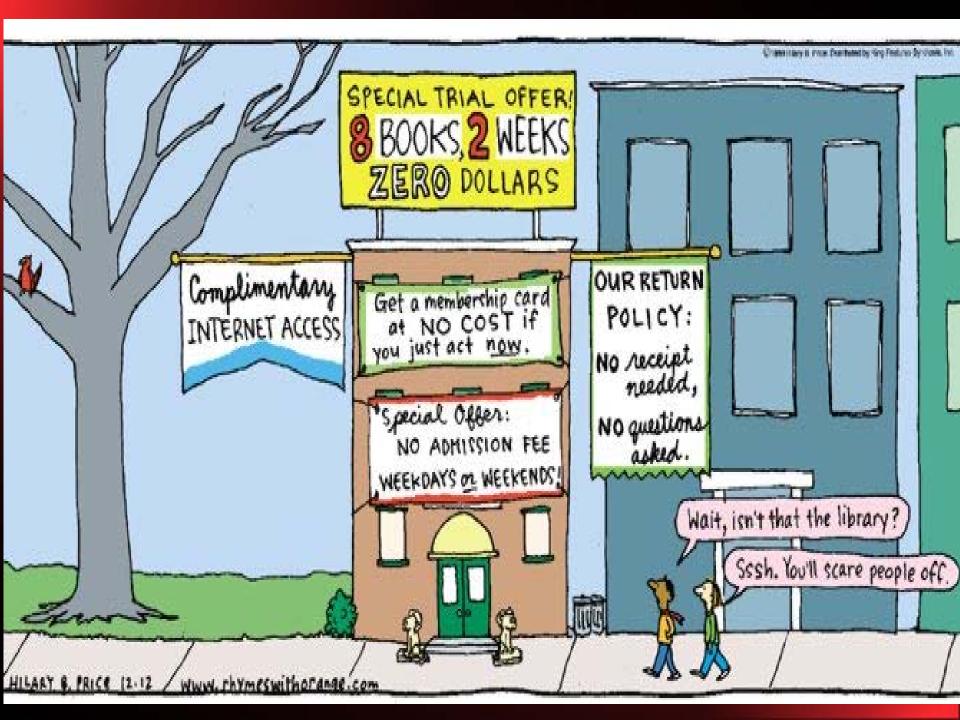
# Content Management Strategies for the Digital Era:

Opportunities and Best Options for (learning) Organizations

Dr. M.G. Sreekumar
UNESCO Coordinator, Greenstone Support for South Asia
Librarian & Head, CDDL, IIM Kozhikode

### Agenda

- The Information Paradigm Some Thoughts
- Information Management Trends in Libraries
- (Best) Practices in LIS
- Library Automation the 'KG' thing of LIS Modernization
- Digital Libraries Overview
- Open Access Overview
- Open Access Archives Institutional Repositories
- Information Aggregation and Integration Strategies
- Emerging Technologies



#### **Future Libraries?**

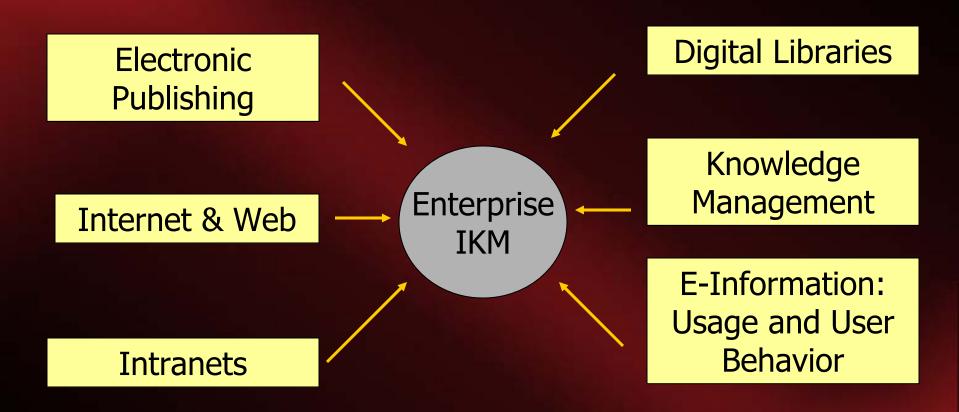
# What is a library and what should it be in 2012, 2020 and beyond...

- What does the academic library of the future look like?
- Where do its walls begin and end?
- On campus?
- On our desktop?
- At home?
- Does it still have a function as a separate and distinct space?
- Or has it become the first step to an all-virtual future?
- Libraries have never been more interesting, difficult and challenging...

### Challenges of the Day

- ✓ Relevance of Libraries in the Google Era
- ✓ Retention of Users, especially the New Gen
- ✓ Proliferation of Content
- ✓ Diverse Datastreams Content Categories, Publication Types
- ✓ Multimedia, Polymedia, Multiformats
- ✓ Collection Building Acquisition, Subscriptions, Licensing...
- ✓ Copyright, Intellectual Property, Fair Use...
- ✓ Technology Complexities, Infrastructure Issues
- ✓ Publishers' Stringent Policies / Monopolies
- ✓ Integration of legacy systems and the new genre

#### Factors of Change



Experiences? Lessons? Impact? Implications?

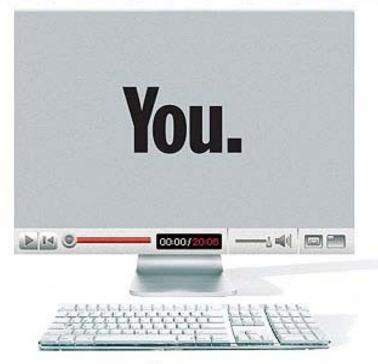
#### IM: Key Goals

Develop and manage a <u>dynamic</u>, <u>unified</u> information resource base (content repository) that gathers and organizes relevant internal and global information resources, based on a taxonomy of information needs of the enterprise, and make these available for learning and informed decision making.

