

**FINAL REPORT OF  
THE  
AQUATIC COMMONS  
IMPLEMENTATION TASK FORCE**

**Submitted by:**

**Stephanie C. Haas, Chair  
Ann Devenish, Jean Collins,  
Snowdy Dodson, Peter Pissierssens,  
Pauline Simpson, Simon Wilkinson,  
Steve Watkins**

**Prepared for the IAMSLIC Executive Board, August 24, 2006**

**CONTENT (pages)**

|  |
|--|
| <b>Report (1- 5)</b>                                   |
| <b>Appendix 1. WHOI Community Proposal (6-14)</b>      |
| <b>                  WHOI Server Proposal (15- 22)</b> |
| <b>Appendix 2. IOC Proposal (23–27)</b>                |
| <b>Appendix 3. FCLA Proposal (28–36)</b>               |
| <b>Appendix 4. Technical criteria (37-45)</b>          |
| <b>Appendix 5. Overall criteria (46–48)</b>            |
| <b>Appendix 6. Cost Comparisons (49-50)</b>            |

This report contains a summary of the discussions and recommendations of the Aquatic Commons Implementation Task Force. The charge of the Task Force was to:

- Implement the "Aquatic Commons" in the most economical fashion for IAMSLIC and make the best use of existing resources (i.e. investigate existing repositories among members and partners like IOC)
- Ensure that we end up with an IAMSLIC-owned product.
- Come up with a plan for raising additional funds (and in the best-case go after donors and grant opportunities themselves).

As initially presented at the 2005 conference in Rome, the Aquatic Commons Model consisted of:

an Aquatic Commons repository,  
existing OAI-capable repositories,  
a harvester,  
an OAI data provider (to interface with the zebra server),  
a search and retrieval interface,  
a database, and  
a zebra Z39.50 server to interface with the IAMSLIC Z39.50 Distributed Library

The current document focuses on the Task Force's work on the repository piece of the model only. This repository is intended for use by marine and aquatic researchers and institutions that do not have electronic infrastructure support at the local level.

The implementation of the complete model was not considered; however, some members expressed the desire to have all of the model components administered by one institution. This institution should have demonstrable expertise and experience in developing and maintaining the various components of the commons harvester/repository/z.39.50 server etc. Also, this institution should have experience in preservation, archiving and migration of digital publications. Initial explorations of open source harvesters were undertaken by Steve Watkins, but professional responsibilities prevented Steve from concluding his investigations and he has no recommendations concerning harvester applications at this time.

Although the RFP was sent to the IAMSLIC list, to date only Task Force members have submitted repository proposals, Discussions included extensive comparisons between the capabilities and sustainability of the hosting institution, software packages, costs, and branding. The complete proposals are contained in Appendix 1: Woods Hole, Appendix 2: IOC; and Appendix 3: FCLA. Woods Hole submitted two proposals: one to create an Aquatic Commons community similar to the current IAMSLIC community which serves the proceedings and newsletters; the other proposal was for the purchase and programming of a separate Aquatic Commons server. Appendix 4 is the technical comparison of functionality between the two software packages; Appendix 5 is a more general comparison of the hosting institution stability, technical support for the repository, and costs; and Appendix 6 contains the final comments from five of the Task Force members.

In addition to the Task Force's review of the proposals, IAMSLIC needs to be cognizant of the trends in the information world. Academic Research Libraries (ARL) [<http://www.arl.org/spec/SPEC292web.pdf>] surveyed 123 large research libraries in the U.S. and Canada about institutional repositories. Of the 53 respondents who identified software, 37 used DSpace. According to the Registry of Open Access Repositories, world wide there are 167 repositories using DSpace and 210 using EPrints. [Accessed 8/20/2006 at <http://archives.eprints.org/>] A current survey of IAMSLIC/EURASLIC member organizations had 15 respondents: 10 use DSpace, 2 EPrints, 1 Fedora, and 1 DigitalCommons.

The learning curve and cost of implementing a chosen software precluded any interest in changing software packages. The figures suggested by Woods Hole for bringing up a new dedicated DSpace server for IAMSLIC are estimated to exceed \$13,000. The average start up costs for repositories given in the ARL study was \$182,500 with an average operating budget of \$113,500. Cost comparisons for all the proposals are given in Appendix 7.

In terms of costs, the least expensive approach is the other Woods Hole's proposal to create an Aquatic Commons subcommunity under IAMSLIC. In turn, each participating agency would become a subcommunity with its own collections. (Please see page 6 of Appendix 1\_WHOICommunity.pdf ) There are no start-up costs for this option, but after the first 50 gigabytes of storage (approximately 16,000 pdfs) additional storage would cost \$100 per gigabyte. There would be a \$500 per year maintenance fee and \$1,000 for the project manager to travel to the IAMSLIC annual meeting for training/administrative purposes. For either proposal, Woods Hole would also offer digitization services at prescribed fees.

The drawback to this proposal is that while IAMSLIC would own the product, the subcommunity level proposed would not give IAMSLIC or the contributing agencies a clearly recognizable branding. This recognition factor was of concern to several Task Force members.

The IOC proposal would create a separate DSpace implementation on an existing server. Advantages to this arrangement are that metadata structures and functionalities can be standardized for all of the repositories: OdinPubCARSA, OdinPubAFRICA, and the proposed IAMSLIC repository, yet individualized branding can be maintained.

The start up costs would be \$4,000 which includes adapting training materials for IAMSLIC. The yearly maintenance cost would be \$1,000. Marc Goovaerts and his staff, University of Hasselt, Belgium would be responsible for the development and deployment of the Aquatic Commons site. Marc has been involved with both ODINPub projects and has considerable expertise in the setup of DSpace repositories. Because all of the ODINPub projects have involved multiple contributing partners from countries throughout the world, the knowledge of the non-technical aspects of repository building, e.g., training, solicitation of items, etc. must be considered of high potential value in this proposal.

The FCLA proposal would provide an Eprint repository. The most attractive aspects of EPrints are ease of implementation and use, the existence of an active user group and development team, and its ability to do full text searching on appropriate file formats. The cost for implementing the system at FCLA would be \$5,700 for the first year; addition years would cost \$2,351 with a built in server replacement allocation, or \$1,750 without the server replacement. In terms of the hosting institution, FCLA is stable and responsible for building collaborative technology-based systems for Florida's public academic libraries including online catalogs, database loading and maintenance, and the creation and serving of digital collections. It has a large staff of programmers and IT people and a policy of forward migration of technologies. It is in the process of creating one of the first true digital archives in the United States.

In assessing the various proposals, the Task Force focused on cost, branding, hosting institution stability, technical capacity, software, and services provided. Depending on the weight of each factor, the proposal advantages change. If training is a key element, FCLA will not provide that except as it relates to the technical implementation of the software. If technical capacity is of primary concern, FCLA has access to a staff of 23 programmers. While four of the Task Force members mentioned a personal preference for EPrints because of ease of use, there is a new version of DSpace in release. Because the Executive Board has more information on the whole gamut of future digital initiatives of IAMSLIC, the Task Force has not formally voted on the proposals, except to eliminate the WHOI proposal that would develop a new DSpace server from scratch.

We are recommending that the Executive Board review Appendix 1: WoodsHole\_community proposal, Appendix 2: IOC proposal, and Appendix 3: FCLA proposal, Appendix 5: Overall criteria chart, and Appendix 6: Task Force member comments.

While the issues of technology and cost are the most apparent at first glance, the true cost of this initiative does not lie there, but rather in the ability of IAMSLIC to recruit and train collaborators to enter content. As we have pointed out above, IOC does have experience in this area and will adopt its training materials, and the project manager at WHOI will come to IAMSLIC to train, but the Task Force members are keenly aware that this will not be enough. Truthfully, the cost of successful implementation probably is not covered in this document except in the comments of individuals. The true cost has to do with content recruitment and that will be dependent on the active solicitation of documents. IAMSLIC members will need to contribute not only their own documents, but identify and request others to submit, and at times, to help digitize legacy materials and create metadata.

In light of the many factors involved in creating a successful repository, the Task Force suggests:

### **Recommendation 1: Establishment of an Aquatic Commons Board**

The Aquatics Commons Board should be established with the express responsibility of making policy and management decisions related to all aspects of coverage, funding of content creation, training, intellectual property rights and access rights. The Board will be responsible for creating working groups as needed to fulfill

repository functions such as soliciting content, working with the hosting institution to create appropriately configured submittal, search, and display interfaces, reviewing submissions of metadata and accompanying digital formats, creating training materials and developing a core of trainers. The Board members should represent the major stakeholders.

**Recommendation 2: IAMSLIC's existing Memorandums of Understanding with FAO and IOC should be used to develop the infrastructure for the Aquatic Commons repository including a base level of funding.**

In light of the existing Memorandums of Understanding between IAMSLIC/IOC and IAMSLIC/FAO, the Task Force believes this project represents an excellent opportunity to share expertise and expenses to create a global resource that transcends the boundaries of each organization. By sharing in the costs and the Board governance, each organization stands to benefit and to synergistically support aquatic/marine researchers throughout the world.

Beyond base funding, each organization should seek the most efficient means of helping their member communities contribute to this initiative. Local needs would dictate what type of help could be offered.

**Conclusion**

The Task Force has solicited appropriate proposals for building an Aquatic Commons repository that will be economically feasible and have clear IAMSLIC ownership. The Memorandums of Understanding offer a means of governance and funding for the repository. The Task Force believes that the minimal amount requested for the creation and maintenance of the repository does not warrant outside funding, but that funding to help build content will be needed.



**Ann Devenish**  
**WHOAS Project Manager**  
**MBLWHOI Library**  
**WHOI - MS 8**  
**Woods Hole, Massachusetts 02543**  
**Telephone: (508) 289-2865**  
**Fax: (508) 457-2183**

|   |   |
|---|---|
| The IAMSLIC Aquatic Commons and WHOAS ..... | 2 |
| About WHOAS .....                           | 2 |
| Branding.....                               | 3 |
| User Authorizations .....                   | 3 |
| Simplified Workflow .....                   | 4 |
| Current Organization .....                  | 5 |
| Collections .....                           | 5 |
| Recommended Organization.....               | 6 |
| Content Acquisition .....                   | 7 |
| Scanning.....                               | 7 |
| Content delivery.....                       | 7 |
| Estimated costs.....                        | 7 |
| Budget.....                                 | 8 |
| Notes .....                                 | 8 |
| Frequently asked questions:.....            | 9 |

## **The IAMSLIC Aquatic Commons and WHOAS**

As the host site for the IAMSLIC Archives, the IAMSLIC web server, and the digital IAMSLIC conference proceedings and newsletters, the following is an invitation to host portions of the Aquatic Commons in the Woods Hole Open Access Server ([WHOAS](#)) of the MBLWHOI Library. WHOAS will host digital content contributed by IAMSLIC and related organizations not currently hosting their own digital repository.

There will be no storage fees for the Aquatic Commons content in WHOAS up to and including 50 gigabytes of storage (approximately 16,000 pdfs.) Additional storage may be acquired at a cost of \$100 per gigabyte.

An annual fee of \$500 will be charged for services related to database back-up and maintenance.

### ***About WHOAS***

WHOAS is an open access institutional repository (IR) hosted by the MBLWHOI Library. All content accepted into WHOAS is open to all visitors and remains in the IR in perpetuity. The WHOAS platform is OAI PMH compliant and metadata records are available for routine harvest by one or more OAI harvesters, including the (proposed) Aquatic Commons metadata harvester. WHOAS is available 24/7/365.

WHOAS utilizes the DSpace platform and is organized into Communities, Sub-communities and Collections. Each Community represents an organization or science community (e.g., IAMSLIC). Each Sub-community may represent a subdivision of the organization (e.g., Aquatic Commons). Each Collection represents a group of digital objects brought together around one or multiple commonalities. Within a given Community there can be an unlimited number of Sub-communities. Within a given Sub-community there can be an unlimited number of Sub-communities. Within a given Community and/or Sub-community, there can be an unlimited number of Collections. Within a Collection there can be an unlimited number of metadata records and their associated digital objects. Each metadata record created in WHOAS has a unique and persistent identifier, known as a Handle.

