Why Your Library Doesn't Do What You Want It To

Stuart L. Weibel

Senior Research Scientist

OCLC Research

weibel@oclc.org

http://weibel-lines.typepad.com/

- OCLC – Online Computer Library Center
  - International Not-for-profit helping libraries remove redundancy from technical processing
  - Shared Cataloging, ILL, reference…
  - Open WorldCat (60+ million records, a billion holdings)
  - Find in a Library (Yahoo, Google, and others)
- OCLC Research
  - The CORE project (Chemistry Online Retrieval)

OC
LC

CORE Project

- ACS, Chemical Abstracts, Bellcore, Cornell, and OCLC

- First generation digital library project

- Pre-Web

- SGML Markup and page images

- On-the-fly formatting on Windows, Unix, and Macintosh platforms

- Lego blocks are esthetic in design and exceptional in engineering
- Design and manufacturing tolerances approach those of internal combustion engines
- Backwards and forwards interoperability
- Sound conceptual model for information systems:
  - standardized interfaces
  - modular design
  - extensibility

OCLC

The darker side of the Lego Metaphor

'Twas the day after Christmas

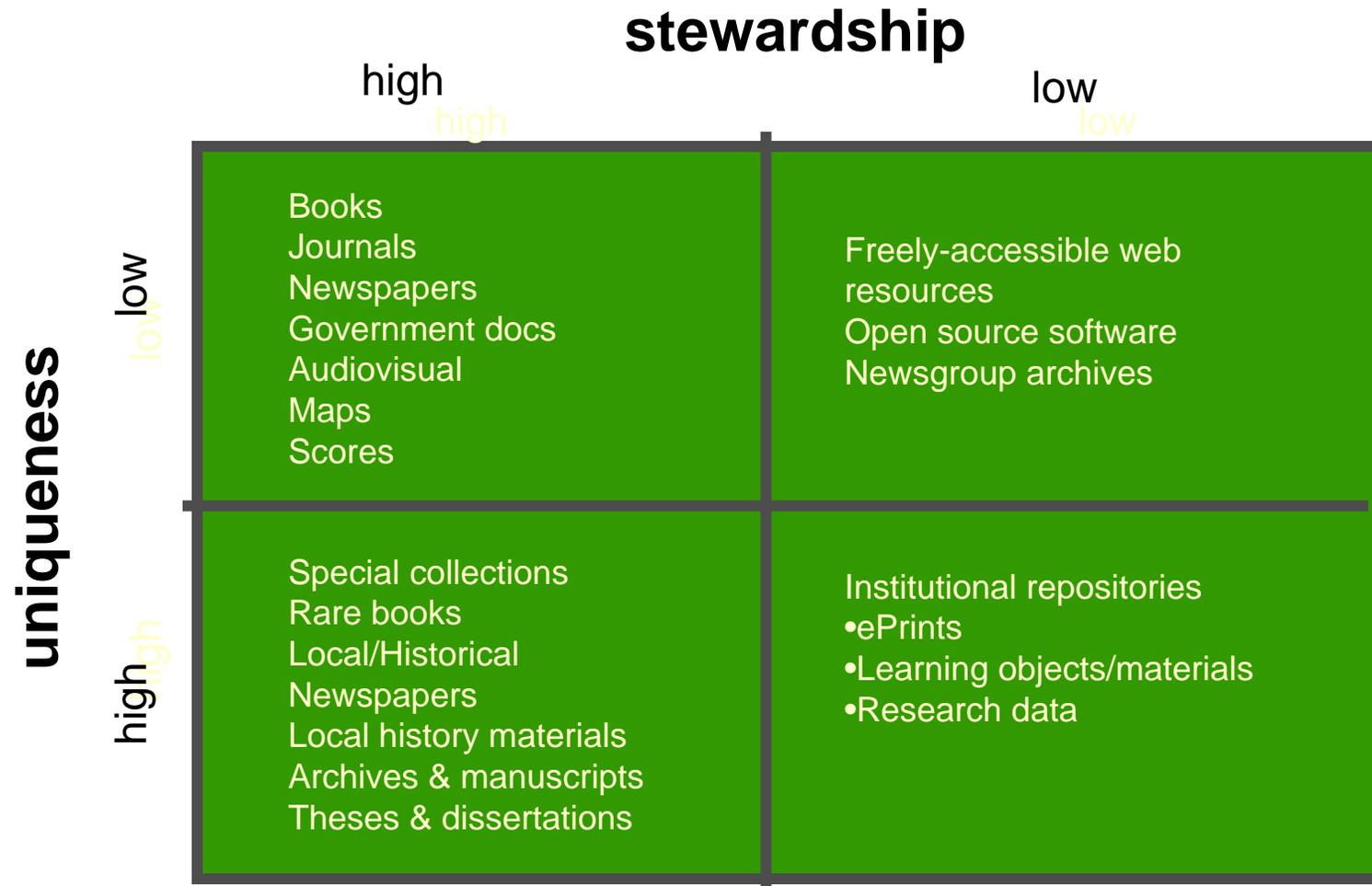and all through the house

not a creature was stirring

for fear of treading on one of 700 sharp and pointy specialized parts of unidentifiable provenance and function (in the absence of the context of the box and the assembly directions, which were burned in the fireplace along with the wrapping paper)

