IU and the BIG DIGITAL MACHINE

Robert H. McDonald
Associate Dean for Library Technologies
Associate Director Data to Insight Center

DLP Brownbag 11.11.09
Hello – Greetings - Welcome – Bonjour - Tag

- Tagging
  - #iudlp
  - #bdm
  - @mcdonald

- GoogleWave
  - with:public IU Digital Library

Robert H. McDonald
OVERVIEW

- What is the Big Digital Machine (BDM)
- BDM Directions
- Implications for IU
- Implications for Scholarly Communication
BIG DIGITAL MACHINE

- Scholarly Communications Publishing as a Service
  - SaaS
- Leverage Services Above Campuses
  - Community Partnership
  - Committee on Institutional Cooperation
- Converge various open and community source software tools to create
  - Open Source Reference Toolkit for Scholarly Communications in the broadest sense.
BDM Directions

- Initial idea came from work with several open source communities
  - DSpace (DuraSpace)
  - Fedora (DuraSpace)
  - OJS (Open Journal Systems) - PKP
  - OCS (Open Conference Systems) – PKP
  - Connexions – Rice University
  - DuraCloud (DuraSpace)
BDM Directions

- How do these systems find community that is cohesive and utilizes their software services at the highest capacity?
- One way is Software as a Service (SaaS)
  - Should leverage mass scale storage across many institutions
  - Should leverage cloud resources
  - Should be easy to use
  - Should be on demand for the scholar
OPEN SCHOLARLY COMMONS

- 2 Key Ingredients
  - Infrastructure
  - Scholars Prepared to Use the Infrastructure
LEVERAGE
COORDINATION

- This is a coordination problem not a money problem
  - Collaboration
  - Co-labor
- Note Educause Presentation by Armstrong and Starratt
  - *The Higher Education Library: What Would $2 Billion Dollars Buy*
Meta-university Collaborations

Library Books  Journals  Textbooks  Learning  Administrative

(Just to name a few…)
IMPLICATIONS FOR IU

- IUScholarWorks
  - Bloomington
  - Indianapolis
- IU Digital Library Program
- Empowering People (IU IT Strategic Plan)
  - Action Item 37
  - Action Item 55
EMPOWERING PEOPLE

http://ovpit.iu.edu/itsp2/
EP – ACTION ITEM 37 – PRESERVING CREATIVE WORKS

- IU should provision a full-featured and robust multimedia utility service to digitize (if needed) and preserve film, audio, and complete creative works. The utility should enable abundant near- and long-term storage, presentation in variety of individual or group settings, and permissions and rights management, and should provide search and retrieval for whole and partial clips based on terms or associations.
IU should provision robust systems and services that enable researchers to easily provide an institutional deposit copy of any scholarly work in any media format.
INFRASTRUCTURE @IU

- Dual Data Centers
  - Bloomington
  - Indianapolis

- Most infrastructure will be redundant via
  - I-LIGHT
  - Data Centers
  - Balanced Applications

- Opportunities
  - http://it.iu.edu/datacenter/media
    groundbreaking1.asx
DATA CURATION

- IU Scholarly Data Services
  - IUScholarWorks Data - Publishing Services
  - Consortial Data Curation/Publishing (HathiTrust)
  - Consortial Data Curation (CIC Research File System)
IU SCHOLARWORKS DATA

Image showing a diagram of a web server system with the following components:

- Item record with URL's of datasets in MDSS
- MDSS web server
- hpssfs filesystem
- HTTP Server
- IU MDSS (HPSS)

Image of a web page with a search interface titled "Endogenous Network Dynamics."
CONSORTIAL DATA - HATHITRUST

Indianapolis  Ann Arbor

Isilon OneFS Currently Supports up to 2.3 PB between Two Nodes

Two Node Federated Repository

INDIANA UNIVERSITY
HATHI TRUST
INSTITUTION
PERVERSIVE TECHNOLOGY INSTITUTE
AT INDIANA UNIVERSITY
CONSORTIAL DATA – CIC RESEARCH FILESYSTEM

- 7 Node Wide Area File System
  - Indiana University
  - Michigan State University
  - Ohio State University
  - Penn State University
  - University of Illinois
  - University of Iowa
  - University of Minnesota
- Fedora Curation Services
- GPFS-WAN – Global Parallel File System Wide Area Network
- SoFS – Scale Out File System
Diagram from IBM
WELLS ON COLLABORATION

“Consistent with this statement of policy, throughout my years of responsibility for administration I have been motivated by a strong belief that the resources of higher education are so insufficient and the opportunities and responsibilities so vast, the only sensible course is to attempt in every way to avoid unnecessary duplication among or with institutions.”

Herman B Wells, Being Lucky, 1980, p. 135
President of Indiana University, 1938-1962
DISCUSSION