Associated with a single metadata record there can be one or more digital objects. An example of multiple objects associated with a single record is individual book chapters comprising the complete monograph described in the metadata record.

The language of WHOAS is English. Digital objects' text may be rendered in a non-English language; the web user interface for content intake and approval, metadata field labels, and help screens are authored in English.

The MBLWHOI Library is a member of [CrossRef](#) and is enabled to deposit DOIs for appropriate content within the IR.

## **Branding**

The home page of WHOAS carries the branding of the MBLWHOI Library as its header. All pages include DSpace navigation links on the left side of the screen. In addition to the MBLWHOI Library header, each Community, Sub-community, and Collection home page may carry additional information, including introductory text, side bar text, copyright text, and/or a logo. The item record page, containing the metadata and link to the associated digital object carries no branding.

## **User Authorizations**

WHOAS supports multiple levels of authorization in order to access its content.

### *Anonymous:*

- These users may view any metadata record and its associated objects within the IR.
- No log in or registration is required.

### *Email notification:*

- These users may view any metadata record and its associated objects within the IR.
- These users have self-selected to be notified when any new content is added to one or more Collections.
- Registration is required; login is required to select/deselect Collection(s) for email notification.

### *Submit:*

- These users may view any metadata record and its associated objects within the IR.
- These users may submit content to one or more Collections, thereby creating a metadata record via the Intake screens (author, title, abstract, keywords, etc.)
- Registration is required; login is required to access the Intake (submit) screens.

### *Edit/Approve:*

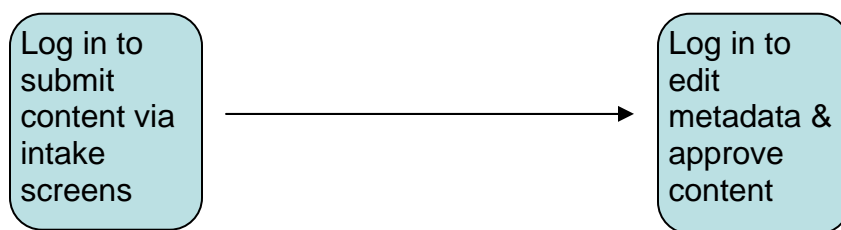
- These users may view any metadata record and its associated objects within the IR.
- These users may review submitted content in one or more Collections, edit the metadata, and approve the content for acceptance into one or more Collections in the IR.
- Registration is required; login is required to access submitted content and metadata and to approve content.

### *Administer:*

- These users may view any metadata record and its associated objects within the IR.
- These users may review accepted content in one or more Collections and edit the metadata within the IR.
- These users may create and/or edit item templates in one or more Collections within the IR.
- Registration is required; login is required to edit metadata and item templates.



## ***Simplified Workflow***



### *Submit*

- ~ Completing the intake screens creates descriptive metadata and some administrative metadata.
- ~ Uploading files generates an email message to Collection editor(s).

### *Edit/Approve*

- ~ Approving the content creates additional administrative metadata, assigns the Handle, generates Email notifications to appropriate users, and makes the content available to all users.

## Current Organization

At present, IAMSLIC is a Community within WHOAS with no Sub-communities and 23 Collections. The Collections represent the 22 Annual conferences for which there exist published proceedings, 1984 to date, and the newsletters published March 1995 to date. All of the content was originally published by IAMSLIC.

Input of content is the responsibility of three authorized content submitters. The edit/approve function is the responsibility of the IAMSLIC web master. The web master also administers all IAMSLIC Collections. These “e-persons” have reliable internet connectivity.

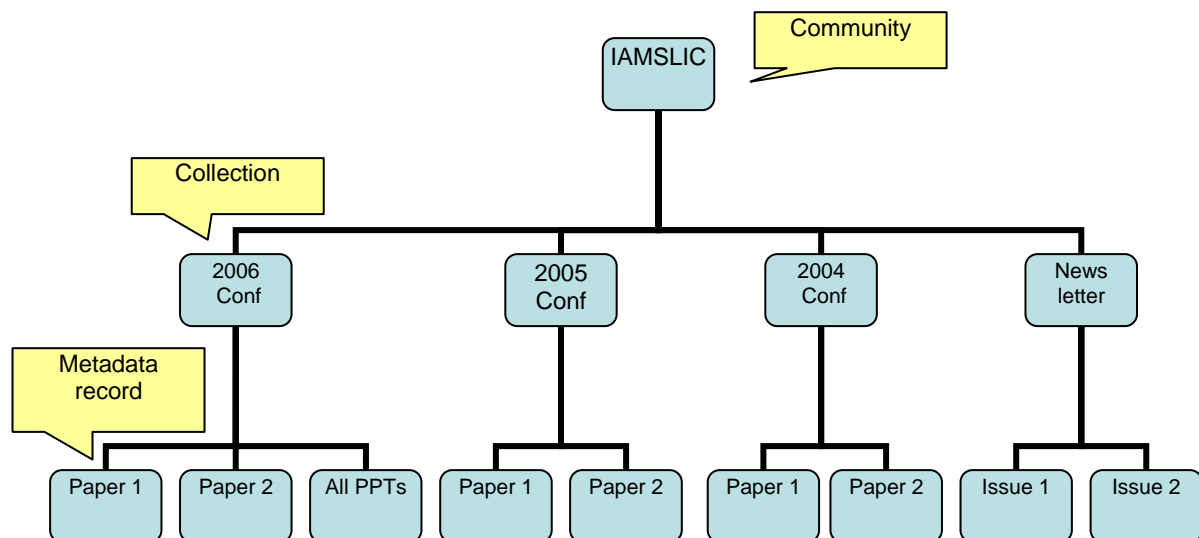
## Collections

*IAMSLIC Newsletter:* Within this Collection are multiple metadata records, each of which represents an issue of the newsletter.

- Each issue has its own metadata record and may consist of multiple digital objects, for example pdf and html versions of the same issue.

*IAMSLIC conferences:* There is Collection for each conference, i.e., 2005 Conference Proceedings, 2004 Conference Proceedings, etc.

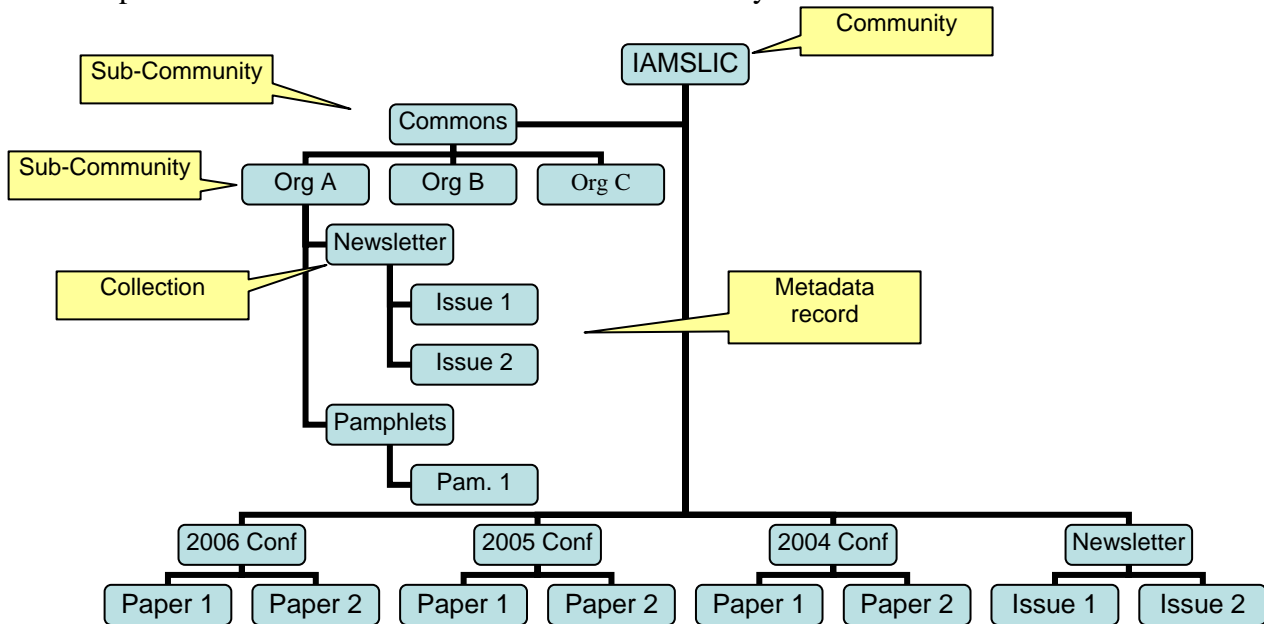
- Each Collection contains multiple metadata records; one metadata record with one associated object, representing a single paper.



## Recommended Organization

The current IAMSLIC Community and Collections will be retained. It is recommended that a Sub-community be created under the IAMSLIC Community to accommodate Aquatic Commons content. The Sub-community, e.g., Aquatic Commons, will have its own home page with appropriate “branding,” subject to the conditions above. Within the Sub-Community there may be one or more Sub-communities and Collections based, for example, on content contributor (Source A, Library B, Organization C, etc.) and content type (serials, books, images, etc)

Each Collection will require a Collection administrator. Each Collection will require at least one content submitter. Each Collection will each require at least one editor/approver, responsible for reviewing the metadata and approving the inclusion of content in the Collection. It is possible that the role of Collection administrator, content submitter and content editor/approver can rest with the same individual. It is possible that one individual may have responsibilities in one or more Collections. These “e-persons” should have reliable internet connectivity.



## **Content Acquisition**

The acquisition of content into the current IAMSLIC Collections is facilitated when the items already exist as digital objects; they were either born digital or have been previously rendered as PDFs. For a portion of the Aquatic Commons community wishing to place content into the Collections of the Sub-community, providing digital objects will not be an issue. DSpace was selected as the WHOAS platform because it supports a wide range of formats including: PDF, Word, Access, Excel, PowerPoint, LaTeX, XML, HTML, JPEG, GIF, TIFF, and WAV. When the contributor has reliable internet connects, content submittal into WHOAS has few obstacles.

For a portion of the Aquatic Commons community, one or both of the following issues may need to be considered: the rendering of print matter into digital files, and the delivery of the digital content to a content submitter.

## **Scanning**

The MBL Digital Processing Center (DPC) offers complete photocopying and high quality color and grayscale digital imaging.

Utilizing a Canon 3300 copier with eCopy software, the DPC scans print documents into multiple TIFF files, which are then converted to searchable PDFs. Color images are scanned separately and inserted in its correct location. Following quality assurance review by DPC staff, the files are delivered to DPC's secure FTP site. There is also the option of requesting these files be copied to archival CD-Rs.

## **Content delivery**

Once the PDFs are placed on the FTP server, the Aquatic Commons e-person submitter is notified via email that the files are ready for download. When the e-person advises that the files are acceptable and have been downloaded to a local workstation, the DPC will remove the files from the FTP site and destroy the print original. If archival CDs are requested, they can be delivered to the originating location or placed in the IAMSLIC archive hosted at the MBLWHOI Library.

Once the PDFs have been acquired, the content submitter may proceed to upload content using the DSpace web user interface; the editor may then proceed to approve the content.

## **Estimated costs**

Unit cost (per page scanned): \$0.40

Unit cost (per 650mb archival CD-R): \$10.00

Postage charges for delivering CD-R: current market rate

## **Budget**

|  |                     |
|--|---------------------|
| Purchase of gigabyte storage per gigabyte  | \$100.00            |
| Annual maintenance fee   | \$500.00            |
| IAMSLIC Annual Conference/training session Project Manager (registration & travel) | \$1000.00           |
| Unit cost (per page scanned)   | \$0.40              |
| Unit cost (per 650mb archival CD-R)  | \$10.00             |
| Postage charges for delivering CD-R  | Current market rate |

## **Notes**

Role of WHOAS project manager:

- Create Community, Sub-communities and Collections as appropriate.
- Authorize users to administer, submit, edit, and/or approve content as appropriate; delete authorizations upon changes in responsibilities.
- Provide training and support at the IAMSLIC Annual conference and through email.
- Deposit DOIs w/CrossRef for appropriate content.