#### IM: Key Goals...

- Support <u>different manifestations of information sources</u> implicit/explicit, print/digital, local/remote, free/commercial, etc.
- Support for delivering <u>personalized</u> <u>information services</u> to staff, both on demand and in anticipation.

# PERSON OF THE YEAR



Yes, you. You control the Information Age. Welcome to your world.

- User Generated
   Internet Content:
  - Blogs
  - YouTube
  - MySpace

 And the same is TRUE with Scholarly Communication too!!

#### Concerns

- In an information world in which Google apparently offers us everything, what place is there for the traditional, and even the digital, library?
- In a library environment which is increasingly moving to the delivery of online rather than print resources, what of the academic library's traditional place at the heart of campus life?
- What about the impact of repositories and open access on the delivery of library resources?
- And the need to digitise and make more widely accessible key scholarly resources?
- And what of the calls for libraries to play a central role in the promotion of 'information literacy'?

#### The Google Generation



#### Magic of the Web

- The universe (which others call the Web) is the place where society keeps the sum total of human knowledge. It's where we learn and play, shop and do business, keep up with old friends and meet new ones ...
- Today we stand at the epicenter of a revolution in how society creates, organizes, locates, presents, and preserves information ...
- .... It's all the Web ....

Ian Witten et al. in "Web Dragons"

#### New Drivers for Change - Environment

- Ubiquitous use of the Web
- "Simpler" search techniques (Google)
- "Information to go"
- Same data, different presentations
- Open access and new publishing models

#### New Drivers for Change - Environment

- Social, personal and work activities blurred
- Information users as information providers
- Interactive do-it-yourself capability of the Web is at odds with highly-wrought library systems
- Technical/licensing/copyright issues
- A volatile landscape of rapid change

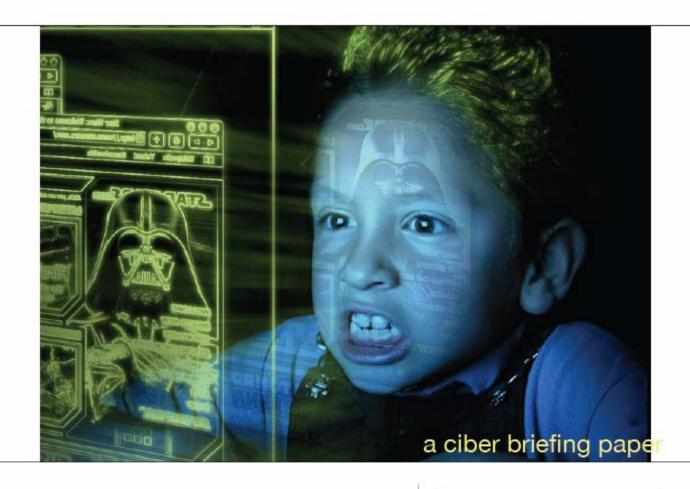
# New Drivers for Change - Students

- Integration of library with VLE
- Web 2 and social networking
- Amazon/supermarket profiling approach
- The digital native generation
- Students as paying customers

## New Drivers for Change - Research

- Access to deep web
- Underlying data important for research
- Re-use of data is possible
- Deposit in institutional repositories
- Mass digitisation (Google, MS, JISC)



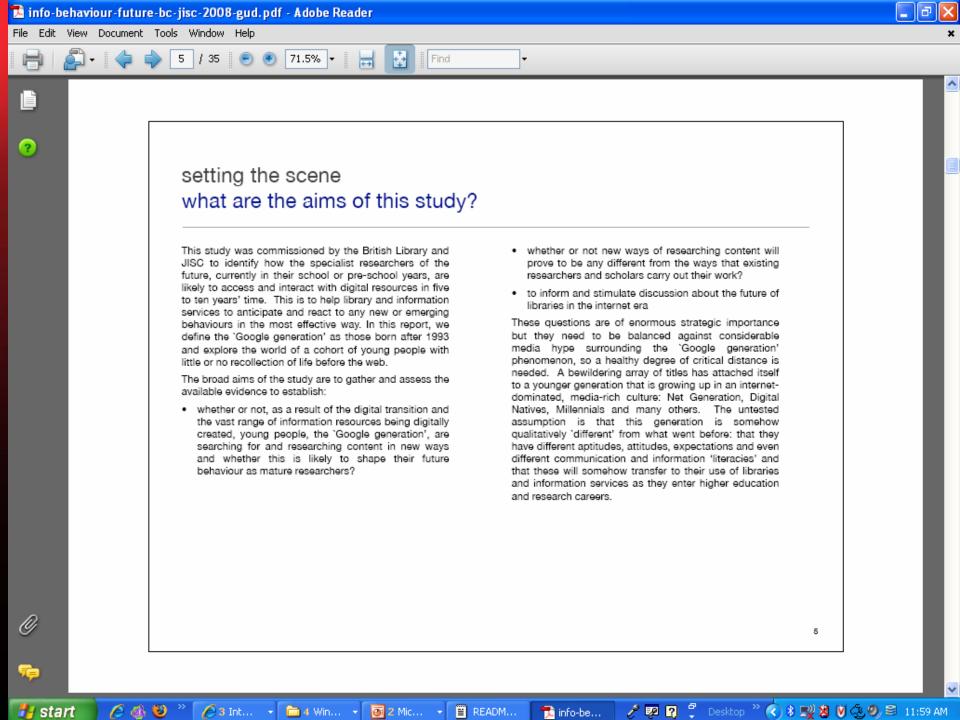


information behaviour of the researcher of the future

11 January 2008







#### Google Syndrome

 Can we assume that a search engine will – intrinsically – teach students how to find, manage and interpret information ??

