

Libraries as Read/Write Services

O C LC Research Distinguished Seminar Series

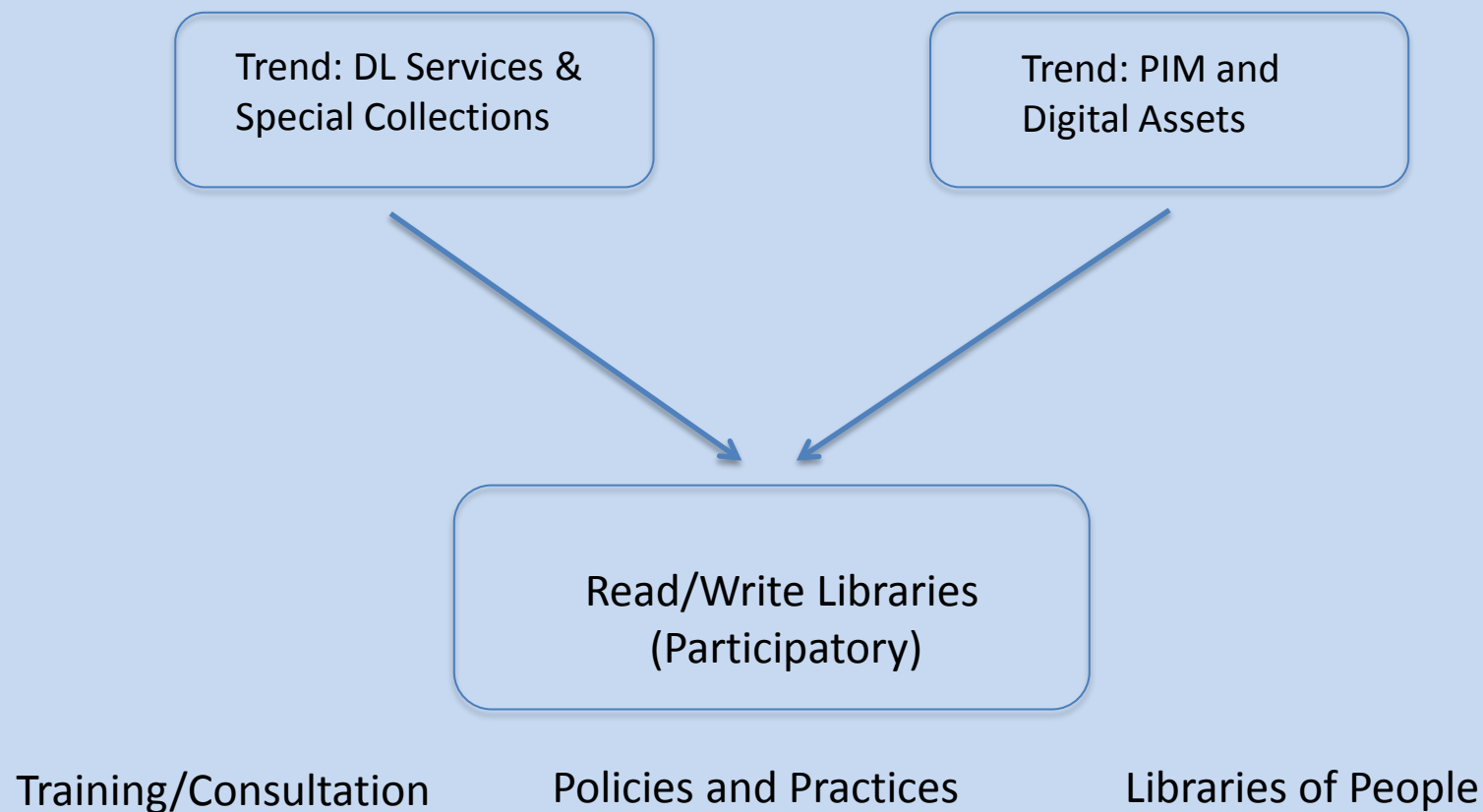
Gary Marchionini, PhD
University of North Carolina at Chapel Hill