To be determined by IAMLSIC:

- What content, if any, is to have DOI deposit?
  - Recommended: books, technical reports, theses, working papers.
- Will non-English content have English language descriptive text in the metadata?
- What will be the input guidelines, i.e., standardized metadata for similar content types.
- What will be the workflow considerations, including but not limited to:
  - Multiple digital objects linked to a single metadata record must be approved for inclusion at the same time.
- Who will be the Collection Administrators, content submitters, and content editors?

Other:

- Newer versions of DSpace permit active url linking from metadata records.
- Future versions of DSpace may permit greater flexibility in branding at the Community, Sub-Community, and/or Collection levels.
- Options for changing the fields that display in the simple item record are to be explored; e.g., are they customizable at the Collections level?
- Options for changing the order of the fields displayed in the full item record are to be explored, e.g., are they customizable at the Collections level?

## ***Frequently asked questions:***

*How long will content be retained?*

It is our intent to maintain contributions in perpetuity.

*Can contributions to WHOAS be withdrawn?*

Once placed in WHOAS, it is not customary to remove content, however, in special circumstances access to the content may be limited.

*Is content limited by format?*

DSpace was selected as the WHOAS platform because it supports a wide range of formats including: PDF, Word, Access, Excel, PowerPoint, LaTeX, XML, HTML, JPEG, GIF, TIFF, and WAV.

*Is content limited by file size?*

The larger the file, the longer it can take to “intake” (upload) into WHOAS. Files beyond 30-40mb can take an exceedingly long time to load into WHOAS and files beyond 50mb may not load at all. To make arrangements to load large files, please contact Ann Devenish, WHOAS project manager, at [adevenish@whoi.edu](mailto:adevenish@whoi.edu), or 508.289.2865.

*If joint author(s) placed this content in their institutional repository, should the same content be in WHOAS?*

Yes! In that policies, access, and competencies can vary from IR to IR, redundancy is a bonus.

*Can the content in WHOAS be easily linked to?*

Yes! Each metadata record and its associated objects in WHOAS are assigned a unique and persistent identifier, known as a Handle. Once assigned, the Handle will not change and the links will remain intact.

*How to get started?*

In order to be authorized as a content submitter or to edit metadata and approve content, users need to register at [WHOAS](#). Select the link to [MyWHOAS](#) and register as a new user. The WHOAS project manager will assign the appropriate authorizations for the appropriate Collections.



**Ann Devenish**  
**WHOAS Project Manager**  
**MBLWHOI Library**  
**WHOI - MS 8**  
**Woods Hole, Massachusetts 02543**  
**Telephone: (508) 289-2865**  
**Fax: (508) 457-2183**

The IAMS LIC Aquatic Commons ..... 2  
  About DSpace ..... 2  
  Branding..... 3  
  User Authorizations ..... 3  
  Recommended Organization..... 5  
  Content Acquisition ..... 6  
    Scanning..... 6  
    Content upload ..... 6  
    Estimated costs..... 6  
  Budget ..... 7  
  Frequently asked questions: ..... 8

## **The IAMSLIC Aquatic Commons**

As the host site for the IAMSLIC Archives, the IAMSLIC web server, and the digital IAMSLIC conference proceedings and newsletters, the following is a proposal to establish and host a DSpace installation to host the Aquatic Commons at the MBLWHOI Library. The installation will host digital content contributed by IAMSLIC and related organizations not currently hosting their own digital repository.

### ***About DSpace***

DSpace is an open source institutional repository platform developed by M.I.T. and Hewlett Packard. DSpace is OAI PMH compliant and metadata records can be made available for routine harvest by one or more OA harvesters, including the (proposed) Aquatic Commons metadata harvester. DSpace is designed to handle “all manner of digital objects,” uses the CNRI Handle System to create persistent URLs, and supports qualified Dublin Core metadata.

DSpace is organized into Communities, Sub-communities and Collections. Each Community may represent an organization or science community (e.g., Aquatic Commons). Each Sub-community may represent a subdivision of the organization (e.g., NACA Contributions). Each Collection represents a group of digital objects brought together around one or multiple commonalities (serial title, books, images, etc.). Within a given Community there can be an unlimited number of Sub-communities. Within a given Sub-community there can be an unlimited number of Sub-communities. Within a given Community and/or Sub-community, there can be an unlimited number of Collections. Within a Collection there can be an unlimited number of metadata records and their associated digital objects.

Each metadata record created has a unique and persistent identifier, known as a Handle. Associated with a single metadata record there can be one or more digital objects. An example of multiple objects associated with a single record is individual book chapters comprising the complete monograph described in the metadata record.



## ***Branding***

The home page of the Aquatic Commons institutional repository may carry the branding of the Aquatic Commons as its header. All pages include DSpace navigation links on the left side of the screen. In addition to the header, each Community, Sub-community, and Collection home page may carry additional information, including introductory text, side bar text, copyright text, and/or a logo. The item record page, containing the metadata and link to the associated digital object carries no branding.

## ***User Authorizations***

DSpace supports multiple levels of authorization in order to access its content.

### *Anonymous:*

- These users may view any metadata record and its associated objects within the IR.
- No log in or registration is required.

### *Email notification:*

- These users may view any metadata record and its associated objects within the IR.
- These users have self-selected to be notified when any new content is added to one or more Collections.
- Registration is required; login is required to select/deselect Collection(s) for email notification.

### *Submit:*

- These users may view any metadata record and its associated objects within the IR.
- These users may submit content to one or more Collections, thereby creating a metadata record via the Intake screens (author, title, abstract, keywords, etc.)
- Registration is required; login is required to access the Intake (submit) screens.

### *Edit/Approve:*

- These users may view any metadata record and its associated objects within the IR.
- These users may review submitted content in one or more Collections, edit the metadata, and approve the content for acceptance into one or more Collections in the IR.
- Registration is required; login is required to access submitted content and metadata and to approve content.

*Administer:*

- These users may view any metadata record and its associated objects within the IR.
- These users may review accepted content in one or more Collections and edit the metadata within the IR.
- These users may create and/or edit item templates in one or more Collections within the IR.
- Registration is required; login is required to access content, metadata, and item templates.

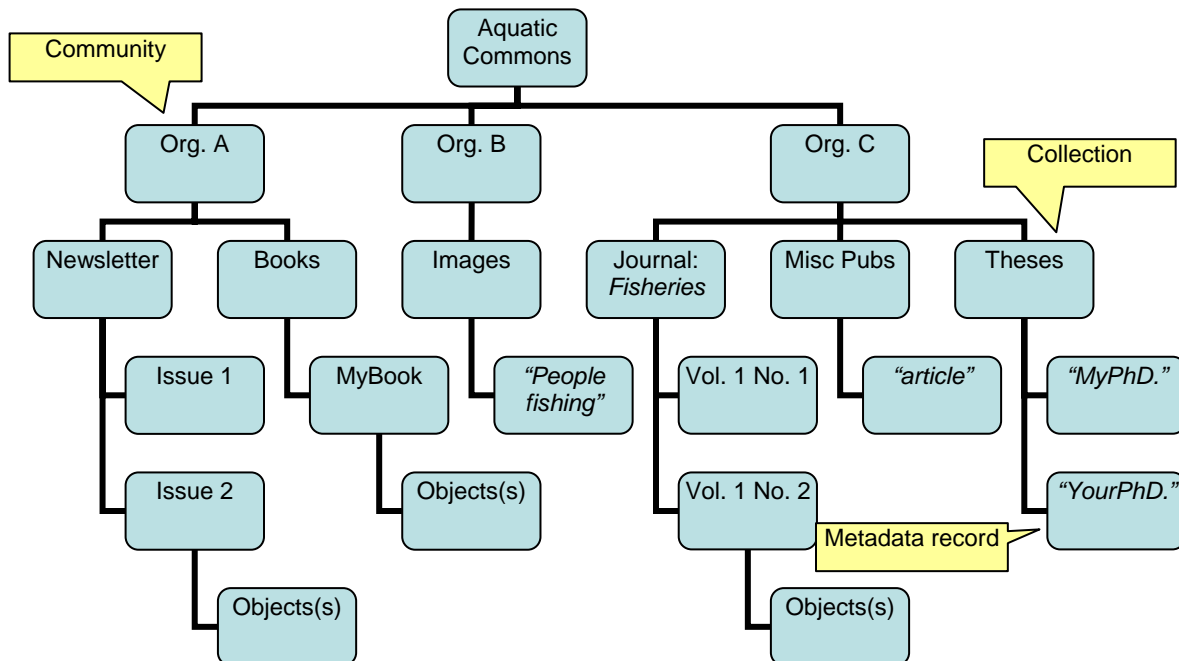
*IR Administrator:*

- These users have all rights, including authorization of Collection administrators, content submitters, and content editors; creation of Communities, Sub-communities and Collections.
- These users may review accepted content in all Collections and edit the metadata.
- These users may create and/or edit item templates in all Collections within the IR.
- Registration is required; login is required to access content, metadata, item templates, and assign authorizations.

## Recommended Organization

It is recommended that a Community be created for each organization to be represented in the Aquatic Commons. Each Community will have its own home page with appropriate “branding,” subject to the conditions above. For each Community there may be one or more Sub-communities and Collections based, for example, on content contributor (Source A, Library B, Organization C, etc.) and content type (serials, books, images, etc)

The Aquatic Commons will require an IR Administrator. The IR Administrator may also serve as Collection administrator or each Collection may have its own e-person in this role. Each Collection will require at least one content submitter. Each Collection will each require at least one editor/approver, responsible for reviewing the metadata and approving the inclusion of content in the Collection. It is possible that the role of content submitter and content editor/approver can rest with the same individual. It is possible that one individual may have responsibilities in one or more Collections. These “e-persons” should have reliable internet connectivity.



## **Content Acquisition**

The acquisition of content into the Aquatic Commons will be facilitated when the items already exist as digital objects; they were either born digital or have been previously rendered as PDFs. For a portion of the Aquatic Commons community wishing to place content into the Collections, providing digital objects will not be an issue. DSpace was selected as the WHOAS platform because it supports a wide range of formats including: PDF, Word, Access, Excel, PowerPoint, LaTeX, XML, HTML, JPEG, GIF, TIFF, and WAV. When the contributor has reliable internet connectivity, content submittal into the Aquatic Commons will have few obstacles.

For a portion of the Aquatic Commons community, one or both of the following issues may need to be considered: the rendering of print matter into digital files, and the delivery of the digital content to a content submitter.

## **Scanning**

The MBL Digital Processing Center (DPC) offers complete photocopying and high quality color and grayscale digital imaging.

Utilizing a Canon 3300 copier with eCopy software, the DPC scans print documents into multiple TIFF files, which are then converted to searchable PDFs. Color images are scanned separately and inserted in its correct location. Following quality assurance review by DPC staff, the files are delivered to DPC's secure FTP site. There is also the option of requesting these files be copied to archival CD-Rs.

## **Content upload**

Once the PDFs are placed on the FTP server, the Aquatic Commons e-person submitter is notified via email that the files are ready for download. When the e-person advises that the files are acceptable and have been downloaded to a local workstation, the DPC will remove the files from the FTP site and destroy the print original. If archival CDs are requested, they can be delivered to the originating location or placed in the IAMSLIC archive hosted at the MBLWHOI Library.

Once the PDFs have been acquired, the content submitter may proceed to upload content using the DSpace web user interface; the editor may then proceed to approve the content.