#### Google Syndrome

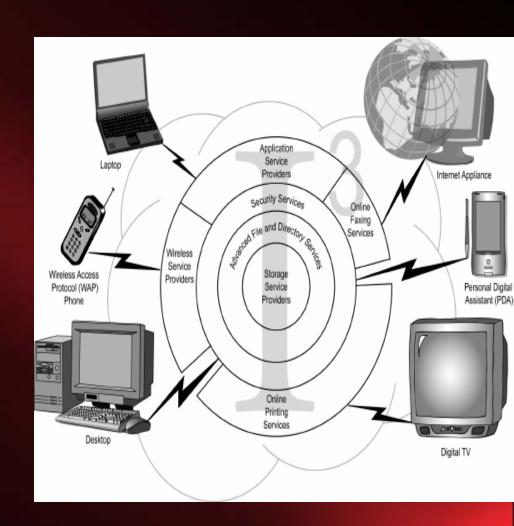
 If a generation of students in primary and secondary schools, particularly in the state sector, are "managing" education without a properly funded library and the help of qualified librarians, then not only will literacy levels and examination results suffer, but so will our universities and workplaces.

#### Top Tech Trends in IT / LIS

- Web 2.0 / Library 2.0
- Blogs / RSS Feeds / Wikis / Podcasts / Webcasts
- Open Source Software, Open Standards, Open URL
- User Tagging, Automated Tagging
- Web OPACs, and Interface Design
- Seamless Integration / Aggregation
- OA -> OAP + OAA
- Open Resource Discovery Tools Google Scholar
- E-Books, E-Journals, E-Resources
- Harvesting, Federation, Metasearching
- Digital Rights Management

## Emerging Publishing/Access Models and the I<sup>3</sup>

- Open Access Archives
  - Open Archives/InitiativeOAi / OAI
- Open Access Publishing
  - OA Journals
- Open Source
- Open Standards
- OpenURL
  - Open URL Framework



### **Information Strategy Tips**

- ✓ Context = Scenarios, Paradigms
- √ Constant = Change
- ✓ Technology = Facilitate, and NO intimidate
- ✓ Information = The Big Picture Landscape
- ✓ Content = Aggregate, Integrate
- ✓ Service = Markup, Market
- ✓ Capital = Human, Tacit, Values, and Users



### **Lib2.0**

L 1.0		L 2.0
Closed stacks	>	Open stacks
Collection development	>	Library suggestion box
Preorganized ILS	>	User tagging
Walk-in services	>	Globally available services
"Read-only" catalog	>	Amazon-style comments
Print newsletter mailed out	>	Team-built blog
Easy = dumb users	>	Easy = smart systems
Limited service options	>	Broad range of options
Information as commodity	>	Information as conversation
Monolithic applications	>	Flexible, adaptive modules
Mission focus is output	>	Mission focus is outcome
Focus on bringing 'em in	>	Focus on finding the user
ILS is core operation	>	User services are core

# Penetration of E-Content in Libraries

#### PUBLICATION TYPES

- E-Books, E-Journals...
- Aggregated Scholarly
   E-Journal Databases
- Databases, CBT/ WBT
- Portals, Vortals...
- Value added services
- Preprints, Eprints, E-Documents....

#### DOCUMENT FORMATS

- ASCII, RTF, HTML, SGML, Postscript, PDF, Proprietary, Native Application Formats
- Images, Graphics
- Audio
- Video
- XHTML, ASP, PHP, XML ...

#### **Major Drivers**

- Information/Knowledge Proliferation
- Technology Progress (push)
- Open Standards / Technologies
- Open Source Software
- Market Demands
- Community / Society

# The Changing (challenging) Role of Information Professionals

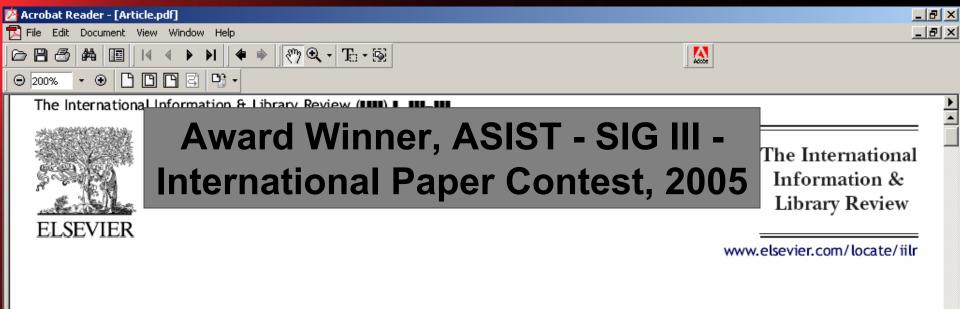
**Book keepers** 

Bit Keepers

(Book/Journal) Stores Officers/Custodians **Information providers Content / knowledge providers** Content / knowledge sellers

#### **Good (Best) Practices**

- Skill Sets Leadership, Communication, Team Work, CE, Perseverance, Persistence....
- Flexibility / Open to Learning / Accept Changes
- Open Source Philosophy
- Open Access Advocacy
- Preach as well as Practice
- Sharing of Knowledge Acquired



Seamless aggregation and integration of diverse datastreams: Essential strategies for building practical digital libraries and electronic information systems

M.G. Sreekumar\*, T. Sunitha

Center for Development of Digital Libraries (CDDL), Indian Institute of Management Kozhikode (IIMK), IIMK Campus P.O., Calicut 673570, Kerala, India















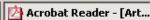














#### Strategy Facets - for an excellent hybrid ambience

- 1. Value and Preserve the Traditional Systems while Embracing Technology for Development
- 2. Aggregation / Integration Strategies
- 3. Library Automation (integration of the traditional library activities/collection)
- 4. Digital Library (creation of efficient information systems using digital collections for effective browse, search, retrieval and long-term preservation)
- 5. OA Archives (pooling the organization's intellectual capital)
- 6. Library Portal (seamless integration of the organization's entire knowledge base—traditional, paper as well as electronic technologies)
- 7. Consortium Approach to Library Subscriptions

