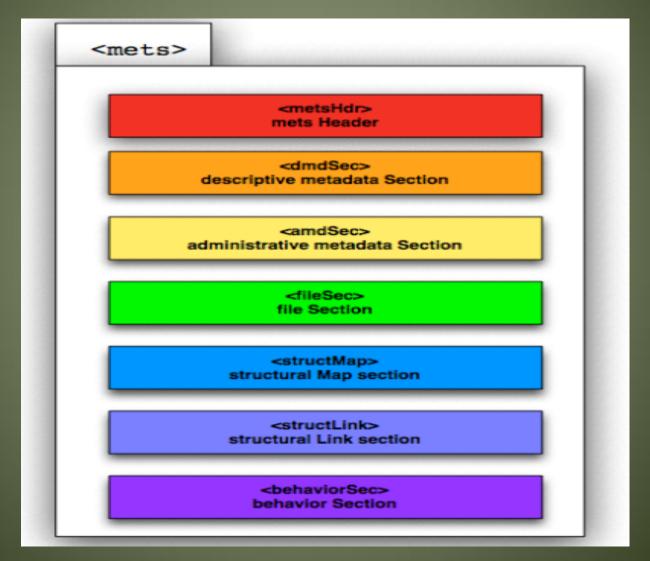


Source: 教育部數位教學資源入口網, https://isp.moe.edu.tw/

#### Trends-3.3

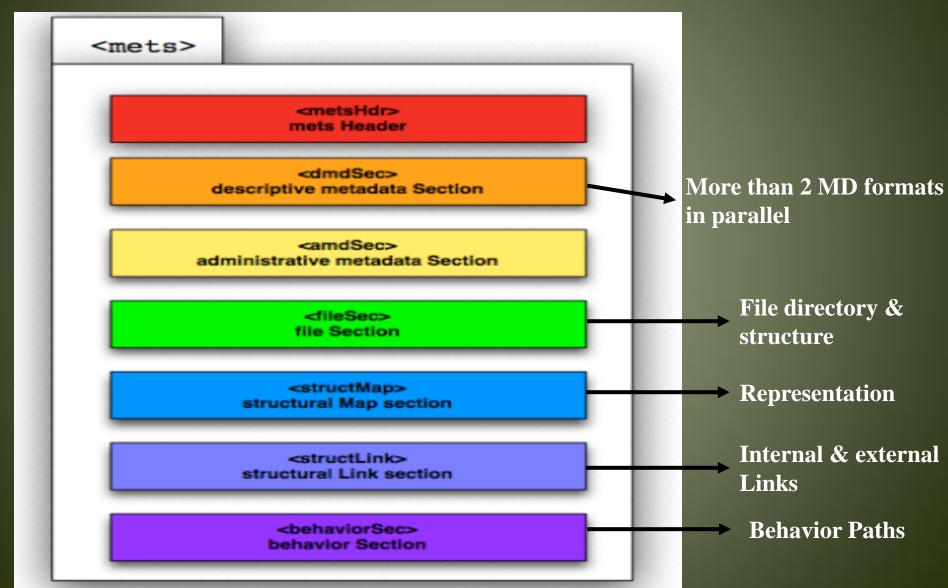
- Interoperability and Integration
  - Harvest: OAI-PMH
  - Crosswalking or mapping
  - Metadata Encoding and Transmission Standard