## **Estimated costs**

Unit cost (per page scanned): \$0.40

Unit cost (per 650mb archival CD-R): \$10.00

Postage charges for delivering CD-R: current market rate

## **Budget**

|  |                     |
|--|---------------------|
| Purchase DSpace server, including Handle server    | \$5,000.00          |
| Install and configure DSpace, 15 hours @ \$75.00   | \$1,125.00          |
| MBL IT programmers, 100 hours @ \$75.00            | \$7,500.00          |
| Consulting w/ WHOAS project manager                | \$1,000.00          |
| Back-up services for database                      | \$500.00 per annum  |
| Scanning services:<br>Unit cost (per page scanned) | \$0.40              |
| Unit cost (per 650mb archival CD-R)                | \$10.00             |
| Postage charges for delivering CD-R                | Current market rate |

## **Notes**

Role of WHOAS project manager:

- Deposit DOIs w/CrossRef for appropriate content.

To be determined by IAMLSIC:

- What content, if any, is to have DOI deposit?
  - Recommended: books, technical reports, theses, working papers.
- Will non-English content have English language descriptive text in the metadata?
- Input guidelines, i.e., standardized metadata for similar content types.
- Workflow considerations, including but not limited to:
  - Multiple digital objects linked to a single metadata record must be approved for inclusion at the same time.
- Who will be the IR Administrator, Collection Administrators, content submitters, and content editors?

## **Other:**

- New versions of DSpace permit active url linking from metadata records.
- Future versions of DSpace may permit greater flexibility in branding at the Community, Sub-Community, and/or Collection levels.
- Options for changing the fields that display in the simple item record are to be explored; e.g. are they customizable at the Collections level?
- Options for changing the order of the fields displayed in the full item record are to be explored, e.g. are they customizable at the Collections level?

## ***Frequently asked questions:***

*Is content limited by format?*

DSpace supports a wide range of formats including: PDF, Word, Access, Excel, PowerPoint, LaTeX, XML, HTML, JPEG, GIF, TIFF, and WAV.

*Is content limited by file size?*

The larger the file, the longer it can take to “intake” (upload) into DSpace. Files beyond 30-40mb can take an exceedingly long time to load and files beyond 50mb may not load at all. To make arrangements to load large files, please contact Ann Devenish, WHOAS project manager, at [adevenish@whoi.edu](mailto:adevenish@whoi.edu), or 508.289.2865.

*Can the content in DSpace be easily linked to?*

Yes! Each metadata record and its associated objects are assigned a unique and persistent identifier, known as a Handle. Once assigned, the Handle will not change and the links will remain intact.

# PROPOSAL OF THE IODE PROJECT OFFICE TO HOST THE AQUATIC COMMONS E- REPOSITORY FOR IAMSLIC

## ***Aquatic Commons***

Based on the experience with the ODINAFRICA repository ODINPubAFRICA (<http://iodeweb1.vliz.be/odin>), several other regions have recommended the development of similar e-repositories. Examples are ODINCARSA: ODINPubCARSA is ready to start (<http://doelib.uhasselt.be/odincarsa> - to be moved to the iodeweb-server after the implementation phase).

IODE has chosen DSpace as its standard repository software.

The Aquatic Commons repository can be hosted on the same server as a separate DSpace installation. The advantages are that the metadata structure and functionalities of the different repositories can be standardized while the branding still can be personalized. Also the further developments planned for the Odin repositories can easily be implemented in the Aquatic Commons Repository.

## ***About DSpace***

DSpace is an open source institutional repository platform developed by M.I.T. and Hewlett Packard. It is based on Linux and uses Java http server/servlet technology (Tomcat, Jetty, ..) and a SQL database (Postgresql or Oracle) .

DSpace supports qualified Dublin Core metadata. The announced version (1.4) will make it possible to support other metadata types. DSpace is OAI PMH compliant and metadata records can be made available in different formats through the crosswalk system. Dublin Core is the basic OAI standard. The ODIN repositories also use a qualified Dublin Core OAI output. A specific Aquatic Commons standard can be developed, besides one for ASFA centers, AGRIS databases, ....

DSpace supports new protocols like SRU/W and LNI, which can be interesting for the development of other specific services on top of Aquatic commons (e.g. bibliographic list for authors , ....)

Specific add-ons are developed in the DSpace community. A selection of these developments are integrated in new versions of DSpace. A statistical package is already included and is now extended as a web service by the University of Minho (Portugal).

DSpace is organized into Communities, Sub-communities and Collections. Metadata records with one or more digital objects are submitted to collections, which are part of

communities or sub communities. A community can have from zero to multiple sub-community levels. Each metadata record created has a unique and persistent identifier, known as a Handle (CNRI Handle System).

As Open Source software, DSpace can be customized on different levels: from layout over workflow processes to database structure. The customization of Aquatic Commons and also of the ODIN projects will always be in line with the DSpace core structure. They will always be upgradeable to new versions of DSpace.

### ***Branding***

Logo and lay-out of IAMSLIC can be used for the Aquatic Commons repository. A specific style sheet can be used defining colors and letter types.

From the community level on every participating institute can use their own logo and description and adapt these pages (only if they get the authorization).

### ***User Authorizations***

Basically DSpace has 4 levels of authorizations: view – write – add – edit. These authorizations are given to user groups for the different levels of the DSpace structure: community – collection – record – attached file.

There are special groups:

*Anonymous* which only has read authorizations for the communities-collections-records-files which does not have a specific authorization policy.

*Administrator* has all the rights on all the levels of Dspace. He can manage the repository through a specific interface:

- Setting up communities and collections
- Managing the authorizations
- Delegating management tasks to community – collection administrators
- ...

Everybody can register in DSpace. But only the administrators can give registered users specific authorizations. A registered user has the possibility to set up email notification on one or more collections.

### ***Organization - proposal***

The major advantage of DSpace is the facility to create, in the repository, different entities based on the community/collection structure. But from our experience in ODINPubAFRICA and also from experience with the University of Hasselt (Belgium), it is clear that the number of levels has to be limited to a maximum of 3. For ODINPubAFRICA we have country/region – institute – collection. At University of Hasselt we have research group – collection.

Our experience with more elaborate structures is that it costs more management time (every subdivision surely for living structures like a journal) needs permanent adaptation



of the collections. It demands a permanent change of habits from the submitters, finding out the new input structure and making it very complex for the user who is searching for an article, not for a search tree.

The goal of a repository is to make documents available and to archive. Specific presentations have to be created as an external service. Therefore the metadata has to be well adapted. For a research group at University of Hasselt we used the OAI-based XML to convert through an XSLT style sheet and integrated in another format in the web page of the institute. Some UTF-8 problems have still to be fixed, but it works. See <https://doclib.uhasselt.be/dspace/handle/1942/874> - <http://www.uhasselt.be/sein/publicaties/>

An effective structure can be:

- Community – Institute
- Sub-community – Institute departments (only for large institutes)
- Collection: based on
  - Type: articles, book (+ chapters), conference papers, theses, ...
  - Content
  - Special collections: a journal, news letter, conference, ...Certainly for these collections the metadata must be granular enough to build specific services.

## ***Services and Standards***

The Aquatic Commons repository must be ready to deliver services for/as:

- Harvesters – IAMSLIC in the first place
- ASFA input center
- AGRIS – AGROVOC
- OceanPortal – OceanExpert
- Lists:
  - Bibliographic lists for authors
  - Journal – conference web pages
- ...

For these services different sort of OAI-output has to be defined. But therefore the metadata structure has to be in place.

- Dublin Core Qualified – or more refined – which fields are necessary
- Thesauri: ASFA – other ones
- Use of language

This has to be developed with all the parties involved.

*File formats - size*

DSpace supports a wide range of formats including: PDF, Word, Access, Excel, PowerPoint, LaTeX, XML, HTML, JPEG, GIF, TIFF, and WAV. For ODINPubAFRICA we limit ourselves to PDF for published documents. But the original can also be stocked , not visible, for the sake of long term archiving.

Older documents are sometimes only available in paper format. Scanning is the solution. Flexible quality levels have to be installed. But it is in the end the responsibility of the submitting institute.

Large files can be a problem to submit, surely if the institute has limited Internet access. Some limits can be proposed based on Internet capacity.

**Development and administration**

- IOC will set up the repository and deliver the technical support for the Aquatic Commons Repository. It is the task of IAMSLIC to develop a policy to attract institutes to participate in Aquatic Commons.
- IOC cannot give the administration support after the set up period. The different institutes can administer their own communities, but still there is a need for a global management. This needs further discussion.
- Interested institutes need training. Modules have been developed for ODINPubAFRICA. But still they have to be more oriented on self-learning. The content of these packages are:
  - Repository policy development
  - Guidelines for submission – administration procedures

**Activities and Budget**

|   |  |
|---|--|
| Specific Aq. Comm.DSpace installation, use of repository server at IOC-secretariat  | Offered by IOC/IODE  |
| Install and configure DSpace  | Offered by IOC/IODE  |
| Customization DSpace –programming   | \$ 3000  |
| Project management<br>- consultations with IAMSLIC/FAO/...<br>- development standards<br>- management + support during the project period | Preferably to be done by IAMSLIC members and requires further discussion |
| Back-up services for database   | Offered by IOC/IODE  |
| Adaptation of self-training package to IAMSLIC needs  | \$ 1000  |
| Extra:<br>Service development: bibliographic list – journal interface – conference interface  | To be discussed as relevant  |
| <b>TOTAL DEPLOYMENT COST</b>  | <b>US\$ 4000</b>   |

**Recurrent (maintenance cost):** US\$ 1000/year (as from 13<sup>th</sup> month after deployment and acceptance of the product by IAMSLIC)

**Period:** September 2006 – May 2007

- Sept – Oct: Requirements – standards development
- Nov – Dec: Customization DSpace
- Jan 2007: Set up Aquatic Commons Repository
- Feb – Mar: Development self-training packages
- Apr – May: Evaluation and finalization repository

**Expertise** for the development and deployment will be provided by Marc Goovaerts, University of Hasselt, Belgium.

**Hosting services** will be provided by the IOC Project Office for IODE, Oostende, Belgium

Joint Proposal to the **Executive Board** of the **International Association of Aquatic and Marine Science Libraries and Information Centers (IAMSLIC)** and to the **Advisory Board** of the **Aquatic Science and Fisheries Abstracts (ASFA)** for the development of the **Aquatic Commons Initiative**.

Written by:

Stephanie C. Haas, Assistant Director, Digital Library Center, University of Florida Libraries  
[haas@uflib.ufl.edu](mailto:haas@uflib.ufl.edu)

Reviewed by: Priscilla Caplan, Assistant Director, Digital Library Services, Florida Center for Library Automation (FCLA) [pcaplan@ufl.edu](mailto:pcaplan@ufl.edu); Pauline Simpson, National Oceanography Centre, University of Southampton, UK [ps@noc.soton.ac.uk](mailto:ps@noc.soton.ac.uk); Steve Watkins, California State University, Monterey Bay [steve\\_watkins@csumb.edu](mailto:steve_watkins@csumb.edu)

## **History**

For the last three years, IAMSLIC has acknowledged the value of institutional repositories and encouraged members to implement IRs within their own institutions. Repositories are now spreading because they provide stewardship and global access to the digital assets of institutions and organizations. Impetus for the development of IRs has also come from the Open Access and Open Archives movements. Similarly, under the guidance of the ASFA Advisory Board, the ASFA Partnership has recognized the value of IRs, and Cambridge Scientific Abstracts (CSA) has increased the harvesting of metadata for digital objects for inclusion in ASFA. For many years, the need to provide access to full text documents has been recognized by the ASFA Partnership and IAMSLIC. Technological advances and the development of IRs now make it possible to address access to digitized texts.

## **Purpose and Stakeholders**

The current proposal is intended to establish an Aquatic Commons initiative to provide a central portal to the literature in marine and aquatic sciences. This initiative will provide an Aquatics EPrint repository for the deposit of metadata and digital texts where IT support at the local level is unstable, and a harvester that will harvest and aggregate metadata from digital repositories that serve OAI compliant metadata including the Aquatics e-print repository. Because of the extensive experience of the staff at the Florida Center for Library Automation (FCLA) in developing harvesting functionality, digital collection and metadata development, and collaborative project development and management, they will be responsible for the implementation and ongoing maintenance of the Aquatic Commons model elucidated here.