OCLC

Formats and resource types are increasing

- Many new content formats, each has a particular constellation of characteristics having to do with
  - discovery,
  - access,
  - rendering,
  - IPR,
  - provenance,
  - Stewardship and preservation

The Resource Grid

Lorcan Dempsey

stewardship

	high	low
low	Books Journals Newspapers Government docs Audiovisual Maps Scores	Freely-accessible web resources Open source software Newsgroup archives
high	Special collections Rare books Local/Historical Newspapers Local history materials Archives & manuscripts Theses & dissertations	Institutional repositories •ePrints •Learning objects/materials •Research data

uniqueness

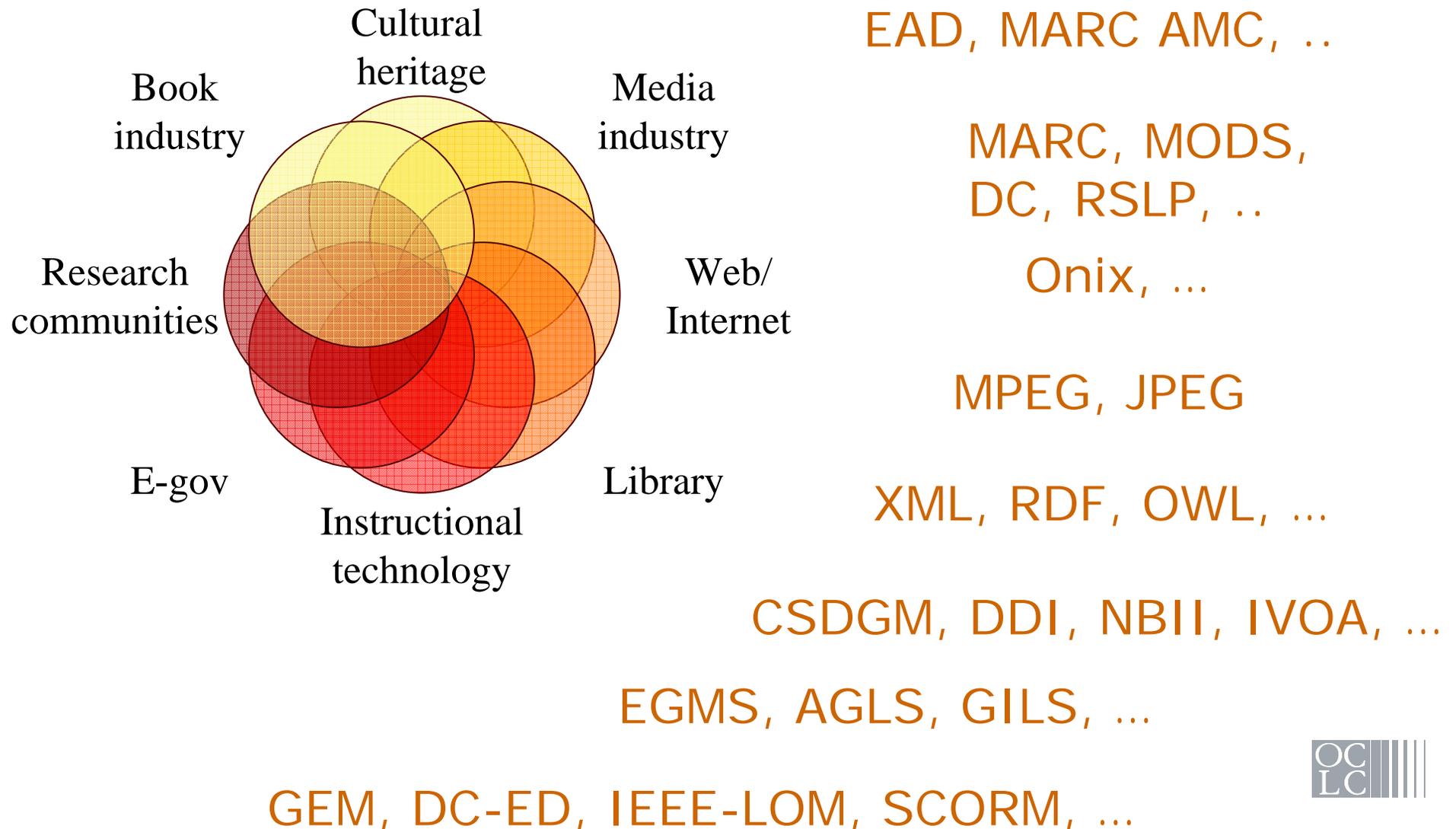OCLC

Lots of Stuff
Uncertain Roles

- Publishers, libraries, archives, professional societies, and IT departments are all uncertain of their once-stable positions in the information supply chain.

