# Aquifer Gap Analysis Task Group Final Report

August 15, 2006

## Charge:

The Aquifer Implementation Group charged the Gap Analysis task group to identify obstacles to making collections available for DLF Aquifer OAI harvesting and to recommend solutions for removing those obstacles.

## Members:


- Naomi Nelson, Emory University, Collections Working Group (libnn@emory.edu)
- Melanie Feltner-Reichert, University of Tennessee, Knoxville, Metadata Working Group (feltner@email.lib.utk.edu)
- Tom Habing, University of Illinois, Urbana-Champaign, Technical Consultant (thabing@uiuc.edu)


## *Executive Summary*


The task group received lists of potential Aquifer participants from the Collections working group. Using "Barriers to Metadata Sharing via the OAI Protocol: A White Paper" by Sarah L. Shreeves, the group developed a survey designed to surface potential obstacles. The survey was divided into the following sections 1) Content, 2) Metadata, 3) Technical Infrastructure, 4) Rights, and 5) Institutional Information. After soliciting feedback on the survey and revising it, the task group invited the institutions that had been identified as potential participants in Aquifer to complete and return it through SurveyMonkey. Thirteen of the 19 institutions returned the survey (64%.) This report summarizes the results of the survey and suggests solutions for the obstacles it surfaced. A list of respondents and the summary report from SurveyMonkey are attached as appendices.


## *Section 1. Content*

The first section of the survey was designed to determine obstacles to aggregation related to the content of the collections. It sought information about the kinds of filters that might need to be applied during aggregation and assessed how difficult it might be to do the filtering.

### Discussion

As one respondent noted in a comments field, "Probably the biggest challenge will be determining what is in scope for Aquifer and then figuring out how to filter based on that." The DLF Aquifer Collections Working Group had proposed that Aquifer be a database of materials related to American history and culture. Some questions sought to gauge the extent of out-of-scope materials that might need to be removed from the harvested sets and the difficulty of doing this filtering. The task group recognized that some DLF members are themselves aggregators.

Additional questions probed the difficulty of filtering out duplicate records that might occur if both the aggregator and originating repository contributed records to Aquifer.

***Out-of-Scope Records.*** The first group of questions in this section sought to gauge the need to filter out out-of-scope records. Questions 1 and 2 asked whether the repositories had organized their records into sets, and if so, what those sets were based on. More than three-quarters of the repositories reported that their records were in sets (10 of 13,) but most of these repositories listed more than one basis for sets within their institutions (e.g., topic and provenance.) One repository noted in the comments field that they would soon be implementing more robust set descriptions for their image sets "which should make it easier to select content appropriate for the Aquifer collection."

Questions 3 and 4 asked repositories to indicate whether their OAI sets would need to be filtered to remove out-of-scope materials and to estimate the percentage of the records that would be in-scope. Eleven of the 13 repositories estimated that only a portion of their records would fall within the Aquifer scope, with the remaining two repositories responding that they are not sure. Most sets, therefore, would need to be filtered for content if the goal was to aggregate only records pertinent to the defined scope. On the other hand, it appears that a majority of all of the records would be within scope. Only two respondents reported that less than 50% of the records in the sets they plan to contribute would fall within scope, and 8 of the respondents reported that at least 75% of their records would fall within the scope.

***Duplicate Records from other Aggregators.*** The second group of questions looked at possible duplication resulting from the contribution of aggregated records. While almost 40% of the respondents are aggregators (5 of 13,) only one of these repositories plans to contribute aggregated records. In addition, those metadata records include a provenance statement, so it should be simple to filter them out if both the originating repository and aggregator are contributing records to Aquifer.

***Repositories' Ability to Filter.*** The final group of questions in this section asked whether the repository might be able to filter records prior to harvesting. Only one repository was willing to commit to filtering records prior to harvesting, while five indicated that they would not be able to do so. One repository commented that "a lot would depend on how the filtering task is defined."

## Obstacles and Solutions

If Aquifer implements the collection scope as defined, the collections contributed by most of the repositories will need to be filtered to remove out-of-scope records. This filtering may be difficult to automate because there is variety both within repositories and between repositories in how sets are determined and, as discussed below, in metadata practices. Most repositories indicated they could not commit to filtering their records prior to contributing them.

The task group proposes the following solutions:

1) Allow a certain percentage of out-of-scope records within a particular set. This policy would obviate the need to filter some sets without compromising the subject strength that would make Aquifer attractive to researchers.

2) Sponsor or encourage the development of automated clustering classifications and filtering (building on the work done at Emory, CDL, and Michigan.) Such systems would bear further fruit down the road as Aquifer moves to a broader base of repositories that may have even greater variety in set definition and metadata practices.
3) Explore the use of social tagging or something similar to Amazon's Turk to do the classification and filtering using volunteers. Aquifer might need to develop an incentive or reward structure to encourage volunteers to participate. (For more on Amazon's Turk, see http://www.mturk.com/mturk/welcome.)