## METS-1



Source: http://www.loc.gov/standards/mets/METSPrimerRevised.pdf

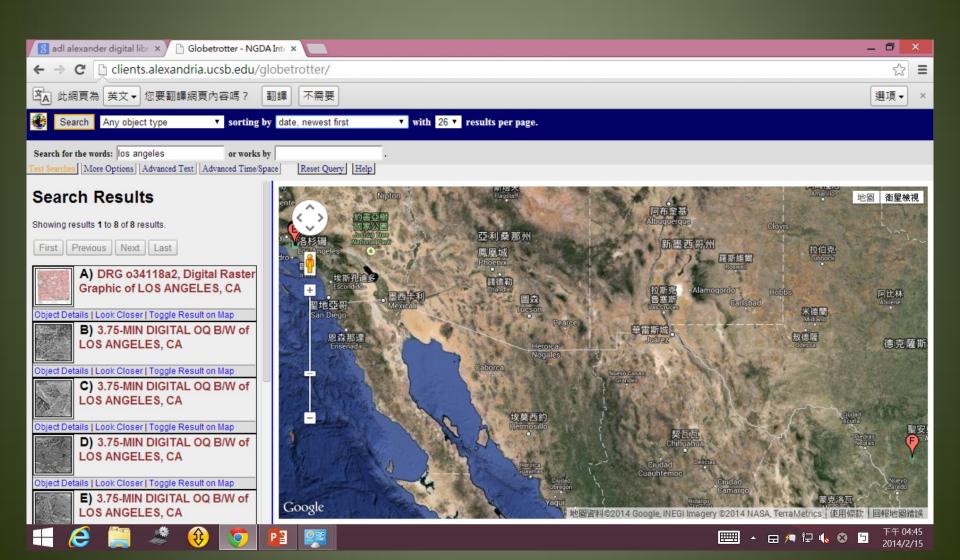
### METS-2



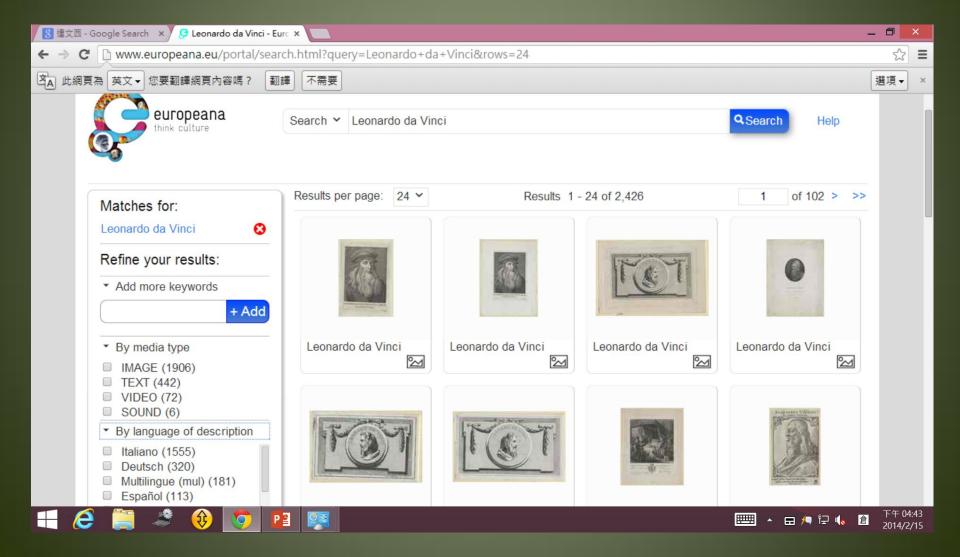
#### Trends-3.4

- Representation
  - GIS orientation: Temporal and Spatial
  - Ontology orientation: European

### ADL



## Europeana



### Trend-4

- Web 2.0
  - Mashup and remix
  - Collective intelligence
  - Social network

### Trend-5

- Semantic web and Linked Data
  - HTTP
  - URI
  - RDF
  - Link(ontology)

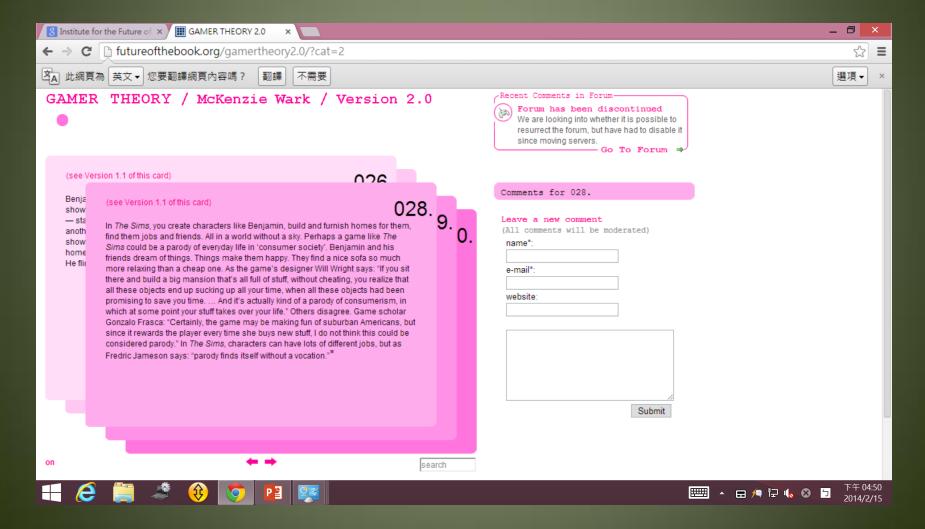
# Issues-1 Creation and Type of information resources

- Acquisition: internal vs. external, and single vs. multiple sources
- Object: physical v.s. digital (digitization and born digital)
- Genre
  - Full-texts: such as books, journals.
  - Ref.: such as abstracts, indexes, and citation.
  - Web 2.0: Images, videos, and tags.
  - Online games
  - Raw data and datasets
  - Behaviors
    - Business Intelligence: such as Amazon
    - Collective Intelligence: Web 2.0 and Books provided by Institute for the Future of the Book