### Strategic Planning

- How to map / scale (assess) the information needs / requirements of the Organization
- How to Communicate this to the Management
- How to build infrastructure Planning, Execution
- How to Collect, Consolidate, Build and Preserve Local Content
- How to build E-Resources / Published Content
- Information Literacy to Information Professionals
- Information Literacy for Users

#### **Libraries - Shifts**

- Traditional / Automated
  - » Organization is physical
  - » Shelving of documents Based on Subject Cln
  - » Key Index / Catalogues / Cards / Digital Catalgs
  - » Cards Real/Virtual Author, Title, Descriptions
- Digital
- » Organization in terms of digital files /objects
- » Contains material digitized form
- » Contains digital material
- » Architecture
- » Key Metadata

#### Shift in Approaches

Traditional

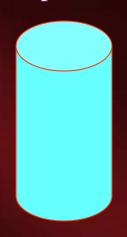
Limited/ Rigid



AACR2
CCC
CC / LCCS
DDC / UDC
Thesauri/LCSH

Automated

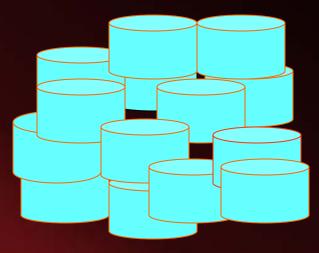
**Improved** 



AACR2
ISO 2709
CCF
MARC
Thesauri

Dig. Library

**Efficient/ Flexible** 



Metadata
DCMI -- W3C
EAD, TEI, DTD
METS,MODS,
Z39.50
MARC21

**OAI-PMH** 

#### Digital libraries ... features

- Knowledge/content management
  - Manage and access internal information assets
- Scholarly communication, education, research
  - E-journals, e-prints, e-books, data sets, e-learning
- Access to cultural collections
  - Cultural, heritage, historical & special collections, museums, biodiversity
- E-governance
  - Improved access to government policies, plans, procedures, rules and regulations
- Archiving and preservation
- Many more ...

#### The Greenstone Software

- Software suite for building, maintaining, and distributing digital library collections
- Comprehensive, open-source
- Developed by New Zealand Digital Library Project at the University of Waikato
- Distribution and promotion partners:
  - UNESCO
  - Human Info NGO, Belgium
  - NCSI, Bangalore; UCT, Cape Town;
     Dakar, Senegal; Almaty, Kazakhstan; ...
  - You!



#### **Features of Greenstone**

- Open Source Philosophy
- Interfacing & Content Delivery via Web
- Multi S/W Platform
- Multi Lingual Support
- Multi Formats
- Structured Metadata in XML using DC
- Metadata Extraction
- Searching & Browsing
- Plug-ins for Documents

- Full-text mirroring
- Text Level Penetration
- Data Compression
- Password protection
- Administrative Functions
- Concurrent & Dynamic Content Development
- Uniform Presentation
- Publishing on CDROMs
- International Presence

#### Greenstone Features contd...

- Easy Installation
- Easy Maintenance
- Content Development (3 alternate ways)
- Predominantly GLI now since (V. 2.41)
- Hierarchy Structure
- Interface Customization
  - Front Page Design, Header for the Digital Library, Collection Icon, Cover Images
- Collection Configuration (Collect.cfg) File
- Scalability, Flexibility
- Interoperability (Crosswalk), OAI Compliance
- Lifeline: Listsery / E-Group / Archives

#### Greenstone DL Software

- Access ✓ Accessible via any Web browser
  - ✓ Server runs on Windows and Unix
  - ✓ Collections can be published on CD-ROM
- Searching/ Full-text and fielded search
  - browsing <a href="#">Flexible browsing facilities</a>
    - ✓ Metadata-based (Dublin Core)
    - ✓ Collection-specific
    - √ Hierarchical phrase browsing supported
    - ✓ Creates all access structures automatically

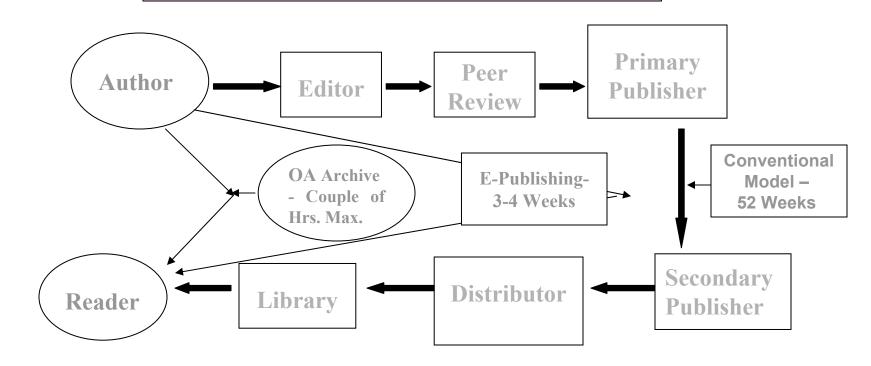
  - Extensible ✓ Plugins new document, metadata formats
    - √ Classifiers new metadata browsers
- Multilingual \( \sqrt{Documents} \) and interfaces
  - ✓ Chinese, Arabic, Maori, Russian etc (+ European, Indian Languages)
  - ✓ Multimedia: video, audio collections exist

## A Word About Open Access

#### **Ground Realities**

- World Over, Institutions / Agencies Fund Research (mostly Govt.s & NGOs)
- Authors want to publish/share their findings/thoughts with peers
- Authors give Copyright to Publishers
- Publishers do not give Authors access to the paper by default
- If your Institute/library is not subscribing to the Journal, you will have to be satisfied with the Offprints the publishers give

## **Upcoming Models of Scholarly Communication Value Chain**



OA Archive (2 - 3 hours)

E-Publishing (3 - 4 weeks)