The model was designed to integrate the efforts of the total community by harvesting metadata where available and by creating repository and harvesting opportunities where needed.

Many larger institutions and agencies are creating repositories related to their own missions. The IOC has created OdinPubAfrica which is aimed at collecting and serving digitized scientific and technical publications on African marine science from Ghana, Indian Ocean commissions, Kenya, Mauritania, Morocco, Mozambique, Senegla, Seychelles, Tansania, and Tunisia. Woods Hole, Scripps, Oregon Institute of Marine Biology, University of Oregon and Southampton all have digital collections related to their oceanographic/marine missions. Presently, NOAA is working on a repository, and Jan Haspelslagh is advancing a marine science repository for the

Netherlands. All of these efforts are supported by stable IT infrastructures with both adequate and ongoing financial and human resources.

To reiterate, the main goals of the Aquatic Commons are: 1) to harvest into a searchable database the metadata from the repositories that exist and those that are currently under development, and 2) to create the Aquatic EPrint repository for smaller institutions and research endeavors that do not have access to stable IT support. Existing repositories with OAI capability include Woods Hole, Scripps, Oregon Institute of Marine Biology, University of Oregon, and Southampton; metadata from additional repositories can be added to the system as they become OAI compatible. The Aquatic e-print repository will accept all aquatic related digital texts. It is hoped that an active collaboration with the FAO Fisheries Department and with the ASFA Partners will bring a particular focus to freshwater and estuarine materials from developing countries. Similarly, it is hoped that the IAMSLIC members will also use the Aquatics EPrint repository for the grey literature developed in the hidden centers and smaller academic units of their institutions.

The initial stakeholders that have been identified include:

- 1) Researchers and research institutions in the marine and aquatic sciences;
- 2) UN, International, and National ASFA partners;
- 3) CSA;
- 4) FAO ASFA Secretariat;
- 5) FAO Fisheries Department;
- 6) Other marine research agencies such as IOC, NOAA, etc.;
- 7) IAMSLIC and its affiliated regional groups; and
- 8) Florida Center for Library Automation (FCLA).

### **Justification of Need**

One of the recurrent needs expressed by users of the ASFA database (and consistently vocalized by IAMSLIC members over the years) has been the inability to obtain the grey literature documents indexed by the ASFA partners. Because most documents are now created electronically whether on a networked or stand alone computer, these electronic files offer the opportunity to readily share the science globally through repository networks.

While it is difficult to estimate the total number of digital documents that might be deposited in the Aquatics e-print repository, three individuals indicated immediate interest at the IAMSLIC Conference in Rome: Anton Immink, Communications Officer, Aquaculture and Fish Genetics Research Programme, Stirling, UK; Simon Wilkinson, NACA, Thailand; and Catalina Lopez-Alvarez, Universidad Autonoma de Baja, Mexico. As part of the summer testing of the repository software, Guillermina Cosulich, INIDEP, Argentina is actively participating in the planning.

### **System Architecture**

The components of Aquatic Commons include:

an Aquatics EPrint repository,  
existing OAI-capable repositories,  
a harvester,  
an OAI data provider (to interface with the zebra server),  
a search and retrieval interface,  
a database, and

a zebra Z39.50 server to interface with the IAMSLIC Z39.50 Distributed Library

During the summer of 2005 as part of the planning phase, FCLA implemented and tested all of the component parts. At that time the software used for the search and retrieval interface was proprietary. It was decided that if full implementation was reached that only open source software would be used in the final architecture configuration.

### **FCLA Qualifications**

The Florida Center for Library Automation provides computer services that assist Florida's university libraries in their daily operations and record keeping as well as providing students and faculty with electronic access to scholarly materials. These services include the operation of a shared Integrated Library Management System, licensing of electronic resources, and providing the repository and support for digital versions of library-owned collections. The university system of Florida consists of ten universities, one college, twelve off-campus centers, seven agricultural research and education centers, and sixty-nine county cooperative extension programs located throughout the state. Additionally, FCLA has taken on a leadership role in the development of harvesting initiatives in the state in its work with the State Library to develop Florida Electronic Library (<http://www.flelibrary.org/index.cfm>).

The staff of FCLA has extensive expertise in mounting standards based digital collections and in creating tools for the management of digital collections over time. It is one of the first U.S. based organizations to have created a viable digital preservation repository (FCLA Digital Archive <http://www.fcla.edu/digitalArchive/index.htm>) and is a recognized leader in the U.S. for its innovative and collaborative approaches to building digital resources.

Perhaps the greatest benefit of involving FCLA in this project is the assured stability of its continued performance beyond the careers of any individual staff member of FAO, ASFA, or IAMSLIC. As a legal entity, FCLA has the ability to commit to binding agreements for work performance and system support.

The Aquatic Commons development would be supervised by Priscilla Caplan as the Assistant Director for Digital Library Services.

Succinctly, the hosting requirements for the Aquatic Commons model include:

1. Unix systems administration and secure network and computer server facilities,
2. technical knowledge and programming skills needed to develop an integrated repository/harvester system,
3. ability to evaluate and implement open source software for the harvester, EPrint repository, search interface, and database functions,
4. ability to work with existing OAI-compliant data providers to integrate them into the Aquatic Commons,
5. knowledge of different metadata formats,
6. understanding of the server requirements associated with diverse digital formats, and
7. ability to work collaboratively with a diverse clientele.

## Development Phases

The development of the Aquatic Commons will occur in three phases.

### **Phase 1: Development of the Aquatics e-print Repository.**

(Estimated implementation time: 6 months from funding)

The open source EPrints software is available from the University of Southampton and has a large community of users. The FCLA staff has already loaded and done initial testing of this software. During Phase 1, a dedicated server would be purchased, installed, and the software reloaded. Because of the large number of options offered by the software, extensive configuration is possible, requiring active participation by intended contributors. All setup requirements including appropriate authority listings, preferred formatting for fields, etc. would be determined in consultation with IAMSLIC and the technical staff at ASFA. Training materials will be developed by IAMSLIC members and administrative duties associated with metadata review and object acceptance would be delineated and assigned to IAMSLIC members.

### **Phase 2: Development of the Aquatic Commons harvester, an OAI provider, a search interface, and database**

(Estimated implementation time: 8 months from funding)

FCLA will investigate the open source *Arc* software as a potential product for providing search and retrieval capability, and possibly harvesting. Arc functionality is described in an article by Liu Ziaoming and others, entitled *Arc—An OAI Service Provider for Cross—Archive Searching* [accessed 11/9/05 at <http://www.ils.unc.edu/~mln/jcdl-arc.pdf>]

“Arc harvests metadata from several OAI compliant archives, normalizes them, and stores them in a search service based on a relational database (MySQL or Oracle).” Because of the rapidity with which software evolves, if a more effective software appears, FCLA retains the right to change open source software packages.

The harvester system can be run on the same server as the Aquatics e-print repository, thus costs will be for staff for development and ongoing maintenance only.

### **Phase 3: Implementation of zebra Z39.50 server to interface with IAMSLIC Z39.50 Distributed Library**

(Estimated implementation time: 2 months following establishment of either the Aquatics EPrint repository or Aquatic Commons database)

FCLA has a zebra Z39.50 server in use for another application. It can be configured to work with the Aquatic EPrint repository or with the more comprehensive Aquatic Commons database once it is established.

By implementing this server interface, any search query launched in the IAMSLIC Z39.50 Distributed Library will also be searching the metadata of the repositories. Effectively, this means a user will be searching across the Aquatic Commons, the distributed catalogues of more than 40 libraries plus their branches, and the serials listings from 40 institutions

Pauline Simpson suggests “That the Aquatic Commons is one more building block toward a superb one stop shop with document delivery at the hub. We are building a thematic portal with built in cross searching.”

Only funding required is in staff costs of FCLA and Steve Watkins to configure the interface.

## BUDGET

The majority of costs associated with Aquatic Commons are related to the hardware and programming to setup and maintain the system. Additional costs are associated with administrative functions related to the submitting of records to the Aquatics EPrint repository, development of training materials, and the promotion and education of potential users of the repository. Because IAMSLIC is a wholly volunteer organization, it would be most appropriate if it could meet its fiscal responsibilities through cost share of members' time. The major activities requiring members' time are delineated below; however, no time has been included for determining and testing configuration options for Aquatic EPrint repository and a minimal amount of time has been included for the development of training materials for Phase 2 and none for Phase 3. Ongoing promotion and education has only been costed in Phase 1, but obviously will be continuous.

| <b>Phase 1: Aquatic EPrint Repository</b>  | ASFA/FAO Funding | IAMSLIC (cost share) |
|--|------------------|----------------------|
| <b>Personnel</b>   |                  |                      |
| FCLA Programmers 120 hrs @ \$15  | \$1,800          |                      |
| Purchase, install and configure server computer 15 hours @ \$20  | \$ 300           |                      |
| Develop training materials in English, French, and Spanish by IAMSLIC members 20 hrs/ language @ \$12                            |                  | \$ 720               |
| Administration of submittals and coordinating setup between ASFA/FAO, IAMSLIC and FCLA 10% of Stephanie Haas' time for one year  |                  | \$6,700              |
| Consulting on e-repository software and setup with Pauline Simpson, Southampton 10 hrs   |                  | \$1,000              |
| Promotion of and education about the repository by IAMSLIC Resource Sharing Committee (sessions at 5 regional conferences, etc.) |                  | \$1,000              |
| <b>Hardware/Software</b>   |                  |                      |
| Server, dual cpu, 4GB memory, 156GB internal disk (projected to serve about 75,000 pdf files)                                    | \$5,000          |                      |
| Tape cartridge for backup  | \$ 200           |                      |



|                        |                |                |
|------------------------|----------------|----------------|
|                        |                |                |
| Red Hat Linux (OS)     | \$ 50          |                |
| Tivoli (backup server) | \$ 50          |                |
| Tripwire (security)    | \$ 300         |                |
| <b>Total</b>           | <b>\$7,700</b> | <b>\$9,420</b> |

|   |                  |                      |
|---|------------------|----------------------|
| <b>Phase 2: Development of the Aquatic Commons harvester, an OAI provider, a search interface, and database</b> | ASFA/FAO Funding | IAMSLIC (cost share) |
| <b>Personnel</b>  |                  |                      |
| FCLA programmers<br>200 hrs @ \$15  | \$3,000          |                      |
| Develop training materials in English, French, and Spanish by IAMSLIC members<br>20 hrs/ language @ \$12        |                  | \$720                |
| <b>Total</b>  | <b>\$3,000</b>   | <b>\$720</b>         |

|  |              |                      |
|--|--------------|----------------------|
| <b>Phase 3: Implementation of zebra Z39.50 server to interface with IAMSLIC Z39.50 Distributed Library</b> |              | IAMSLIC (cost share) |
| <b>Personnel</b>   |              |                      |
| FCLA programming<br>40 hrs @\$15   | \$ 600       |                      |
| Steve Watkins' time to implement interface estimated at 10 hrs   |              | \$1,000              |
| <b>Total</b>   | <b>\$600</b> | <b>\$1,000</b>       |

The costs listed above represent the start-up costs for the various system functions.

Annual ongoing costs for FCLA system maintenance for the *Aquatic Commons* are:

|   |                 |
|---|-----------------|
| Hardware / network                          |                 |
| Server maintenance                          | \$ 500          |
| Hardware replacement cost                   | \$ 1,000        |
| Network cost                                | \$ 86           |
| Software                                    |                 |
| Red Hat Linux (OS)                          | \$ 50           |
| Tivoli (backup server)                      | \$ 50           |
| Tripwire (security)                         | \$ 165          |
| Staff                                       |                 |
| Ongoing maintenance and support (20 hrs/mo) | \$ 3,600        |
| <b>Total annual ongoing costs</b>           | <b>\$ 5,451</b> |

This includes maintenance for the Aquatics e-print repository, a harvester, an OAI provider, interfaces to existing OAI-compliant servers, a search and retrieval interface, a database, and a zebra Z39.50 server to interface with the IAMSLIC Z39.50 Distributed Library. For the Aquatics EPrint repository alone, hardware and software costs would be the same, but annual staff support would be 5 hrs/mo (\$900). The hardware replacement cost is intended to provide a sufficient amount over a five year period to replace the server.