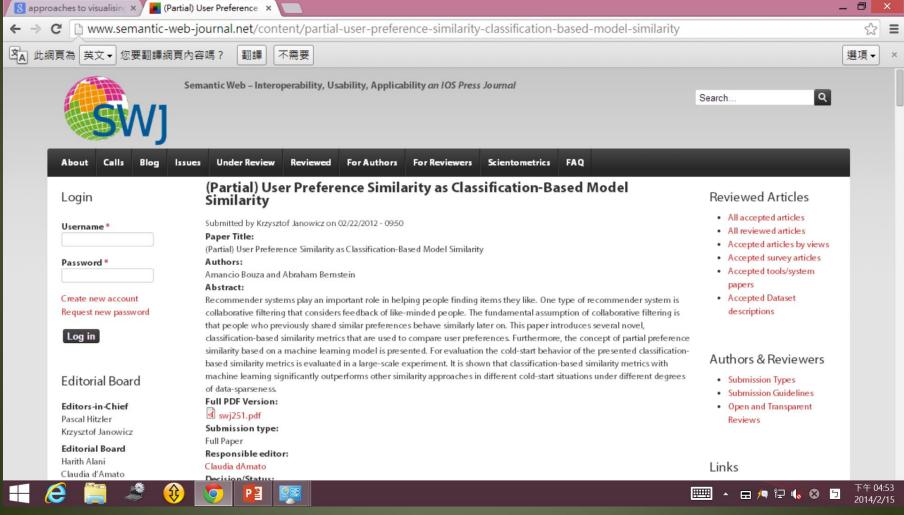
#### Institute for the Future of the Book-1



#### Institute for the Future of the Book-2



### SWJ-1



Source: http://www.semantic-web-journal.net/content/partial-user-preference-similarity-classification-based-model-similarity

#### SWJ-2



#### Reviews:

#### Review 1 by Vojtech Svatek

翻譯

The paper presents a method and extensive experiment in using machine learning for user-similarity-based recommendation.

My main problem with the current version of the paper is (following the experience of a previous reviewer) the cumbersome way in which the authors approach is explained. I had to read most of the text 2-3 times, and finally I would say that it can be summarized as a straightforward sequence of activities from the ML point of view:

- For each user, a classifier is learned from labeled examples.
- 2) Unlabeled examples are labeled by this classifier.

不需要

- 3) Based on the classified examples (hypothetical ratings) the similarity of users is computed.
- 4) When recommending items to a certain user, the ratings of similar users are taken into account.

I dont see any point in the lengthy formal treatment of the problem, with 24 numbered formulae, where most statements in both symbols and text are inherently trivial but explained in complex terms. On the other hand I still miss a well worked-out running example, which was demanded by the previous reviewers.

The formal part itself has quite a few flaws, for example:

- In (4) c is introduced as if it were invariant wrt. q, which is definitely not the case.
- In the itemlist between (5) and (6), the set of items Gi is certainly not part of experience of a user by itself
- (7) is trivial any probability in the world must be smaller or equal than 1!
- The hypothesis space H is only mentioned once what does it mean?
- The left side of equation (20) doesnt contain s why?

I dont see much novelty in the probabilistic nor correlation-based similarity. Furthermore, the claim that the unified rated item set approach in the former is a good approximation is not entirely convincing. It may give too much weight to the relative size of the common rated item set. For example, if the two users rated 55 objects each, 10 of them are shared, and all of these are rated same, the value of the metric is only 0,1. In contrast, if the two users rated 15 objects each, 10 of them are shared, and 5 (i.e. only half) of these are rated same, the value is 0,4.

Even simple notions such as interval rating and ratio rating are explained in an absolutely unclear way.

An interesting observation is the directionality of partial similarity, meaning that e.g. similarity of user A and user B is measured wrt, a category C considered by user A. However, its impact is not studied further.

Another problem that remains is the limited relevance to semantic web topics. The paper now mentions the YouLike ontology and shows an RDF fragment associated with it, however the ontology is not explained and there is even no link to its code (it was only published as part of the first author's thesis?).

- SWJ online issues at IOS Press
- SWJ Scientometrics Portal
- SWJ at Twitter
- SWJ at DBLP
- · Book series 'Studies on the Semantic Web'
- · Linked Data Hub

#### Recent blog posts

- · Call for papers: Special Issue of the Semantic Web Journal on Scalable Semantic Graph Systems
- Scopus listing
- · A Linked-Data-driven and Semantically-enabled Journal Portal for Scientometrics
- Semantic Web Journal Outstanding Reviewer and Outstanding Paper Awards 2013
- Some Visual Analytics Results for ISWC 2013
- SWJ @ ICLP2013
- Call for papers: Special issue of the Semantic Web Journal on Web Reasoning and Rule