OCLC

Dublin, Ohio
February 21, 2012



In a Nutshell





The Institutional DL Milieu

- 20 years of DL research and development
- Hybrid libraries
 - Academic, research
 - Publishing (e.g., ACM DL)
 - Secondary (e.g., OCLC, Research Index, ISI, institutional repositories)
- Born Digital
 - Special library projects (e.g., Perseus, Open Video)
 - E-science databases (e.g., Genbank, EOS, sensor nets) Petabytes per day, exobytes per year
 - Indexes (e.g., Google, Bing, Yahoo)
 - Contributor Run (e.g., Wikipedia, Ebird, ibiblio)
 - Social Networks (e.g., Twitter, FB)
- Trend: special collections and customized services
 - Academic libraries: institutional repositories, research data, specialized DLs, instruction
 - Public libraries: local interest exhibits, instruction and consultation



The Personal DL Milieu

- Multiple capture of streams of personal memories
 - Files: photos, music, videos, texts, e-books
 - Communiqués: emails, blog posts, wall posts, tweets
 - Streams: SMS feeds, sensor streams (lifelogs, smart devices)
 - Secondary: annotations, hyperlinks, friend networks
 - Profiles, passwords, purchases, access and activity logs
 - Personal health records (cumulative or virtual)
- Multiple generations
 - Annotations, edits, versions of objects
 - Formats and applications
- Multiple devices and formats: Cloud solution?
 - How many devices do you use? Have you lost data?



Institutional and Personal Time

- Libraries are memory institutions: avoid rediscovering knowledge
 - Access implies persistence
 - There is substantial responsibility in distinguishing between what is ephemeral and what is worth keeping
 - Digital libraries often emphasize local content (the popular materials are everywhere)
- Digital age now spans 3 generations
- Personal memories increase in value with age
- Digital assets require attention: no reliable attics
- **Claim: Digital libraries must participate in collection, management, and perpetuation of personal memories**



Web 2.0 Drivers

- Web 2.0 is about more active participation (pull and push)
 - Communication and collaboration
 - Synchronous (e.g., IM, skype, IRC, RSS feeds)
 - Asynchronous (e.g., FB, Twitter, Flickr, wikis)
 - Collaboration gives new opportunities to libraries
 - Tagging and bookmarks (e.g., delicious, stumbleupon)
 - LibraryThing

