A DSpace Institutional Repository: WHOAS

Ann Devenish
ICSTI Public Conference
National Library of Medicine, Bethesda, MD
8 June 2006
Agenda

Background
- Woods Hole science, the MBLWHOI Library

Building an institutional repository with DSpace
- How we got to where we are
- How to get to where we’re going

Lessons learned along the way
41°31.2’N, 70°40.0’W
Woods Hole science

MBL, formerly, Marine Biological Laboratory (1888): an international center for research, education, and training in biology, biomedicine, and ecology

National Marine Fisheries Service/Northeast Fisheries Science Center (1871): conducts research in support of commercial fisheries in U.S. waters off the northeastern United States

Sea Education Association (1975): deep-water research and education program for college students
Woods Hole science

U.S. Geological Survey (1962): geological and
depth research and mapping investigations of
the submerged continental margins of the United
States and worldwide

Woods Hole Oceanographic Institution (1930):
dedicated to research and higher education at the
frontiers of ocean science

Woods Hole Research Center (1975): address the
issues of environment through scientific research
and education and through applications of science
in public affairs
MBLWHOI Library

Joint library serving the Woods Hole science institutions

Staffed by MBL and WHOI library employees

Physical service points:

- Main ~ Lillie Laboratory (MBL); accessible 24/7
- Data Library & Archives ~ McLean Laboratory (WHOI, Quissett campus)
- Library Service Center @ 25 Falmouth Technology Park (compact shelving)
MBLWHOI Library

Virtual service points on every computer in the IP range (128.128. … )

- Databases (98+)
  - i.e., ASFA, GeoRef, Science Direct, Web of Science
- E-journals (1300+)
- PDF delivery for locally held print and ILL articles
Dreaming the IR

“Quiet” phase began in 2002
Conversations regarding the value of an IR
  – “Concentrates” the intellectual product
  – Archiving scholarly communication
Conversations regarding the value of open access
  – Need to re-shape scholarly communication
  – Champion for authors’ rights
Dreams shaped by reality

"Build within existing budget dollars and staff"
  - “Herbarium” model
  - Open source

Voluntary participation by authors
  - No institutional directives for self-archiving

Support all types of digital formats

Content had to be “Googleable”

Content had to be open access

Initial focus on MBL and WHOI authors & content
Building the IR

*Pilot project Dec 2003-July 2004*

Questions
- Platform
- Metadata scheme
- Workflow
- DOI deposit
DSpace (library view)

Open source platform
Focused on long-term preservation, using the CNRI Handle System
Supports qualified Dublin Core Metadata
Supports OAI PMH (open archives initiative protocol for metadata harvest)
Full administrative capabilities; decisions can be made at the institutional & departmental levels
DSpace (user view)

Web user interface

Designed to handle “all manner of digital formats”

Minimal metadata requirements for input
Growing the Woods Hole Open Access Server, WHOAS

*KISS (keep it simple)*

- For DSpace & IT
  - Minimum customization
  - No software development
  - Migration to new versions only as necessary
Growing WHOAS

KISS

- For researchers
  - Content “as is”
  - Mediated: library staff doing the input
- For library staff
  - No authority control for author names
    - As they appear in the work
  - No authority control for subject terms
    - Abstract: lots and lots of words
Content growth

Total metadata records
Content analysis

Snapshot [929 metadata records ~ 06/01/06]:

Woods Hole content

Technical Reports, Working Papers, and Theses: 276
Articles (published version): 200
Articles (preprint): 140
Books: 19
Presentations/Other: 6
Data sets: 2

IAMSLIC content: 286
Content recruitment

**MBL & WHOI authors have published:**
- 1250* articles Jan 2004 - date
- 725* articles Jan 2005 - date
- *in approximately 380 science, technology, and medical journals*

*represents 85% of known Woods Hole science publications, many in journals that permit self-archiving in IRs of pre-print/post-print (e.g., AGU, Blackwell, Elsevier, Springer)
Recruitment techniques

*Push vs. Pull*

**Push (voluntary contributions):**

<3% of total WH content in WHOAS

*Report on the NIH Public Access Policy, Jan. 2006*

<4% of the total number of articles estimated to be eligible have been added to PMC

**Pull (harvest):**

Harvest: 59% of WH total articles in WHOAS
Pulling content from authors

Pull (solicit):
Library staff:
  Identify
  Verify
  Solicit
  Load
Recruitment analysis

Solicit: 27% → 40% response rate
Additional content acquisition

WHOI Technical Reports
3500+ published since 1941
initial focus: 1990-

WHOI and MIT/WHOI Joint Program theses
approximately 1150 published since 1968
initial focus: 1990-
“Author hereby grants to MIT and WHOI permission to reproduce …”

“On demand”
- Belukha Whale
- Indian Ocean Bubble
- Papers on Physical Oceanography and Meteorology
Implementing DOI deposit

Help expose content to its widest potential audience.
- DOI added to 024 †a in MARC record in our local system, Mariner
- 856 field added
- holding record added
- tbd: update OCLC records?

Broad acceptance of DOI in the STM publishing community

Facilitates linkage to our objects by STM publishers
DOI deposit

DOIs deposited for “original” works:
- WHOI Technical Reports and Technical Memoranda
- WHOI and MIT/WHOI theses
- Digitized books “owned” by WH science community, out of copyright, and/or where we have been granted permission

DOIs are not deposited for:
- articles ~ preprints or published version
Lessons learned

*Involve library and IT colleagues, early and often*

*Be flexible*

*Be patient*

*Be persistent*

*Expect the unexpected*

*Be prepared to un-do and re-do*

*Know your limitations*

*Mistakes are okay*

*Look forward*
Looking ahead

More of the same

Conversations regarding the value of an IR
  • “Concentrates” the intellectual product
  • Archiving scholarly communication

Conversations regarding the value of open access
  • Need to re-shape scholarly communication
  • Champion for authors’ rights

Outreach to recruit content from:
  SEA
  USGS
  WHRC
URLs of interest

**DSpace:**
http://www.dspace.org/

**WHOAS:**
https://darchive.mblwholibrary.org/index.jsp

**MBLWHOI Library:**
http://www.mblwholibrary.org/

**Amendment to Publication agreement:**
http://www.mblwholibrary.org/services/copyright/

**Journal policies – self archiving:**
http://romeo.eprints.org/
Questions?

Ann Devenish
Publishing Services
WHOAS Project Manager
MBLWHOI Library
Woods Hole, MA, USA
adevenish@whoi.edu
508.289.2865