Ongoing matching costs from the IAMSLIC membership would be continued efforts to inform all levels of the aquatic community of the availability of the Aquatic e-print repository, to provide guidance in the submittal of items to the repository, and where appropriate to submit digital items from their own organizations and institutions.

### **Awareness, Education, and Training**

As indicated in the budget outline above, the IAMSLIC Resource Sharing Committee will assume the main responsibility for preparing training materials in French, Spanish, and English related to the Aquatics e-print repository and/or the Aquatics Commons database. Additionally, it is expected that all collaborating partners involved in the Aquatic Commons project will inform their colleagues of these initiatives. It is hoped that through colleague networking interest in contributing to the Aquatic Commons will be sparked much as it has been with the IAMSLIC Z39.50 Distributed Library project. Because Phase 3 of this project will also permit direct searching of the Aquatic Commons database through the developed IAMSLIC Z39.50 Distributed Library, those participating in this established program should find this enhancement of particular value in providing more access to relevant full text.

Jean Collins of FAO, who is very supportive of the Aquatic Commons initiative, wrote:

"I think an IAMSLIC Aquatic Commons is a good idea...I realise that there are constraints and many legal-type issues that would have to be ironed out...It would be the best opportunity for me to help get developing countries involved and their publications included.

It would be good to start with a critical mass of publications - such as those of NACA and MRC. As regional organizations they also have more resources than a lot of national fisheries institutions - they would provide a better testing ground than those who only produce a half dozen documents a

year.

My role (although I don't always determine that myself!) would be providing contacts and promoting the Aquatic Commons, providing guidelines/training for developing country libraries, getting funding."

As indicated in the Justification of Need section above, NACA has already indicated its interest in participating. Jean's willingness to promote and provide training assistance at the local level in developing countries is a key factor for developing a truly valuable repository.

Preliminary discussions with CSA have already begun about harvesting the records from either the Aquatics e-print repository or from the Aquatic Commons database for inclusion in ASFA. Such inclusion will stimulate increased use of this material and is expected to greatly enhance the value of the ASFA database to all researchers.

## **ADMINISTRATION**

In order to provide oversight and to further the goals of the Aquatic Commons initiative, one member from each organization should be appointed to serve on an Aquatic Commons Steering Committee. The exact charges, duties, and communication avenues will be decided later.

In terms of administration, all aspects of the technical functioning of the system will be the responsibility of FCLA staff. The administrative review of submittals to the Aquatics e-print repository will be the responsibility of IAMSLIC members who will agree to serve gratis for a period of at least one year. Outgoing submittal review administrators will train the incoming administrator and provide consultation as needed.

The Aquatic Commons database resulting from the harvesting of extant repositories and the provision of OAI metadata are all automated functions that will be administered by FCLA. It will be the responsibility of all partners to alert the Aquatic Commons Steering Committee to new repositories coming online.

## **SUSTAINABILITY AND EVALUATION**

In terms of financial commitments, the costs associated with the development and maintenance of Aquatic Commons are minimal. IAMSLIC is an all volunteer organization, but its in-kind contributions in terms of the development of the project and the administration of the Aquatic e-print repository are substantial. Initially, Stephanie Haas has agreed to act as the facilitator to coordinate the development of the three phases and to act as the submittal review administrator for the repository. Pauline Simpson has agreed to act as a consultant to Stephanie and FCLA in the initial setup and implementation of the repository system. Steve Watkins will be providing match too in setting up the interface with the zebra Z39.50 server.

The Florida Center for Library Automation projects that \$4,500 will be needed on a yearly basis to maintain the Aquatic Commons components. Current proposed changes in IAMSLIC membership dues are likely to affect IAMSLIC's ability to contribute to this maintenance fee. However, it is also likely that the continued e-print repository administrative duties that are the responsibility of IAMSLIC will come close to matching that amount.

It seems logically that the cost-benefit evaluation of this project will have to be measured qualitatively by those contributing to the e-print repository, and quantitatively by measuring use of the e-print repository and/or the Aquatic Commons database. Harvesting from these two sources will also be a measure of value. A critical mass of documents must be deposited before searching the e-print repository attains efficacy; the same holds true for the creation of the Aquatic Commons database from harvested metadata. During the start up phase of any of the component, any evaluative measures will be inaccurate, although statistics will be an important functionality of the software. If possible, a commitment to fund the ongoing costs for a two year period to commence after the set up year would be highly desirable. At that point, a more accurate assessment of potential benefit can be determined.

As with all initiatives, the sustainability of the Aquatic Commons project is directly tied to its perceived value by the stakeholders. There has been an ongoing need for full text access to the ephemeral, or grey literature, indexed by ASFA for years. Because more and more scientific reports are born digital, the capture and preservation of these files is becoming increasingly critical. IOC has begun to capture a very small proportion of the digital literature related to marine science. It has plans to create similar repositories in other circumscribed regions of the world. The role of Aquatic Commons is to compliment these efforts by addressing the repository needs of the non-circumscribed marine and aquatic regions. The documents related to freshwater and estuarine systems are more critical in many parts of the developing world than are the marine documents. For many years, the activities of IOC, FAO, FAO ASFA, and IAMSLIC have inextricably tied the organizations together. The joint support of an initiative such as Aquatic Commons will not only formally strengthen relationships but will draw on the expertise of each organization in building digital resources of global value.

|   | A  | B   | C  |
|---|--|---|--|
| 1 | <b>Appendix 4: Technical criteria comparison: Dspace and EPrints</b>   |   |  |
| 2 | Criteria   | Dspace (all responses *Dspace* were provided by Brian Voss, NOAA, Seattle Regional Office)  | Eprints (Stephanie supplied information from the demo site <a href="http://demoprints.eprints.org/">http://demoprints.eprints.org/</a> and the Technical Documentation ( <a href="http://www.eprints.org/documentation/tech/php/intro.php">http://www.eprints.org/documentation/tech/php/intro.php</a> ) site. Pauline Simpson's additions are in red.)  |
| 3 | What types of files can the repository handle? From NOAA criteria  | *Dspace* All content types accepted. Bitstream (read: file) Format registry is editable to allow new formats. No mention of conversion utility within DSpace. <a href="http://dspace.org/technology/system-docs/configure.html#registries">http://dspace.org/technology/system-docs/configure.html#registries</a> | An EPrint Repository is a collection of digital documents. Each eprint, be it article, book, dissertation or journal, consists of metadata and files, which may be accessed via the EPrints web site. One or more files in a directory. Most documents are a single file, but some contain multiple files: such as, an HTML document with diagrams or a collection of images of the same item from different angles. Files: HTML, PDF, Postscript, ASCII, Other, Image of Cover<br>Guides for Repository Administrators<br><br>* Howto: Create a Data Repository<br>* Howto: Create a Multi-Language Repository<br>* Howto: Create a Multimedia Repository<br>* Howto: Create RSS Feeds<br>* Howto: Create a Theses Repository See examples of various types of repositories <a href="http://www.eprints.org/software/examples/">http://www.eprints.org/software/examples/</a> -SH |
| 4 | Version control --does system retain and link all versions of a document? From NOAA criteria   | *Dspace* Versioning NOT supported (JHU evaluation), but perhaps managed in other ways ( <a href="http://wiki.dspace.org/PersistentIdentifiers">http://wiki.dspace.org/PersistentIdentifiers</a> )   | Create a new version of a current eprint (copies metadata, documents and links it to the old version): for more information, see Creating new versions of eprints .<br><a href="http://www.eprints.org/documentation/demoprints/wwhelp/wwhimpl/common/html/wwhelp.htm?context=EPrints_User_Guide_Help&amp;file=04%20Registered%20Users%20demo%20site%20features7.html">http://www.eprints.org/documentation/demoprints/wwhelp/wwhimpl/common/html/wwhelp.htm?context=EPrints_User_Guide_Help&amp;file=04%20Registered%20Users%20demo%20site%20features7.html</a> . (In EPv3) every change to a record increases its revision number. Also new records can be added and indicated to be later versions of existing records - PS   |
| 5 | Access control --can system restrict access to repository? to document types or collection? to individual documents?/ From NOAA criteria | *Dspace* See answer to 5. authentication  | Access options for documents is to Anyone, Registered Users Only, Repository Staff only [SH] Decided at the deposit stage  |

|   | A  | B   | C   |
|---|--|---|---|
| 6 | Interface --how easy to use in general, how robust the searching, does system use stylesheets, what design "metaphors" does system use, etc.? From NOAA criteria | <p>*DSpace* Workflow for Submitting materials:<br/> <a href="http://dspace.org/implement/submit-content.html">http://dspace.org/implement/submit-content.html</a></p> <p>Customizable interfaces using CSS, HTML, JSP <a href="http://dspace.org/technology/system-docs/configure.html#customui">http://dspace.org/technology/system-docs/configure.html#customui</a></p> <p>Custom Metadata-entry Pages for Submission<br/> <a href="http://www.dspace.org/technology/system-docs/submission.html">http://www.dspace.org/technology/system-docs/submission.html</a></p> <p>...simple API which allows for indexing new content, regenerating the index, and performing searches on the entire corpus, a community, or collection.</p> <p>Lucene gives us fielded searching, stop word removal, stemming, and the ability to incrementally add new indexed content without regenerating the entire index. As of DSpace 1.2.1 the Lucene search indexes are configurable,<br/> <a href="http://www.dspace.org/technology/system-docs/functional.html#search_browse">http://www.dspace.org/technology/system-docs/functional.html#search_browse</a></p> <p>Option to view simple and full record formats<br/> <a href="http://www.dspace.cam.ac.uk/handle/1810/1425?mode=simple">http://www.dspace.cam.ac.uk/handle/1810/1425?mode=simple</a></p> | <p>Customizable interfaces. all the text in the interface is stored in an XML file,</p>   |
| 7 | Authentication --what method does system use to authenticate users (e.g. LDAP)? are there different levels of approval? can repository staff enhance metadata?/  | <p>*DSpace* Users (e-people) can be given group attributes and communities through files can be associated with same attributes.<br/> <a href="http://www.dspace.org/technology/system-docs/functional.html#auth">http://www.dspace.org/technology/system-docs/functional.html#auth</a></p> <p>LDAP authentication is available<br/> <a href="http://www.dspace.org/technology/system-docs/configure.html#ldap">http://www.dspace.org/technology/system-docs/configure.html#ldap</a></p>  | <p>Can be linked to LDAP or any business systems in house. Manual Registration by approval system which allocates levels or activity. Repository staff can enhance metadata. * Preservation features. eg preservation metadata fields</p> |

|    | A   | B  | C  |
|----|---|--|--|
| 8  | Batch loading/export --what formats can you load? what formats can you get out?/  | <p>*DSpace* Batch import command line. * Does an error in the middle of the ingest mean all work done so far is discarded? Can resume after failure. * How does bulk ingest interact with normal operation? RDBMS takes care of it. Issues with lucene?</p> <p><a href="http://wiki.dspace.org/CrosswalkPlugins">http://wiki.dspace.org/CrosswalkPlugins</a><br/>The Crosswalk Plugin interface described here only addresses XML-based metadata formats. Since OAI-PMI can only export XML, and metadata containers like METS and IMS-CP have a preference for XML metadata, this is not seen as an important limitation at this time. If there is a need, anyone can add a new plugin interface to handle binary or text-based metadata (e.g. old-style MARC). 'Registration' allows batch import type function from existing IRs</p> <p><a href="http://www.dspace.org/technology/system-docs/functional.html#registration">http://www.dspace.org/technology/system-docs/functional.html#registration</a></p> | Not that I can find - SH <b>Yes use XML import - map from original format. Endnote and BibTex exports plus straight listings</b> |
| 9  | Usage statistics --what can you get from system? how does it work (easy or difficult)?/ From NOAA criteria  | <p>*DSpace* report about the contents and use of your system can be automatically generated by the system. generated by analyzing DSpace's log files. can be broken down monthly. either be made public or restricted to administrator access only.</p> <p><a href="http://www.dspace.org/technology/system-docs/functional.html#statistics">http://www.dspace.org/technology/system-docs/functional.html#statistics</a></p>   | <b>standard package awstats works with Eprints - provides all the usual requirements: who what when where how many</b>           |
| 10 | Extensibility/interoperability -- can you extend the metadata format? how does various admin/metadata metadata? can you selectively turn off From NOAA criteria | <p>*DSpace* <code>add/index/display</code> another Dublin Core field?</p> <p><a href="http://wiki.dspace.org/TechnicalFaq#head-e605ed31d8b09520e881c679bbde730a7474a9d8">http://wiki.dspace.org/TechnicalFaq#head-e605ed31d8b09520e881c679bbde730a7474a9d8</a></p>   | Can modify fields in metadata -SH  |