**Conventional Publishing (52 weeks)** 

#### Institutional Scholarly Knowledge

#### PUBLISHED

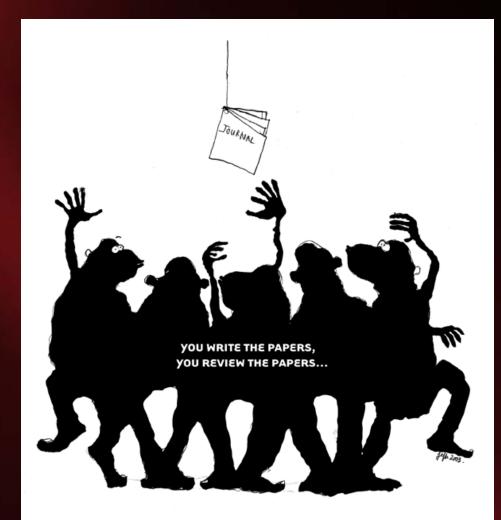
- Journals / Journal Articles
- Conference / Symposium Papers
- Books / Book Chapters
- Patents
- Cases
- **—** ...

#### UNPUBLISHED

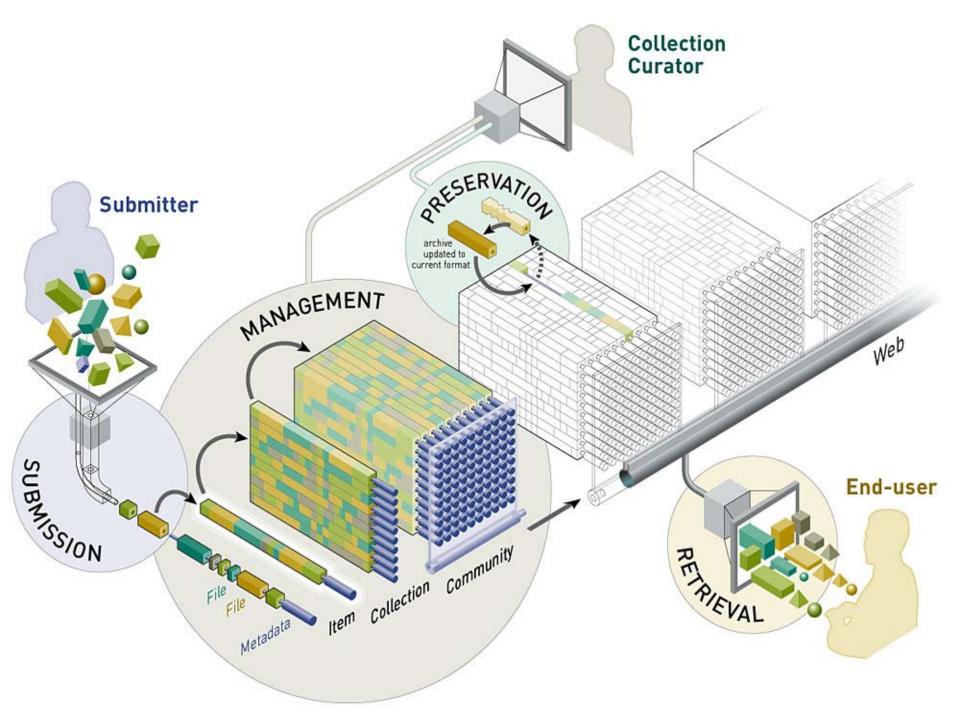
- Theses / Dissertations
- Working Papers
- Technical Reports
- Coursewares
- Classroom Presentations (Videos, Audios, PPTs...)
- Lecture Notes
- Guides / Manuals
- **–** ...

### Open Access

- Open Access Publications (OAP) Golden Road
- Open Access Archives (OAA) Green Road



WHY SHOULD YOU PAY TO READ THEM?







#### Open Access Initiatives In India

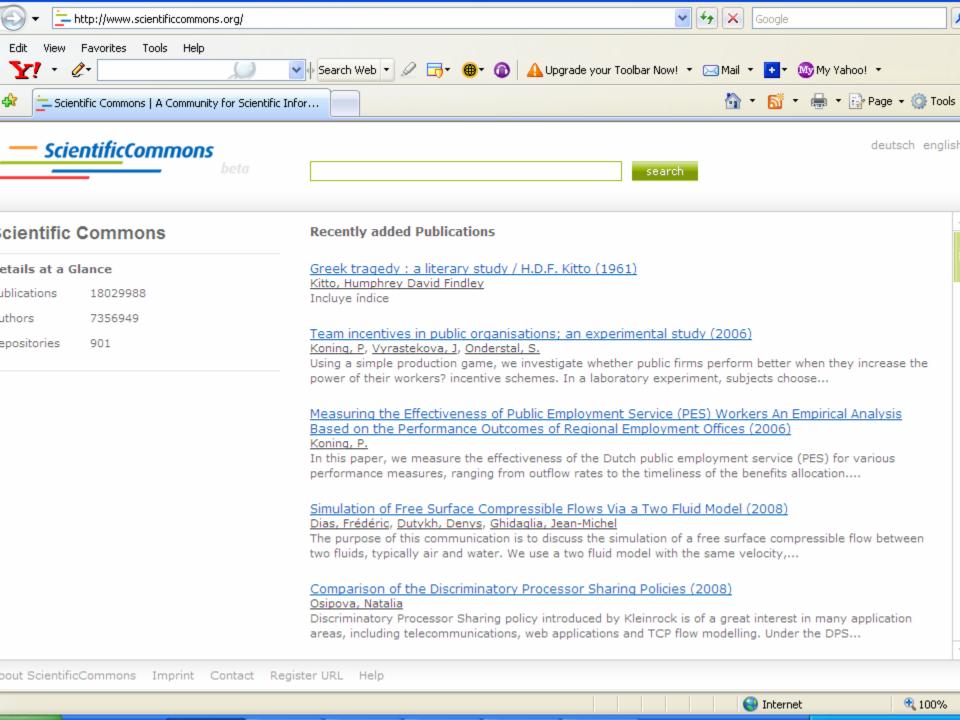
- Institutional Repositories( evolving)
  - http://casin.ncsi.iisc.ernet.in
  - Software: DSpace, E-Prints
- Subject Specific Repositories
  - Science/Technology EPrints@IISc http://eprints.iisc.ernet.in
  - Engineering / Technology http://eprint.iitd.ac.in/dspace/
  - Librarian's Digital Library (LDL)
     <a href="http://drtc.isibang.ac.in/DRTC/index.html">http://drtc.isibang.ac.in/DRTC/index.html</a>
  - OpenMed@NIC <a href="http://openmed.nic.in">http://openmed.nic.in</a>
  - Management DSpace@IIMK http://dspace.iimk.ac.in
- Electronic Theses and Dissertations
  - INFLIBNET and UGC <a href="http://dspace.inflibnet.ac.in">http://dspace.inflibnet.ac.in</a>
  - Vidyanidhi <a href="http://www.vidyanidhi.org.in">http://www.vidyanidhi.org.in</a>
- Cross Archives Search Services for Indian Repositories (CASSIR) http://ardb4.ncsi.iisc.ernet.in/oai