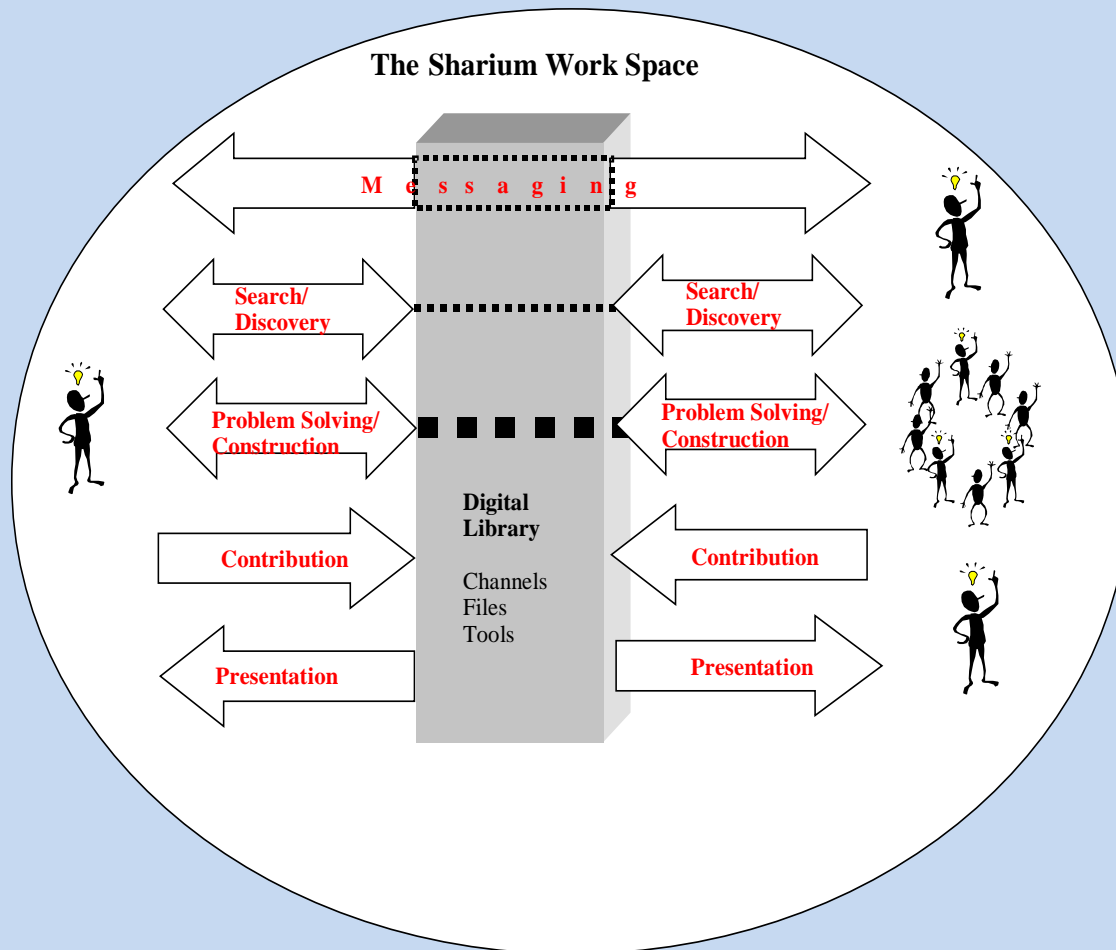


DL Collections and Services: The Sharium

- Digital Libraries are active workspaces in which many stakeholders participate
 - Multimedia streams as well as files/objects
 - Systems exhibit behavior (dynamic and interactive; computational; memorial)
 - Web 2.0 end user interactions: contributions; annotations/tags; crowd sourcing---these become part of the collection and must also be managed



Circa 1995 Sharium Workspace



Marchionini, G. (1999). Augmenting library services: Toward the sharium.
Proceedings of International Symposium on Digital Libraries 1999 (Tsukuba, Japan, September 28-29, 1999). 40-47



Spectrum of Examples

- Valley of the Shadow <http://valley.lib.virginia.edu/>
 - Contributions solicited physically
- Cornell Lab of Ornithology
<http://www.birds.cornell.edu/netcommunity>
 - Several projects where people contribute data
- Baltimore Learning Community [now defunct]
 - Content centralized, Teacher lesson plans and use notes contributed
- arXiv physics preprints <http://arxiv.org/>
 - Registered users, contributor responsibility
- Worm Community System (and other scientific collaboratories)
<http://www.canis.uiuc.edu/projects/wcs/index.html>
 - Data sharing, some with embargo (e.g., dbGaP)
- Ibiblio <http://ibiblio.org/index.html>
 - Contributor run DL with 1500+ collections
- Europeana <http://europeana.eu/portal/>
 - Member libraries, end user feedback
- Wikipedia
 - Evolving policies