|    | A   | B  | C  |
|----|---|--|--|
| 11 | Storage of digital property rights/copyright --can your software do this?/ From NOAA criteria   | *DSpace*<br>Submitters are given an opportunity to select a Creative Commons license to accompany the Item.<br><a href="http://creativecommons.org/">http://creativecommons.org/</a><br>Setting webui.licence_bundle = true in dspace.cfg will result in a hyperlink being rendered on the Item View page that points to the item's licence.<br><a href="http://dspace.org/technology/system-docs/configure.html">http://dspace.org/technology/system-docs/configure.html</a>  | Not integrated into program that I can find SH - <b>Digital Rights Management is coming soon- PS</b>   |
| 12 | Authority control --does the system allow for authorized lists of names, subject terms, etc. Can user pick from a list on entry?/               | *DSpace* No authority control like a library catalog, but pulldown menus may be an option in submission process.<br><a href="http://www.dspace.org/technology/system-docs/business.html">http://www.dspace.org/technology/system-docs/business.html</a>  | Subject terms (in out of the box version) come from LC. Also uncontrolled keywords -SH <b>Any thesaurus can be loaded - PS</b>   |
| 13 | Withdrawals --can system allow for blocking access to superseded/withdrawn objects that we wish to retain in database?/ From NOAA criteria      | *DSpace* Items are marked withdrawn. Items deleted?<br>Collections and communities can be deleted. Note that there is no 'DELETE' action. In order to 'delete' an object (e.g. an item) from the archive, one must have REMOVE permission on all objects (in this case, collection) that contain it. The 'orphaned' item is automatically deleted.<br><a href="http://www.dspace.org/technology/system-docs/functional.html#auth">http://www.dspace.org/technology/system-docs/functional.html#auth</a><br><a href="http://www.dspace.org/technology/system-docs/policy-issues.html#withdraw">http://www.dspace.org/technology/system-docs/policy-issues.html#withdraw</a> | If an eprint becomes obsolete or if an eprint was deposited in error, for example, you can send a request to a Repository Administrator and request that the eprint is removed from the Repository.<br>To request an eprint is removed from the Repository<br><br><ol style="list-style-type: none"> <li>1. From your User Area, click Review your documents in the archive.</li> <li>2. Locate the eprint you want to have removed from the Repository and click Request Removal.</li> <li>3. In the box, enter the reason why you want the eprint removed from the Repository.</li> <li>4. Click Send Request to send your reason to the Repository Administrator. Alternatively, if you want to keep the eprint in the Repository, click Cancel. -SH</li> </ol> |
| 14 | Concurrency...<br>* Can objects/collections be locked for updates?<br>* Are write/write and read/write conflicts handled?<br>From NOAA criteria | *DSpace* Yes. Handled by database.   | <b>Yes</b>   |



|    | A   | B  | C  |
|----|---|--|--|
| 15 | Backup and restore....<br>Are there special facilities for backup and restore? From NOAA criteria         | *DSpace* No  | Not that I can find - SH <b>usual backup of data does allow restore</b>  |
| 16 | Notifications....<br>* Can users register to be notified of object/collection changes? From NOAA criteria | DSpace* Users can subscribe to collections and receive email notification of changes.  | <b>Yes</b> Select Change your subscription options. This allows you to instruct the repository to automatically email you with lists of documents deposited that match your criteria every day, week or month. -SH   |
| 17 | Updating software after modifications.... From NOAA criteria  | *DSpace*...always be some manual merging involved, but CVS can help out a lot. This page describes how to get CVS to help you, and once you've followed this once, your local code base will be a checkout of the SourceForge CVS tree, which means you can keep up-to-date more easily in the future and over time by doing periodic cvs updates.<br><a href="http://wiki.dspace.org/UpdatingLocalCodebase">http://wiki.dspace.org/UpdatingLocalCodebase</a>  | We currently expect to release a few more versions of EPrints 2. We plan to make it as easy as possible to upgrade from earlier versions of EPrints 2. see <a href="http://www.eprints.org/documentation/tech/php/intro.php">http://www.eprints.org/documentation/tech/php/intro.php</a> Upgrading documentation at <a href="http://www.eprints.org/documentation/tech/php/updating.php">http://www.eprints.org/documentation/tech/php/updating.php</a> -SH<br><b>EPv3 is on beta test- PS</b> |
| 18 | Tracking Provenance of items<br>From NOAA criteria  | *DSpace*...While provenance information in the form of prose is very useful, it is not easily programmatically manipulated. The History system captures a time-based record of significant changes in DSpace, in a manner suitable for later 'refactoring' or repurposing. Currently, the History subsystem is explicitly invoked when significant events occur (e.g., DSpace accepts an item into the archive). The History subsystem then creates RDF data describing the current state of the object. The RDF data is modeled using Harmony/ABC, an ontology for describing temporal-based data, and stored in the file system. Some simple indices for unwinding the data are available. | Appears to allow linkage between versions -SH <b>see versioning control. Can track all changes to metadata - PS</b>  |

|    | A   | B  | C  |
|----|---|--|--|
| 19 | Directory Structure... From NOAA criteria   | *DSpace* <a href="http://dspace.org/technology/system-docs/directories.html#sourcedir">http://dspace.org/technology/system-docs/directories.html#sourcedir</a>   | Tech. Documentation - The Archive Configuration Files<br><a href="http://www.eprints.org/documentation/tech/php/configarchive.php">http://www.eprints.org/documentation/tech/php/configarchive.php</a><br>Eprints 2 relational diagram<br><a href="http://www.eprints.org/documentation/tech/php/structure.php#eprints_2_relational_diagram">http://www.eprints.org/documentation/tech/php/structure.php#eprints_2_relational_diagram</a> - SH   |
| 20 | Installation Open Archives Forum Inventories – Open Archives Software Tools<br><a href="http://www.oaforum.org/otherfiles/tv-tools.pdf">http://www.oaforum.org/otherfiles/tv-tools.pdf</a>            | The installation of DSpace requires a little more effort. But in fact DSpace is easy to run and maintain for any experienced systems engineer.<br>In order to run DSpace the following list of Software is necessary to be installed and configured before: Java 1.3, Tomcat 4.0+, Apache 1.3, PostgreSQL 7.3+, Ant 1.5. Details of the requirements can be viewed at: <a href="http://dspace.org/technology/system-docs/install.html#prerequisite">http://dspace.org/technology/system-docs/install.html#prerequisite</a><br>If the programmer follows step by step the installation documentation, Java, Ant and PostgreSQL are easy to install successfully.<br>To set up DSpace man needs to compile the DSpace source code with java tool Ant. The Tomcat server must be started by user “dspace” and user “dspace” should then create a database named “dspace”.<br>With the installation some common problems arose, e.g. that Tomcat doesn't work when the DSpace is connected to Tomcat. Some changes in the configuration script solved that problem.<br>There is no support service for the DSpace installation. But there is a detailed system documentation at: | Eprints is easy to set up: An installation script automates most of the installation processes.<br>It is possible to chose between a source- or binary-installation. With the source one the software has to be compiled by the programmer. The binary one is precompiled for special architectures like Solaris Sparc systems. The programmer only need to configure the software. MySQL, Apache and mod_Perl, the components which are necessary for implementation are smooth installations - no matter if source- or binary-installation is chosen. The installation of additional required Perl modules need more time to resolve the dependencies.<br>There are two possibilities to support the system: One installation variant is a Solaris environment. The second variant, Linux, is easier to maintain.<br>If any installation problems are arising a comprehensive support is ensured. GNU Eprints has a separate website containing documentation, downloads, demonstration server and mailing lists: <a href="http://software.eprints.org/">http://software.eprints.org/</a> - SH |
| 21 | Programming language -- Open Archives Forum Inventories – Open Archives Software Tools<br><a href="http://www.oaforum.org/otherfiles/tv-tools.pdf">http://www.oaforum.org/otherfiles/tv-tools.pdf</a> | DSpace had been tested on Linux Suse 7.3. In general DSpace can run on Solaris, Linux and Windows systems.   | Both environment variants had been tested: Solaris and Linux. Furthermore it is also possible to install Eprints2 on any computer that is running with GNU/Linux or UNIX operating system.   |

|    | A  | B   | C  |
|----|--|---|--|
| 22 | <p>Functions-- Open Archives Forum Inventories – Open Archives Software Tools<br/> <a href="http://www.oaforum.org/otherfiles/tv-tools.pdf">http://www.oaforum.org/otherfiles/tv-tools.pdf</a></p> | <p>DSpace can be used for self archiving by institutions and faculties. It provides long-term physical storage and management of digital items in a repository.</p> <p>DSpace is organised into "Communities" and "Collections", each of which retains its identity within the repository. It supports a variety of digital formats and content types including text, images, audio, and video and allows contributors to limit access to items in DSpace. All these items can be organised by an administration interface.</p> <p>DSpace supports the OAI protocol 2.0 as a data provider. This OAI support was implemented using OCLC's OAICat open-source software to make DSpace item records available for harvesting.</p> <p>Currently DSpace supports only the Dublin Core metadata element set with a few qualifications conforming to the library application profile. But there are still developing plans to support a subset of the IMS/SCORM element set (for describing education material) in the coming year.</p> <p>More details of DSpace functionality can be founded at <a href="http://libraries.mit.edu/dspace-">http://libraries.mit.edu/dspace-</a></p> | <p>Eprints is free software which creates online archives:</p> <p>It is possible to store documents in any common format that the archive administrator defined to be accepted. Each individual research paper/ eprint/ ... can be stored in more than one document format.</p> <p>The archive can use any metadata schema; the administrator decides what metadata fields are held about each eprint. This is specified in three or four stages:</p> <ul style="list-style-type: none"> <li>• Definition of a maximal set of metadata fields that should be stored (e.g. authors, title, journal, journal volume, etc.)</li> <li>• Definition of different types of eprints (e.g. refereed journal article, thesis, technical report, unpublished preprint, etc.)</li> <li>• Specification for each type which metadata fields should be stored, and which of those fields are mandatory.</li> <li>• Decide how these metadata fields should be projected into the Open Archives world. (If necessary, interoperability can be switched off, but this is strongly discouraged.)</li> </ul> <p>More functions can be viewed at <a href="http://software.eprints.org/">http://software.eprints.org/</a> -SH</p> |
| 23 | <p>Reusage --Open Archives Forum Inventories – Open Archives Software Tools<br/> <a href="http://www.oaforum.org/otherfiles/tv-tools.pdf">http://www.oaforum.org/otherfiles/tv-tools.pdf</a></p>   | <p>It is not reported how many archives are running DSpace software. One example of an European repository that implemented DSpace is "Erasmus University: Research Online".</p>  | <p>Eprints is widespread all over the world. In Jun 2006 there are 210 worldwide archives running Eprints software officially listed (<a href="http://software.eprints.org/">http://software.eprints.org/</a>).</p>  |