#### Syndicate

Kunal Sengupta















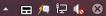


















選項▼

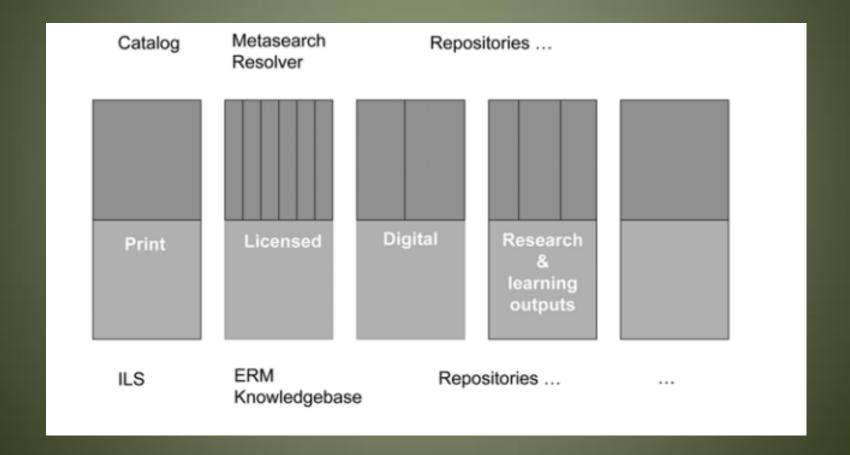
# Issues-2 Granularity and Identification

- Divergence and Convergence
  - Remix and Composite/Compound Object
  - FB example
- Digital Libraries/Archives and eLearning
- Identifier
  - Object: ARK, DOI, ISBN, ISSN, PURL, & URI/URL, etc.
  - Record: NBN, OAI, arXiv ID
  - Subject: LCSH
  - Person: VIAF, ResearchID, ORCID
  - DBpedia: knowledge entry/unit
  - ADL: geographic name

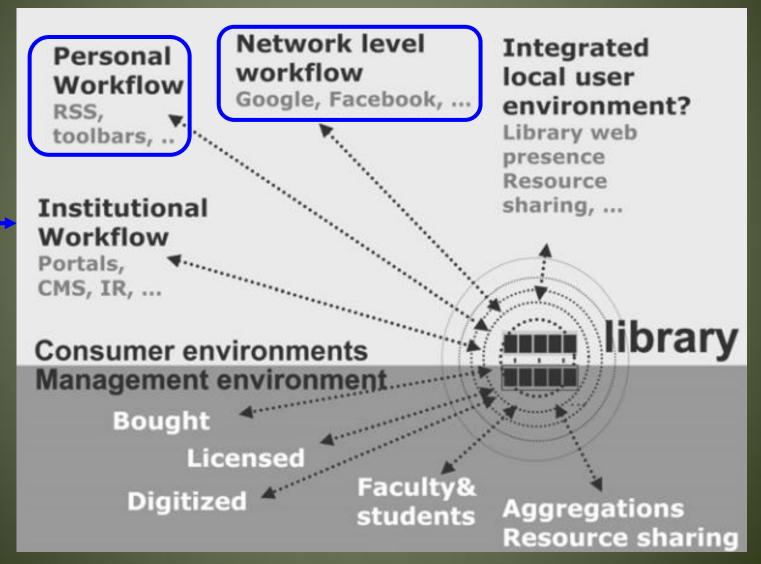
# Issues-3 Organization and Management

- Unified organization and management
  - Amazon and Google
  - Unified Resource Management: BL and Ex Libris
- Dempsey's viewpoints

# Dempsey's Viewpoints-1



## Dempsey's Viewpoints-2



Source: http://www.oclc.org/research/publications/library/2008/dempsey-portal.pdf

# Issues-4 Association

- BI orientation: such as Amazon
- CI orientation: such as FB through by FOAF
- Linked Data-based orientation
- Ontology-based orientation
  - FRBR, FRSAD, and FRAD(Libraries)
  - Provenance and Fonds(Archives)
  - CIDOC CRM(Museums)
  - FRBRoo(Libraries and Museums)
  - W3C: RDF, OWL, URI, etc.
  - Others

# Issues-5 Strategic Alliance

- Homogeneous
  - Hathitrust
  - Open Content Alliance
  - Open Library
  - Evergreen
  - Koha
- Heterogeneous
  - LibraryThing, LC, and Bowker
  - Sierra/OLE/WorldShare Management and Publishers
  - LAMs

# Issues-6 Automation

- Information Flow
  - Business Process/Workflow
  - Metadata: AMeGA
  - Full-texts: Million Book Project
  - Behaviors

### Conclusion

	Intangible	Tangible
Entity	Information-as- knowledge Knowledge	Information-as-thing Data, document, recorded knowledge
Process	Information-as-process Becoming informed	Information Processing Data processing, document processing, knowledge engineering

Source: Buckland, M. (1991). Information and Information Systems. New York: Praeger.

Thanks for your attention.