Based on this survey, it does not appear that duplication resulting from the contribution of aggregated collections will be an obstacle during the initial harvests for Aquifer. As Aquifer grows and continues to accept new content, however, this could become a more significant obstacle.

The task group proposes the following solution:
1) Require that repositories contributing aggregated records identify the provenance of the records in the metadata. This policy would ensure that the records could easily be filtered out if the institution that contributed the records to the aggregator decided to contribute the records directly to Aquifer.

## *Section 2. Metadata*

The Gap Analysis Task Group recognizes that most institutions are not already creating MODS records adhering to the draft DLF Aquifer MODS Implementation Guidelines. Hence, this section of the survey sought to identify obstacles to simply providing MODS metadata, rather than obstacles to conforming specifically to the DLF guidelines.

### Discussion

The Metadata section explored five components of metadata practice in the repositories: (1) collection- or set-level metadata; (2) item-level metadata; (3) adherence to standards; (4) impact of DLF Aquifer MODS Implementation Guidelines; and (5) staff resources for metadata transformation. It should be noted that only 12 of the 13 survey respondents answered questions in the metadata section. The figures cited throughout this section of the report are based on the group of 12 respondents.

***Collection- or Set-level Metadata.*** Collection- and set-level metadata play a critical role in filtering records for inclusion in the DLF Aquifer aggregation. Questions 10 and 11 addressed this topic. Question 10 asked whether institutions create collection- or set-level metadata, while Question 11 identified which schemas the repositories use. Twelve institutions responded to these two questions. One hundred percent of respondents create collection- or set-level metadata. Half of the institutions (6 of 12) use Dublin Core for set-level description; five of 12 institutions use MARC. Other schemas utilized by respondents for set-level description include 1) EAD (3 of 12); 2) TEI (2 of 12); 3) MODS (1 of 12); and 4) local format (1 of 12.)

***Item-level metadata.*** Question 15 asked respondents which metadata schemas are in use for item-level description. All 12 respondents answered this question. While most repositories use more than one descriptive metadata schema, Dublin Core is overwhelmingly the most heavily used item-level schema with 83.3% of respondents (10 of 12) indicating that they use DC. Two schemas tied as next most heavily used: MODS and TEI (41.7% or 5 of 12 respondents,) along with local formats which also are utilized by 36.4% of the institutions surveyed. Other formats identified were MARC, used by 33.3% of respondents, and VRA Core, used by 16.7% of respondents.

***Adherence to Standards.*** This part of the Metadata section explored the extent to which repositories adhere to community standards that impact metadata quality and consistency, such as controlled vocabularies and data content standards. Question 12 inquired whether repositories use controlled vocabularies, while Question 13 identified which ones were in use, and whether the authorities are indicated in the markup. All 12 respondents answered Question 12; only 11 answered Question 13. Nearly all respondents (91.7% or 11 of 12) do use controlled vocabularies for populating metadata element content, while only 8.3% do not (1 of 12.) Controlled vocabularies in use by 11 repositories that gave feedback on this question include: LCSH (90.9%,) TGM (54.5%,) AAT (45.5%,) LCNAF (36.4%) and TGN (27.3%.) Additionally, 18% of respondents (2 of 11) indicated that locally controlled vocabularies were used in their collections.

Questions 16 and 17 dealt with data content standards, asking respondents to indicate whether they adhere to any such standards, and if so, to identify which standards were in use. All 12 respondents answered Question 16, indicating whether or not they use data content standards. The majority (66.7% or 8 of 12) do use data content standards for formulating metadata element content. A small group (25% or 3 of 12) does not use content standards; one respondent was unsure. When asked to identify which data content standards were in use, only eight repositories responded. Of those eight institutions, AACR2 was the most prevalent with 87.5% repositories adhering to this standard. The next most common standard was DACS (37.5% or 3 of 8 respondents), followed by local standards (25%); APPM (25%); and CCO (12.5%.) It should also be noted here that within individual institutions, a variety of content standards are in use across a variety of collections.

***Impact of DLF Aquifer MODS Implementation Guidelines.*** Question 18 posed an open-ended question, "What impact, if any, will the proposed Aquifer MODS Guidelines have on your repository's development of local MODS implementation?" Eleven of the 12 repositories responded to this question. While some respondents were unsure how the DLF guidelines would impact local practices, many expressed interest in creating and sharing Aquifer-compliant MODS records. One respondent indicated "Has already had a large impact. We are trying to implement robust MODS records in our digital collections from here forward." Another said, "[The guidelines] will help us iteratively improve our practices to create better shareable metadata."