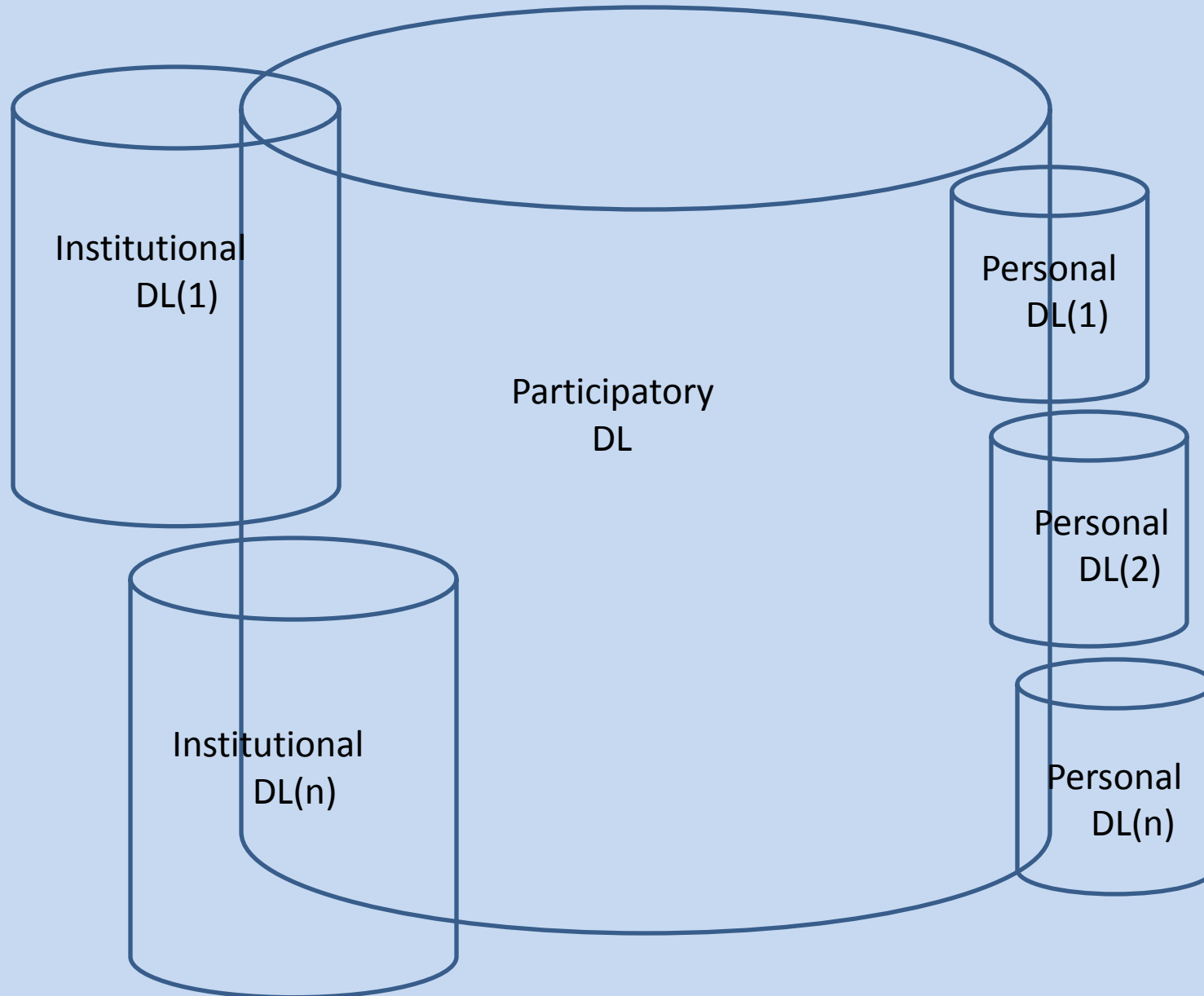


BitTorrent

- Virtual communities/co-ops
- **What.cd** for music sharing
 - Participation is payment
 - 100,000 participants; ~400,000 music albums
 - Invitation only; heavily regulated (one must earn rights to invite); the actions of invitees propagate to inviter (invite losers, you lose and vice versa)
 - Ratio system to modulate participation (including incentives and punishments)
 - Allow sharing as well as seeding new torrents
 - File integrity standards (e.g., format, bit rate) as well as metadata/authority standards
 - These quality standards are a hallmark of carefully controlled private sites like **what.cd**
 - Requests can be made and voted on/discussed by community—a kind of collection development mechanism



Participatory Trend





What are the roles of institutional DLs in Personal DLs?

- Storage? [economies of scale; digital estates; trusted imprimaturs]
- Technical and informational consulting? [reference and training]
- Customized services? [indexes, apps]
- Local history/cultural memory? [personal digital heritage blends into local heritage]
- Vocabulary and open source standards? [connect my digital life to other digital lives]



UNC

SCHOOL OF INFORMATION
AND LIBRARY SCIENCE



Gary Marchionini, UNC-CH



UNC

SCHOOL OF INFORMATION
AND LIBRARY SCIENCES





Personal and Cultural Identity

- PIM bleeds into GIM into DL
 - Implications for preservation
 - Implications for privacy
 - Implications for identity
- Sensor streams, click streams, and personal histories (e.g., MyLifeBits)
- Projections+Reflections=Proflections
- Institutional networks (e.g., DLs) recapitulate our personal networks
- DLs can become trusted personal repositories



From the Vision to the Implementation

- Many research and development challenges
- Many social challenges
- Information professionals must preserve the values of libraries (e.g., access, cooperation, stewardship) while adapting to powerful trends of computationally-augmented human behavior and knowledge generation.



Personal Identity: Your Proflection

- What are the surrogates for you?
- What are your surrogates in cyberspace? (there are more than you might think)
- Proflection as surrogate for self
 - Projection of surrogates into cyberspace
 - Conscious (e.g., your webpage, Facebook profile)
 - Implicit (your click streams, your exoinformation)
 - Reflections of self in cyberspace
 - What others say/link about/to you
 - You are aware or not
 - What machines and algorithms say/link about you and all of your projections and reflections
- A fundamental issue of cyberinfrastructure
 - It is a stimulating and challenging new world for information professionals—the curse of interesting times and the euphoria of life in these times

Marchionini, G. 2010. Information Concepts: From Books to Cyberspace Identities. Morgan-Claypool Synthesis Lectures on Information Concepts, Retrieval, and Services, (doi:10.2200/S00306ED1V01Y201010ICR016)




Digital Surrogates are Malleable





Facebook Profile



Wall





Info


Photos (5)

Friends

Friends (337)


Gary Marchionini


 Born on September 12  Add your current work information  Add your education information  Edit Profile



Education and Work Edit





Share Your Experiences

 Add Your Work Information

 Add Your School Information

Arts and Entertainment Edit

Share Your Interests





From candyschwart...