#### Open Access Initiatives In India...

- Open Access Journals
  - Indian Academy of Sciences (11 journals)
     <a href="http://www.ias.ac.in/pubs/journals">http://www.ias.ac.in/pubs/journals</a>
  - Indian National Science Academy (4 jls)
     <a href="http://www.insa.ac.in">http://www.insa.ac.in</a>
  - Indian MEDLARS Centre (NIC) (38 jls)
     <a href="http://medind.nic.in/">http://medind.nic.in/</a>
  - MedKnow Publications (28 jls)
     <a href="http://www.medknow.com/journals.asp">http://www.medknow.com/journals.asp</a>
  - Indianjournals.com (7 jls)
     <a href="http://www.indianjournals.com">http://www.indianjournals.com</a>
  - Kamala-Raj Enterprises (5 jls)
     http://www.krepublishers.com/KRE-New-J

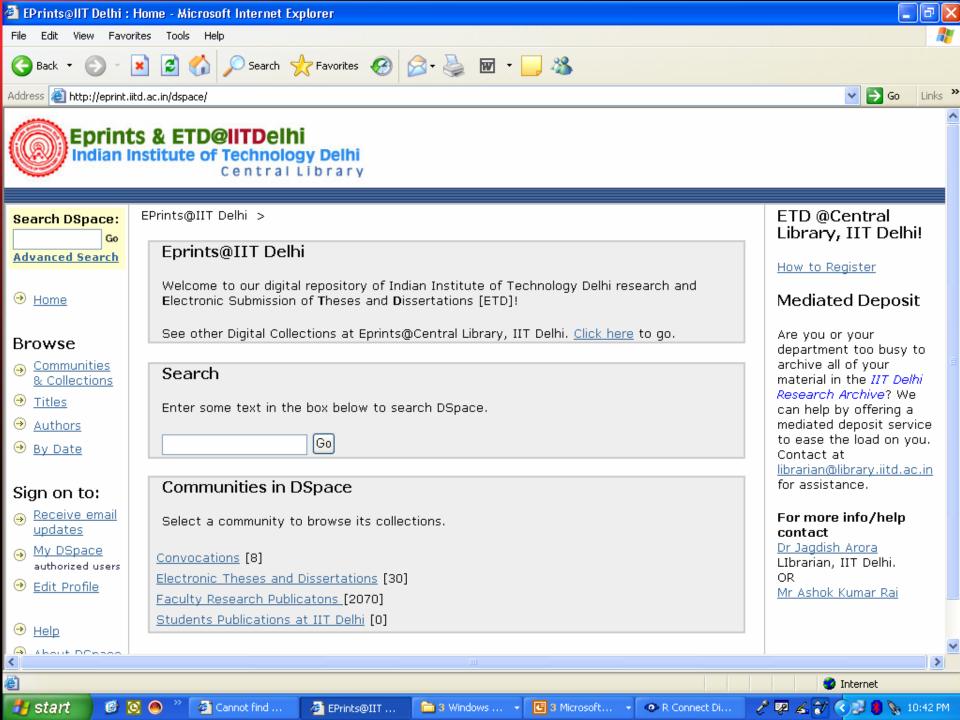
## Open Access Literature (Resource Discovery)

- Directory of Open Access Journals (DOAJ) 3270+ http://www.doaj.org/
- Open J-Gate 3000+ Journal, Millions of Citations http://openj-gate.com/
- Open DOAR 1000+ OA Archives http://www.opendoar.org/
- ROAR: Registry of Open Access Repositories 1000+OAA <a href="http://archives.eprints.org/">http://archives.eprints.org/</a>
- Scientific Commons 901 OAA, 18029988 Articles, 2356949 Authors <a href="http://www.scientificcommons.org/">http://www.scientificcommons.org/</a>
- OAlster 934 OAA + 14,673,867 records <u>http://oaister.umdl.umich.edu/o/oaister/</u>
- DMOZ 59 OAA http://dmoz.org/Science/Publications/Archives/



#### **DSpace Software / Utilities**

- 1. Java SDK 1.4.2
- 2. Apache 2.0.54
- 3. Tomcat 5.0.28
- 4. mod\_jk 2.0.4
- **5. Apache Ant 1.6.5**
- 6. PostgreSQL 8.0.2
- 7. DSpace 1.5x / 2.x



# Information (Content) Aggregation / Integration

- Library Automation
- Library Portal
- Digital Library
- Institutional Repository
- Library Consortia