- The Web has turned things upside down and created both opportunity and risk for all concerned.

- Everyone is looking for ways to consolidate their own precarious position and wondering if they might co-opt some part of someone else's in the bargain

The distributed model of the Web brings us all into the Internet Commons

- Information silos are breaking down, but innovation on one hand and the challenge of interoperability on the other have spawned many standards in many standards arenas
- They don't all work seamlessly (alert the media!)

Interoperability is a difficult objective

- Traditional views of interoperability depend on **pre-coordinated agreements** about every aspect of syntax, structure, and semantics - everyone spoke the same language
- In the Internet Commons, this degree of **coordination is hard to come by**
- Schema declarations allow applications to search, access, display, and link metadata and content from arbitrary **schemas that share declaration conventions**
- **Registries are the online dictionaries** that support such applications
- Much of this remains in the **research domain** (what if the book had been invented before the book shelf?)

OCLC

The Solutions are a major part of the problem

- Metadata languages
- Controlled vocabularies
- Encoding conventions
- Metadata registries (online metadata dictionaries)
- Schema languages
- Rights Management Languages
- Knowledge representation languages
- Ontologies
- … and more ….

Business Models in flux

- Scholarly publishing: a slow motion train wreck
- Because of shifting roles, and a *presumption* of greater efficiency, consolidation is expected to take place.

Some Ponderables

- Do publishers add sufficient value to retain their place (and revenue) in the information supply chain?
- Who will manage the IPR and how?
- Are libraries useful as aggregators and distributors of content?
- What will the role of the Professional Society become?
- Will authors or institutions self-archive?
- Who will look after the stuff when it is 5 years or 50 years or 100 years old?

Follow the money

- **Chemistry is the most costly subject for journal subscriptions average price per journal:**
  - **chemistry $2,868**
  - **physics $2,719**
  - **engineering $1,683**
  - **biology $1,494**
  - **technology $1,460**
  - **math and $1,267
    computer
    science)**

**Periodical Price Survey 2005,
Library Journal, April 15, 2005**

Heard on the street in Libraryville…

- "…mismatch between the growing cost of scholarly literature and our ability to buy it"
- "growing economic crisis in scholarly communications, and chemistry is at the forefront"
- "We cannot continue to rob one discipline for another"
- "faculty are long overdue in understanding the economics of scholarly publishing"

The slow train wreck

- Publishing is growing
- Costs are rising
- Demand is rising
- $upport is flat or declining
- Distributed networking promotes 'disintermediation'

- Conflicting trends place stakeholders at odds; change is inevitable (and inevitably painful)

OCLC

What about other models?

- Open access journals
- OAI repositories
- Greater diversity in publishing vehicles to support expanding scholarship
- Changes in academic recognition and reward systems
- Hybrid models incorporating managed repositories, blogs, *peer-to-peer review…?*

Open Access journals as an alternative to traditional scholarly publishing?

- "…impact factors prove that BioMed Central's Open Access journals are high quality and widely read and cited.

- Journals published by BioMed Central … compare well with equivalent subscription titles… with five titles in the top five of their specialty.

  http://www.biomedcentral.com/info/about/pr-releases?pr=20050623

Peer-to-peer Review publishing?

- Self-publish to a peer-to-peer journal (*ala* epistulary journals such as initiated scholarship)
- "Subscription" to the journal constitutes an established identity with blog-like response rights
- A paper accrues legitimacy by attracting reviews
- An author has revision rights that might be continuous, or might be discrete and limited
- Stewardship of the journal includes management of data stores, the article, supporting infrastructure

Summing up…

- The Web destabilizes the balance among libraries, publishers, professional societies, content creators, and users (and the institutions that support these entities)

- Throwing technology at the problem helps in some respects and complicates in others

- The *new equilibrium* will be catalyzed by actors who recognize how to bring the greatest benefits to the stakeholders – the researchers and the consumers of that research

Many Thanks…

…to Lorrin Garson for asking me to be here

Lorrin helped make the American Chemical Society a leader in this most tumultuous of times

Whatever the *New Equilibrium* turns out to be, he will have played a memorable part in bringing it to pass

OCLC