From ario_



From candyschwart...



From MC-IGN



From SIGIR2006



From david.portab...



From EvaMen



From ario_



From fstutzman



From anikarenina



From Flickmor



From anikarenina



From ECDL 2009



From jcdl2009



From ECDL 2009



From anderbill



From UNC SILS



From UNC SILS



From UNC SILS



From UNC SILS

Gary Marchionini, UNC-CH



Search Engine Results Pages

Videos for **gary marchionini**



YouTube, Open Video and
digital video repositories 🔍

2 min - 14 Oct 2008

Uploaded by uskala
[youtube.com](#)



Gary Marchionini: Informatics
in Action ... 🔍

20 min - 24 Oct 2007

Uploaded by UNCCChapelHill
[youtube.com](#)

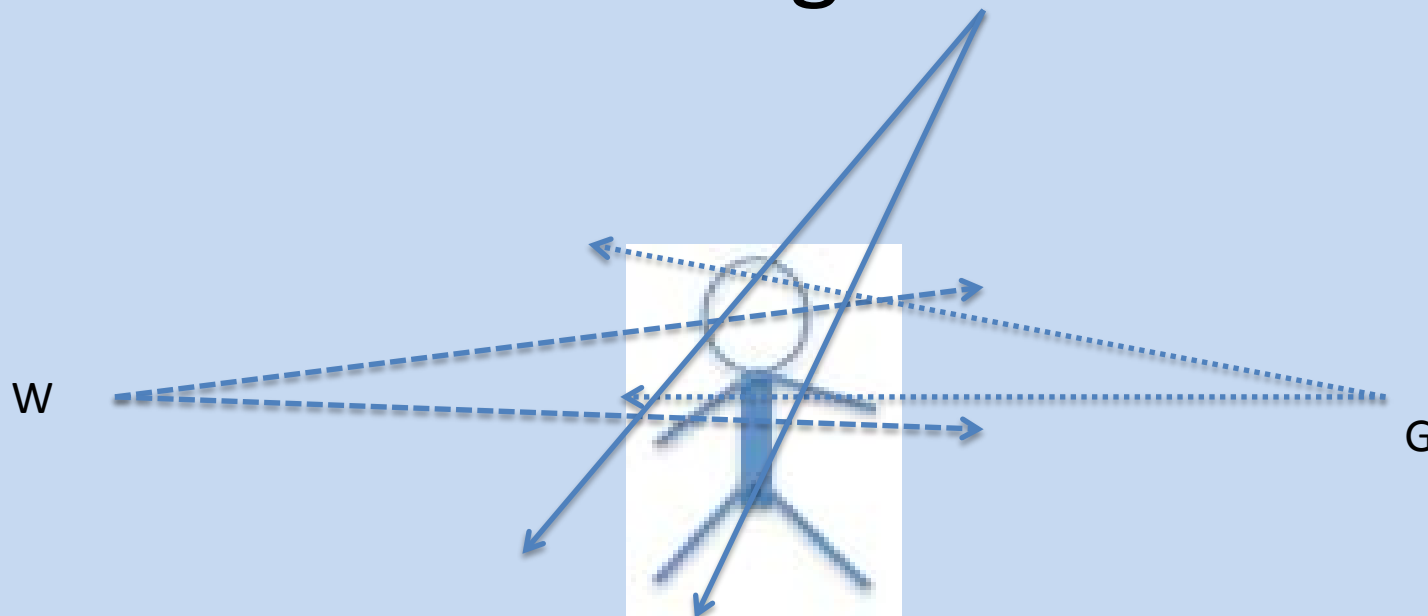


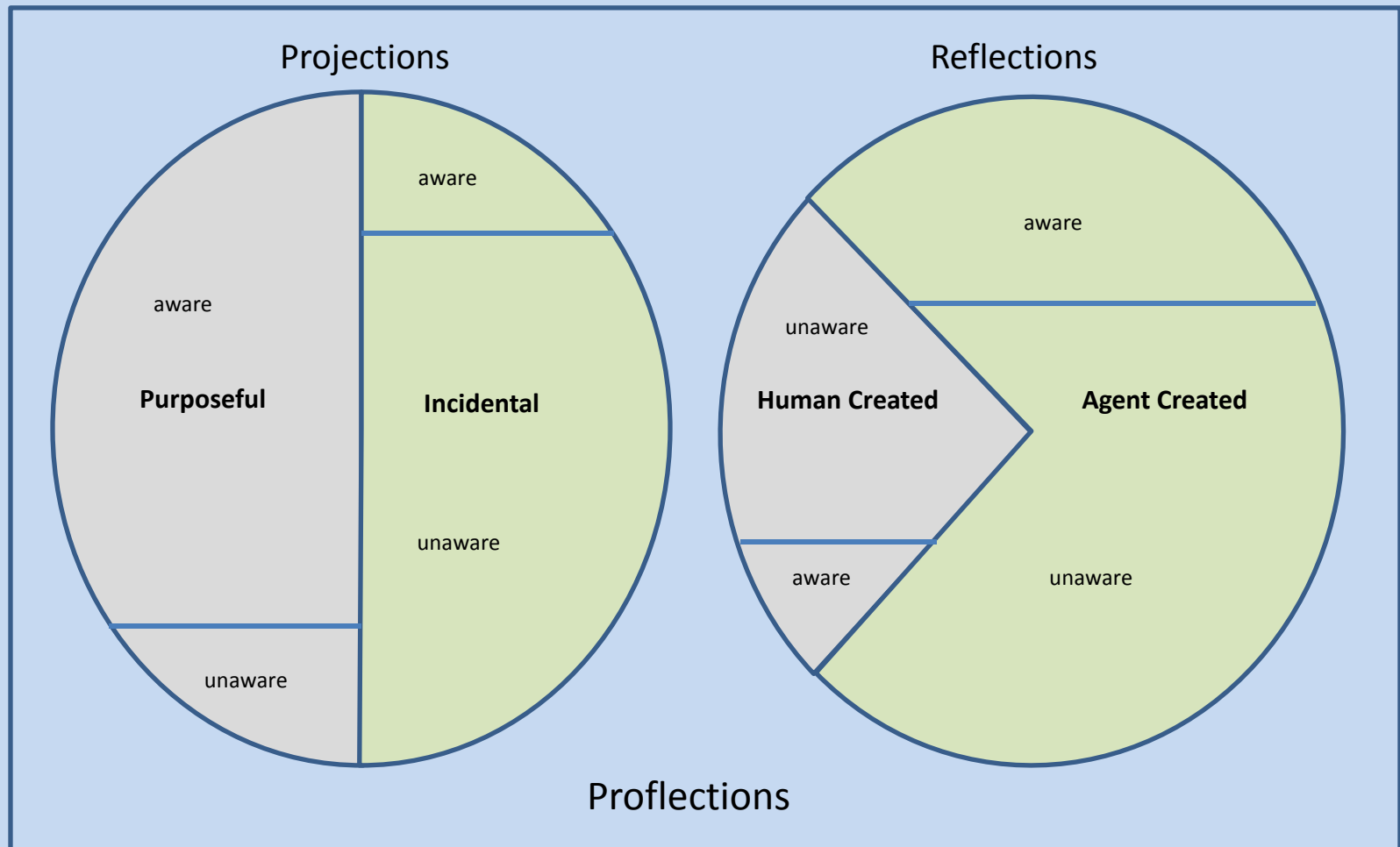
What does YOUR avatar say about you?