# Need for Content Integration / Organization

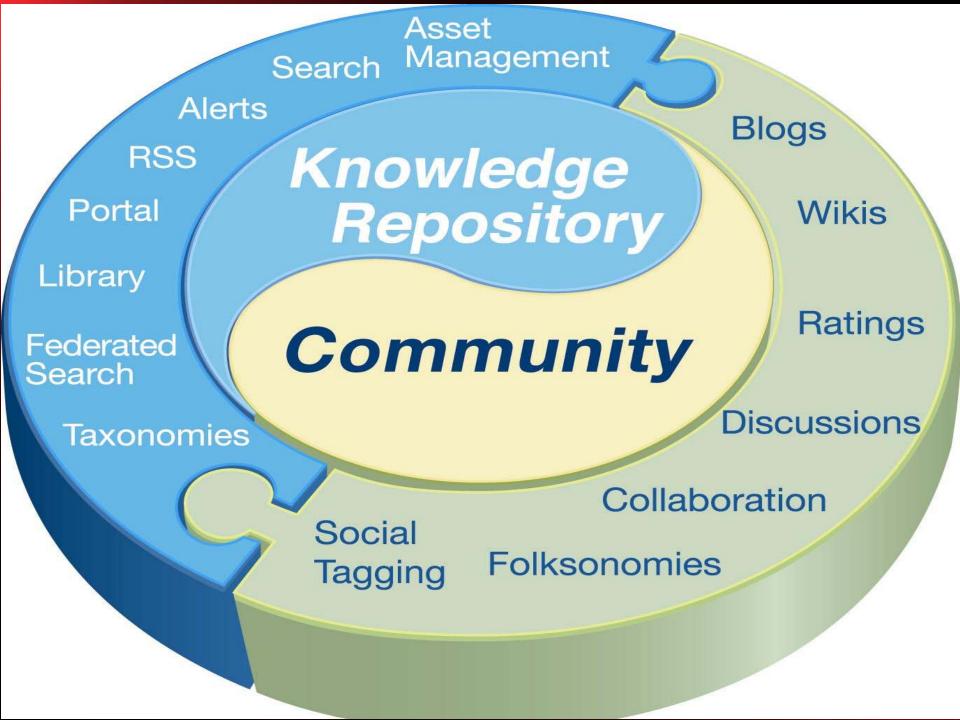
- Assuring Seamless Access to the Content
- Need for a single Info. Gateway / Access Point
- Multi Formats, Media, Platforms (Content / Data in different formats)
- Data encoding (role of markup languages)
- Role of Metadata (role of Standards)
- Structured Metadata (role of XML)
- Need for Interoperability
- Interface / Delivery / Presentation
- Exorbitant cost of proprietary DL S/W

#### KM Systems - Flaws

- Historically, the goal of knowledge management has been to create, capture, preserve, and distribute information to drive finite and tangible objectives, such as improving competitiveness, fostering innovation, and streamlining organizational effectiveness.
- However, most knowledge management solutions exist merely as passive repositories.

#### KM – Next Progression

- It is the confluence of content management systems, knowledge management systems, and social tools to create networks of knowledge.
- Once an organization's knowledge repositories combine with the "wisdom of the community," a fundamentally new approach to socializing knowledge and enhancing an organization's innate intelligence is created.
- Called "Social Intelligence".
- That taps into an organization's social intelligence through the creation of social knowledge networks



#### 2.0-style Service Examples

- Blogging
- User tagging
- Folksonomies
- User comments
- User ratings
- RSS feeds
- Feed aggregators

- Wikis
- Skypecasts
- Community citation services
- Community photo services
- Community book services

#### Other Websites with 2.0 Characteristics

- Flickr
- eBay
- De.licio.us
- Second Life
- CiteULike
- Librarything

- Connotea
- LiveJournal
- MySpace
- Technorati
- Netflix

#### Key Library 2.0 Concepts

- The library is everywhere\*
- The library has no barriers\*
- The library invites participation
- The library uses flexible, best-of-breed, component-based systems
- The library is a human-centered organization
- \* Also stated as the concept of the library as a place of unrestricted access to information (McDonald, 2006)

#### Common Traits of 2.0 Services

- Interactivity
- Respects and leverages user contributions
- Complementary/compatible/cross-referential
- Treats information as a conversation
- Emphasis on ease of use
- Sharing—use/reuse/remix/mashups encouraged

### Another angle on 1.0 versus 2.0...

- Library 1.0 is a <u>pull</u> model
  - Acquire / store / lend / find

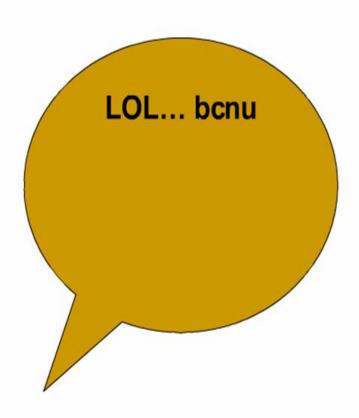
- Library 2.0 is a <u>push</u> model
  - Acquire / store / broadcast / converse