|    | A  | B   | C   |
|----|--|---|---|
| 24 | <p>Technology -- Open Archives Forum Inventories – Open Archives Software Tools<br/> <a href="http://www.oaforum.org/otherfiles/tv-tools.pdf">http://www.oaforum.org/otherfiles/tv-tools.pdf</a></p>       | <p>DSpace operates with new technologies such as the Postgres database, that is more advanced than mySQL and Tomcat for jsp/java web application, that has higher performance than eprints.<br/> Dspace supports and includes also handle server, which ensures that each document has unique and persistent URL.<br/> Optionally, DSpace can be protected by the security features (SSL) of Tomcat. It is also possible to use the redirect function (port number can be omitted) from Apache referring to Tomcat.</p>   | <p>Eprints uses traditional technologies and runs on pure Open Source systems: mySQL is the world's most popular open source database, recognized for its speed and reliability and Apache has been the most popular web server on the Internet since April of 1996.<br/> Eprints is programmed by using the script language "Perl", that is low level but powerful.</p>  |
| 25 | <p>Interoperability -- Open Archives Forum Inventories – Open Archives Software Tools<br/> <a href="http://www.oaforum.org/otherfiles/tv-tools.pdf">http://www.oaforum.org/otherfiles/tv-tools.pdf</a></p> | <p>The DSpace system is freely available as open-source software. This allows to make any necessary changes to the downloaded copy.<br/> The system was designed to make adaptations for individual organisations as easy as possible. In fact, several modules in DSpace will probably be customised by organizations using this tool (e.g. it might be necessary to get authorization and authentication for more than one person). Or some organisations may want to adapt a different environment than recommended (e.g. replace postgresQL by mySQL or Oracle). At the moment, substituting a different relational database than postgresQL will require just a few changes to the system's Browse module.<br/> Java provides documented Java APIs that can be enhanced to allow interoperation with other systems that an institution might be running (e.g. auto-depositing in DSpace a department's web document system, or the campus data warehouse).</p> | <p>Eprints is freely distributable and subject to the GNU General Public License. This means that its source code is open and freely modifiable by any programmer who wishes to modify it (on condition that modifications are all free and open).<br/> Therefore in principle an adjustment to every environment is possible even if it is different than the recommended. Naturally this may be connected with substantial expenditure.<br/> However Eprints offers no supporting documents there are nevertheless mailing lists for support. -SH</p> |

|    | A  | B   | C   |
|----|--|---|---|
| 26 | <p>Search -- Open Archives Forum Inventories – Open Archives Software Tools<br/> <a href="http://www.oaforum.org/otherfiles/tv-tools.pdf">http://www.oaforum.org/otherfiles/tv-tools.pdf</a></p> | <p>DSpace offers two levels of text search: simple and advanced search. It's submission process also allows to use a qualified version of the Dublin Core metadata schema for the description of each item. These descriptions are stored in a relational database, which is used by the search engine to retrieve items.</p> | <p>Eprints allows to scan each of the metadata field types in the database by simple or advanced search. Any metadata field can be searched with fine granularity by SQL querying the database. Further information could be found at <a href="http://software.eprints.org/files/eprints1/docs/system.html#SECTION00060000000000000000">http://software.eprints.org/files/eprints1/docs/system.html#SECTION00060000000000000000</a></p> |

**Appendix 5: Overall Criteria Comparisons between Proposals**

| <b>CRITERIA</b>                          | <b>FCLA</b>  | <b>WHOI_Community</b>   | <b>WHOI_Server</b>  | <b>IOC</b>  |
|--|--|---|---|---|
| <b>Host</b>                              |  |   |   |   |
| <i>member of Iamslic</i>                 | No--affiliated through UF  | Yes   | Yes   | Yes   |
| <i>funding route for inhouse support</i> | Funding is external only; contracted through IAMSLIC   | External  | External  | External  |
| <i>length of agreed support</i>          | indefinitely with annual ongoing system maintenance costs of \$2,351   | indefinite with annual ongoing maintenance costs of \$500   | indefinite with annual ongoing backup services of \$500   | indefinite with recurring annual maintenance of \$1,000   |
| <i>number of support posts</i>           | FCLA has 23 programmers on staff who work with projects ranging from digital collections to the OPAC for the Florida state university system | 3: system administrator, web, server admin available to support Dspace; none are full-time for this project | 3: system administrator, web, server admin available to support Dspace; none are full-time for this project | 3: 2 persons of Hasselt University Library for DSpace support + development supported by Informatics Faculty; 1 IOC IT-specialist as system administrator |
| <i>IAMSLIC branded service</i>           | yes  | no  | yes   | yes   |
| <b>Server</b>                            |  |   |   |   |
| <i>dedicated server</i>                  | yes  | no  | yes   | hosted as a separate DSpace installation on the IODE DSpace server (Raid: 2x 76Gb)  |
| <i>size</i>                              | up to 75,000 pdf files   | up to 50 gigabytes (approx 16,000 pdf files); \$100/additional gigabyte                                     | per IAMSLIC specifications  | 25Gb immediately available (extension if necessary)   |
| <i>service availability/downtime</i>     | none anticipated   | none anticipated  | none anticipated  | none anticipated  |
| <i>backup</i>                            | tape cartridge/ Tivoi (backup server)  | tape  | tape  | tape  |
| <i>connectivity/access</i>               | Internet   | Internet  | Internet  | Internet  |

|   |   |                 |                           |  |
|---|---|-----------------|---------------------------|--|
| <b>Software</b>   | E-print   | Dspace          | Dspace                    | Dspace 1.4<br>(+add-ons:<br>e.g. ASFA-<br>thesaurus in<br>development)   |
| <i>open source</i>  | yes   | yes             | yes                       | yes  |
| <i>OAI compliant</i>  | yes   | yes             | yes                       | yes -<br><i>metadata:Dubl<br/>in Core-<br/>Dublin Core<br/>Qualified -<br/>Agris AP (end<br/>2006) - Mods<br/>(2007) +<br/>SRW/SRU</i> |
| <i>upgrades timescale</i>                                       | as available  | as available    | as available              | as available   |
| <i>Number of technical<br/>support in house<br/>with skills</i> | 3   | 2               | 2                         | IOC: 1 -<br>UHasselt: 2  |
| <i>multilingual<br/>capabilities</i>                            | yes see<br><a href="http://www.eprints.org/software/howto/multi-language/">http://www.eprints.org/software/howto/multi-<br/>language/</a> | no              | yes                       | yes - main<br>languages<br>already<br>available)   |
|   |   |                 |                           |  |
| <b>Content</b>  |   |                 |                           |  |
| <i>admin</i>  | IAMSLIC responsibility  | MBLWHOI Library | IAMSLIC<br>responsibility | IAMSLIC<br>responsibility  |
| <i>preservation</i>   | tape cartridge  | tape            | tape                      | tape   |
|   |   |                 |                           |  |
| <b>Services</b>   |   |                 |                           |  |
| <i>help desk</i>  | technical issues only   | yes             | technical<br>issues only  | set up,<br>technical and<br>organizational<br>support  |
| <i>format conversion<br/>service</i>                            | no  | yes             | yes                       |  |
| <i>training courses</i>   | no  | possible        | no                        | available in<br>IOC  |
| <i>training<br/>documentation</i>                               | no  | possible        | no                        | Based on the<br>results of<br>OdinPubAfrica<br>a multilingual<br>package will<br>be developed  |
| <i>User Group</i>   | yes   | yes             | yes                       | yes  |
|   |   |                 |                           |  |

|                                      |  |                                     |  |   |
|--------------------------------------|--|-------------------------------------|--|---|
| <b>Cost</b>                          |  |                                     |  |   |
| <i>implementation and first year</i> | \$5,700                                      | \$1,500                             | \$14,625   | \$4,000   |
| <i>2nd year and ongoing</i>          | \$2,351 includes computer repl after 5 years | \$500                               | \$500 back-up services; other services at \$75.00/hr | \$1,000   |
| <i>host contribution</i>             | general technical support                    | general tech support; Dspace admiin | general technical support                            | Expertise in setting up similar repositories for Odinafrica and OdinCarsa |



**Appendix 6: COMPARISON OF COSTS FOR START UP AND MAINTENANCE of the AQUATIC COMMONS REPOSITORY**  
**Aquatic Commons Repository at Woods Hole based on shared server. (Community option)**

|  |                     |
|--|---------------------|
| Purchase of gigabyte storage per gigabyte  | \$100.00            |
| Annual maintenance fee   | \$500.00            |
| IAMSLIC Annual Conference/training session Project Manager (registration & travel) | \$1000.00           |
| Unit cost (per page scanned)   | \$0.40              |
| Unit cost (per 650mb archival CD-R)  | \$10.00             |
| Postage charges for delivering CD-R  | Current market rate |

**Aquatic Commons Repository at Woods Hole based on separate Dspace server**

|  |                     |
|--|---------------------|
| Purchase DSpace server, including Handle server    | \$5,000.00          |
| Install and configure DSpace, 15 hours @ \$75.00   | \$1,125.00          |
| MBL IT programmers, 100 hours @ \$75.00            | \$7,500.00          |
| Consulting w/ WHOAS project manager                | \$1,000.00          |
| Back-up services for database                      | \$500.00 per annum  |
| Scanning services:<br>Unit cost (per page scanned) | \$0.40              |
| Unit cost (per 650mb archival CD-R)                | \$10.00             |
| Postage charges for delivering CD-R                | Current market rate |

**Aquatic Commons Repository at FCLA based on Eprints**

| <b>Phase 1: Aquatic EPrint Repository</b>   | <b>Funding Required</b> |
|---|-------------------------|
| <b>Personnel</b>  |                         |
| FCLA Programmers 120 hrs @ \$15   | \$1,800                 |
| Purchase, install and configure server computer 15 hours @ \$20                               | \$ 300                  |
| <b>Hardware/Software</b>  |                         |
| Server, dual cpu, 4GB memory, 156GB internal disk (projected to serve about 75,000 pdf files) | \$3,100                 |
| Tape cartridge for backup   | \$ 200                  |
| Red Hat Linux (OS)  | \$ 50                   |
| Tivoli (backup server)  | \$ 50                   |
| Tripwire (security)   | \$ 300                  |
| <b>Total</b>  | <b>\$5,700</b>          |

Annual ongoing costs for FCLA system maintenance for the *Aquatic Common Repository* are:

|   |                 |
|---|-----------------|
| Hardware / network                          |                 |
| Server maintenance                          | \$ 500          |
| Hardware replacement cost                   | \$ 600          |
| Network cost                                | \$ 86           |
| Software                                    |                 |
| Red Hat Linux (OS)                          | \$ 50           |
| Tivoli (backup server)                      | \$ 50           |
| Tripwire (security)                         | \$ 165          |
| Staff                                       |                 |
| Ongoing maintenance and support (10 hrs/mo) | \$ 900          |
| <b>Total</b>                                | <b>\$ 2,351</b> |

The hardware replacement cost is intended to provide a sufficient amount over a five year period to replace the server. If no replacement funds are accrued, the yearly amount would be \$1,750.

#### **Aquatic Commons Repository at IOC based on DSpace**

|  |  |
|--|--|
| Specific Aq. Comm.DSpace installation, use of repository server at IOC-secretariat   | Offered by IOC/IODE  |
| Install and configure DSpace   | Offered by IOC/IODE  |
| Customization DSpace –programming  | \$ 3000  |
| Project management <ul style="list-style-type: none"> <li>- consultations with IAMSLIC/FAO/...</li> <li>- development standards</li> <li>- management + support during the project period</li> </ul> | Preferably to be done by IAMSLIC members and requires further discussion |
| Back-up services for database  | Offered by IOC/IODE  |
| Adaptation of self-training package to IAMSLIC needs   | \$ 1000  |
| Extra:<br>Service development: bibliographic list – journal interface – conference interface   | To be discussed as relevant  |
| <b>TOTAL DEPLOYMENT COST</b>   | <b>US\$ 4000</b>   |

**Recurrent (maintenance cost):** US\$ 1000/year (as from 13<sup>th</sup> month after deployment and acceptance of the product by IAMSLIC)