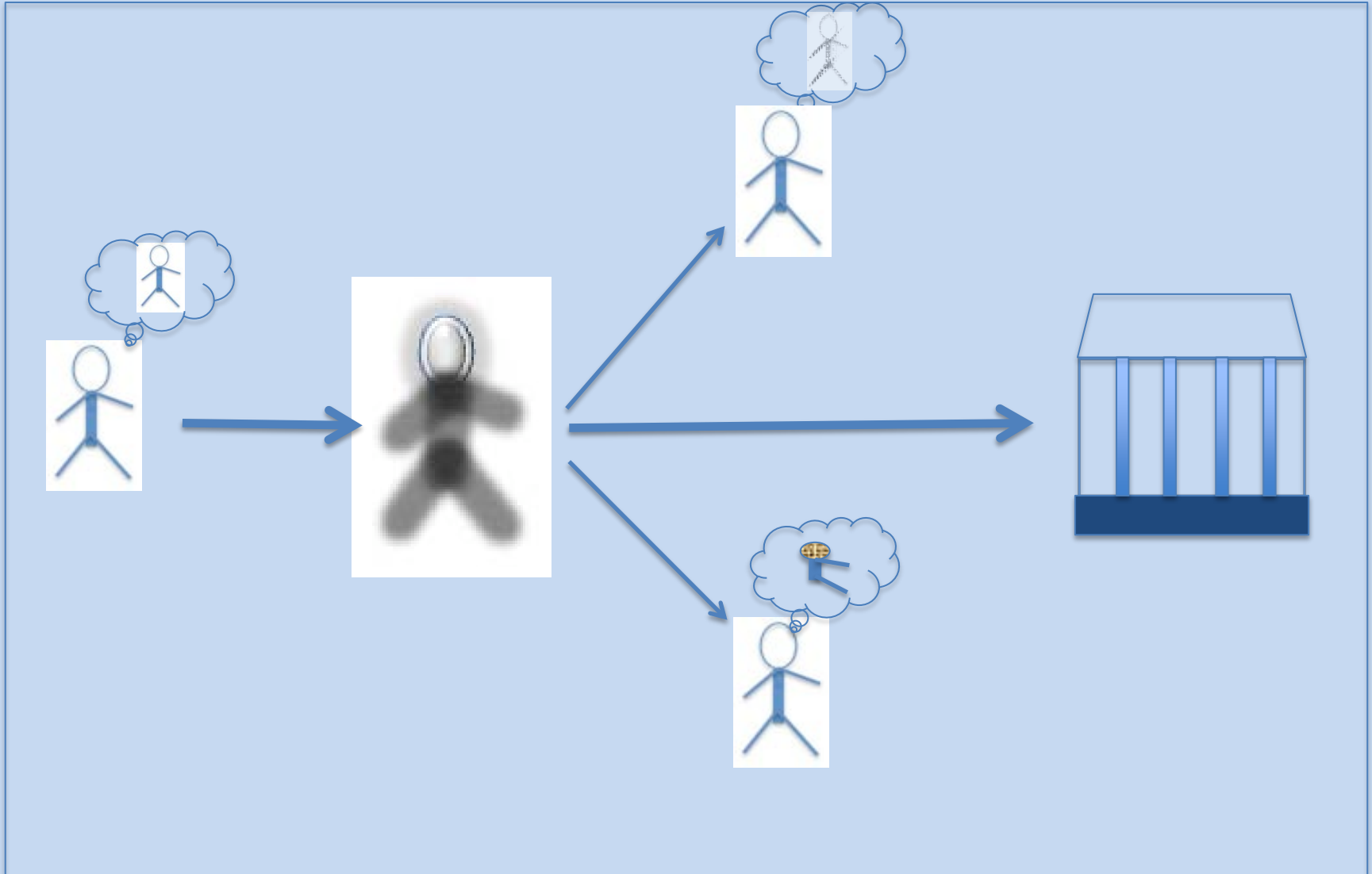
A Scholar's Identity to ISI, Scopus, Google