#### Blog: like a personal newspaper

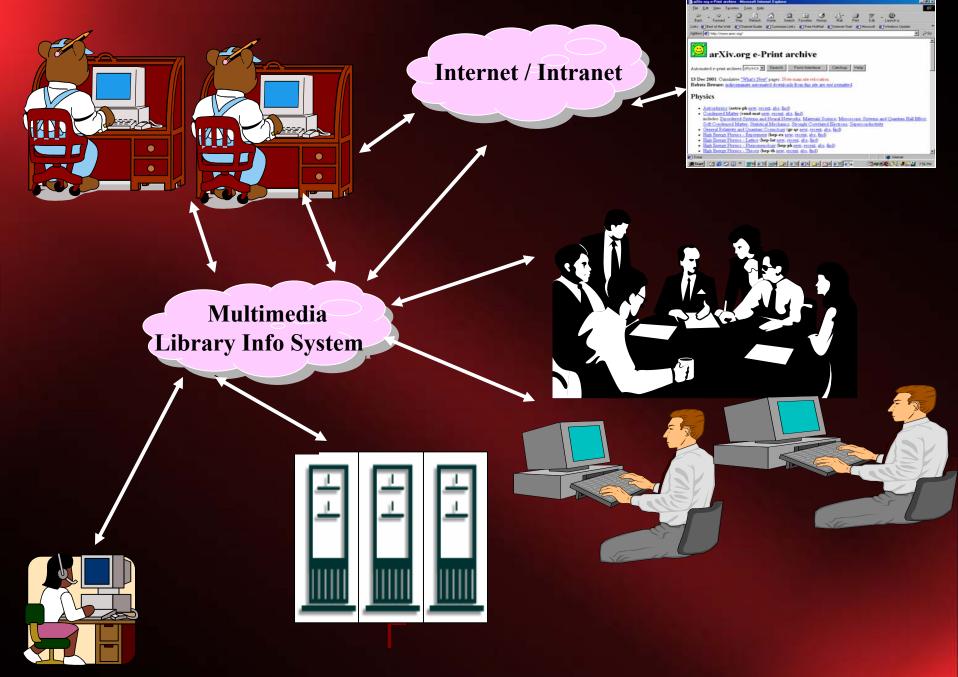


- Easy to use
- Free (at basic level)
- Easy to stay updated
- Encourages sharing and engagement

#### IM (Instant Messaging):

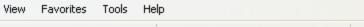


- Easy to use
- Free (at basic level)
- Encourages sharing and engagement



USER @ anywhere (access to information from anywhere)



















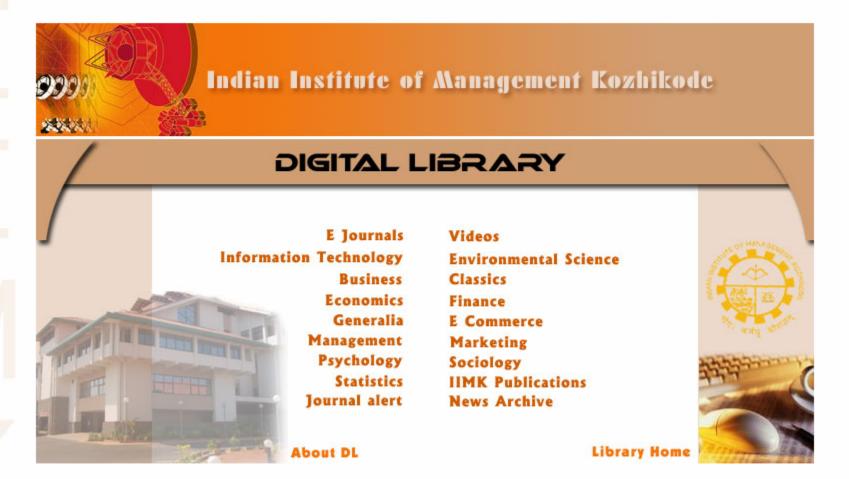




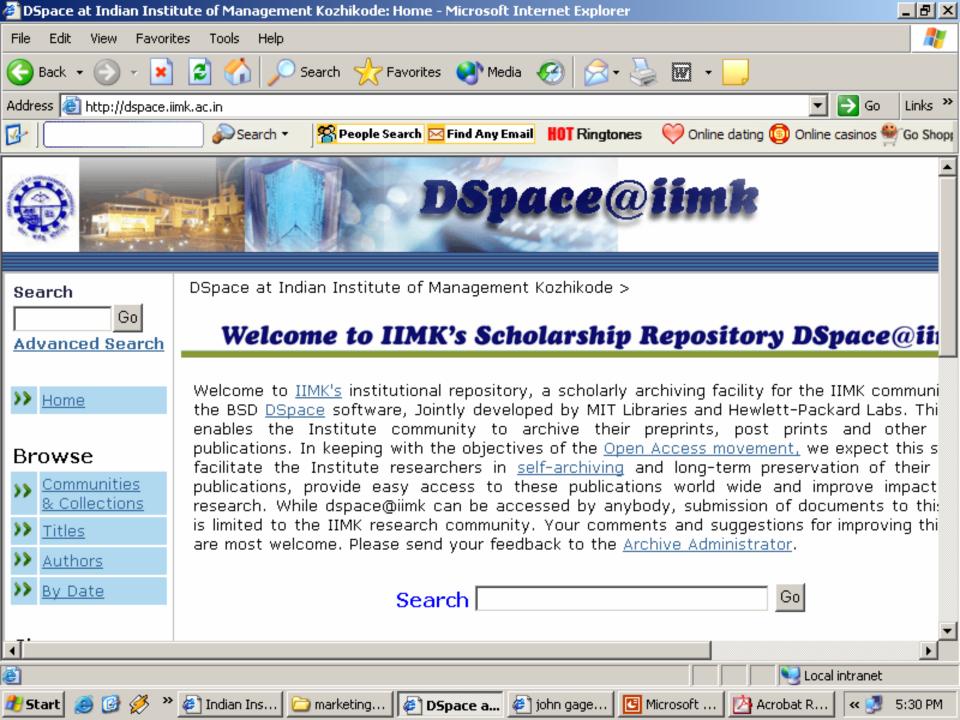








Powered by Greenstone



# Library 2.0

# LIBRARY 2.0

## Library

that fits that suggests that learns that gathers that combines that organizes

### STAFF

Creation of an Emerging Technology Committee

Integration with (e) learning environment

Library is a framework for integrating change into all levels of library operations

# Library that LETS

The library invites participation

### OPAC

- Federated search
- RSS for cataloging records & search results
- Records tagging
- User reviews

1. User-centricity

- 2. Technology-savvy environment
- 3. Reaching of the patrons long tail
- 4. Content for more than one device
- 5. Component-based software, not monolithic ILS
- 6. Constant change
- 7. Use of Web 2.0 apps and services
- 8. Open standards

The library has no barriers

The library is human

Social computing apps to meet users' need when, where and howthey need it

# THE PHYSICAL LIBRARY

Loud spaces for collaboration & conversation Mobile devices for users The library is everywhere

The library uses flexible, best-of-breed systems

Patron 2.0 = from content consumer to content creator