Identity Views



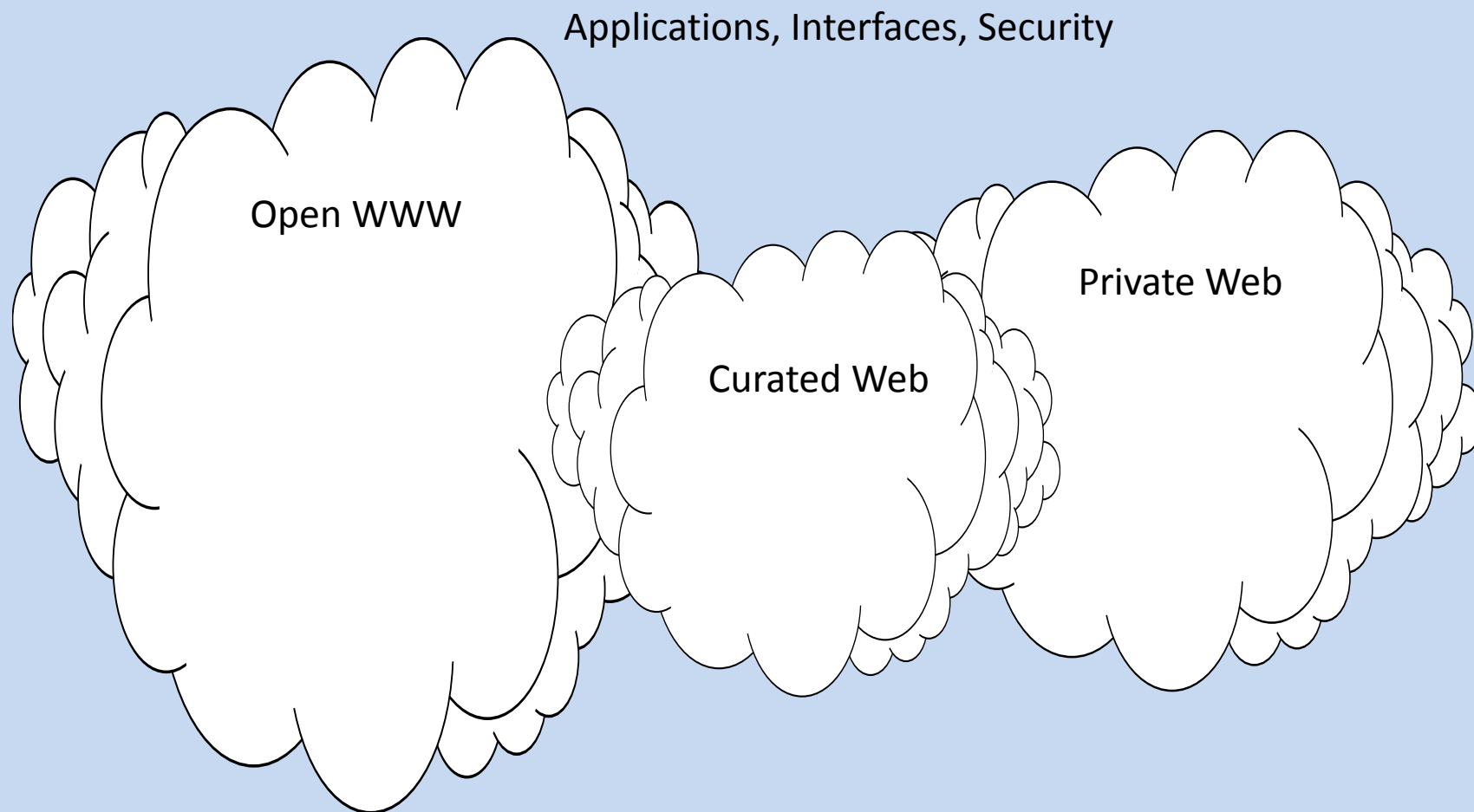


Library Responsibilities and Opportunities

- Balance public and private DLs
 - Privacy policies (e.g., blur my grandchildren's faces on the public site)
 - Ownership policies (e.g., legacy and inheritance rights; right to be forgotten)
 - User assistance and consultation on PIM
 - Facilitating writing as well as reading (access)



Challenge: Managing Blurred Boundaries





Libraries of People

- Digital Assets
 - Private, secure
 - Selectively shared
 - Digital afterlife
- Quid pro quo services
 - Trust, authority
 - Shared interests
 - Customized apps



The SILS LifeTime Library: Living and Learning Read/Write Libraries

- Trusted repository
 - Storage for life
 - Applications platform
- Shared Resources
 - Member resources
 - Connection to institution and community
- Pilot test in Fall 2010 and Spring 2011 at UNC-CH
- All incoming SILS students Fall 2011





LifeTime Library Circa Feb 2012

- 40 TB Grid
- iRods policy-based middleware
- iDrop Java application; iDrop web version
- 150 users
 - .25 TB storage per user
 - ~.5 TB total used
 - ~300,000 files
 - ~4 GB/user (range 0-155 GB)
 - ~2000 files/user
 - ~ 15,000 files/month uploaded
- SILS courses (personal DL, metadata)
- Harvester apps (iTunes, Flickr, FB)
- Shared resources: SILS photo library (2600 photos in 64 sets harvested from Flickr; video and special collections coming)



LifeTime Library 2012 Issues

- Synchronization
 - Devices, user control
- How to manage sharing?
 - Profile management, groups
- Sustainability and cost modeling
 - Storage, bandwidth, # Replicas, checksum frequency
- Customized apps
 - Indexes, metadata augmenters, privacy tools
- Organizational schemes
 - Public, private



Thank You

gary@ils.unc.edu

<http://ils.unc.edu/~march>