While most of the respondents maintained an open posture towards the DLF Guidelines, the level of receptivity varied. Among the four institutions already creating MODS records, three indicated great interest in adhering to the DLF Guidelines once they are completed. However,

one of those four institutions indicated that the DLF Guidelines will have no impact on their implementation of MODS. Among the institutions not already creating MODS records, their responses fell primarily into two categories: (1) uncertainty about what how the guidelines might impact their local practices since the guidelines are still in draft status; and (2) hope that they will be able to adhere to the guidelines. Uncertainty was also sometimes coupled with interest in developing MODS metadata. As one respondent said, "Not sure at the moment, but we are very interested in supporting MODS." Yet another respondent indicated, "We will attempt to meet the minimal requirements outlined in the DLF Aquifer MODS Implementation Guidelines." Still another respondent indicated the repository was very close to sharing MODS records, but software as well as the draft status of the DLF guidelines poses significant obstacles, "We're close to making our metadata available via MODS. Bottlenecks are changing our software (broker) and awaiting the newly revised Aquifer MODS profile, based on the recent changes to the MODS profile."

***Staff Resources for Metadata Transformation***. Question 19 asked whether the repositories could devote staff time to the process of transforming non-MODS metadata into MODS. The respondents were very much divided in terms of whether their institutions could devote the human resources necessary to transform existing records into MODS. Forty percent indicated they could commit staff time to transforming records. Another forty percent said they were "Not sure" whether staff time could be allocating for this transformation. Twenty percent indicated they were not able to commit human resources to this work.

## Obstacles and Solutions

Most institutions do not have MODS records to share. This represents the most significant metadata-related obstacle to participating in the DLF Aquifer aggregation. Additionally, a variety of metadata standards are in use both between and within individual institutions. Surfacing MODS records for aggregation is also complicated by software limitations (discussed further in Section 4. of the report) and the lack of human resources to devote to transforming non-MODS records into MODS. Lastly, the MODS standard itself changes often, while the DLF Aquifer MODS Implementation Guidelines are still in development. Keeping up with the many changes in the standard itself, as well as following the development of the DLF guidelines presents unique challenges for institutions with limited resources.

The task group proposes the following solutions:

1) Release the completed DLF Aquifer MODS Implementation Guidelines. These guidelines will clarify the requirements. Institutions waiting for the final guidelines will then be able to move forward with implementation.
2) Develop levels of compliance for DLF Aquifer MODS Implementation Guidelines so institutions have clear guidance on what metadata is minimally required to participate in the aggregation.
3) Explore offering training in DLF Aquifer MODS Implementation Guidelines after finalized.
4) Point repositories to existing tools for metadata transformation, such as the LOC's MARC to MODS stylesheet.
5) Encourage and sponsor the development of additional tools for metadata transformation, such as stylesheets for transforming other commonly used descriptive metadata standards into MODS (for instance, TEI Header.)

## *Section 3. Technical Infrastructure*

This section of the survey was designed to investigate the OAI-PMH technical infrastructure available to possible Aquifer content providers. One respondent failed to make it this far in the survey, so results are based on 12 responses.

### Discussion

***Tool Limitations.*** Based on previous work, we knew that various digital repository systems have limitations relative to the OAI-PMH that could affect a participant's ability to meet the aggregation requirements of Aquifer. The most likely limitation would be a lack of support for the MODS metadata format. There were a few questions on the survey designed to address the issue. These included Question 21, "What tools are you using to implement your OAI data provider," Question 27, "Are there any hardware or software limitations that are preventing you from exposing MODS," and Question 24, "If your current system does not allow you to expose MODS records for harvest, is there someone with expertise in XML who could build a static repository for your MODS records?"

In general, the responses to these questions were mildly encouraging. Not surprisingly there are many different tools in use by the participants with some participants using different tools for their various OAI data providers. Interestingly, over half the respondents indicate that at least one of their OAI data providers was developed locally or was a local customization of a third-party OAI data provider. This is encouraging since it implies that given the right motivation and resources those institutions should be able to modify those tools to meet Aquifer requirements. The assumption being that it is usually much easier to modify a locally developed tool than it is to modify a third party tool.

There were only three yes responses to Question 27, "Are there any hardware or software limitations that are preventing you from exposing MODS?" However, there were five responses to the "If yes, please describe." text field. Most of these responses indicated that the particular tool being used, namely CONTENTdm, DSpace, or DLXS, did not support MODS. Michigan indicated that they are in the process of adding MODS support to DLXS. Another repository stated that since they are an aggregator they have little control over the format of the records, so MODS support would be problematic.

***Ability to Overcome Tool Limitations.*** Question 24 asked "If your current system does not allow you to expose MODS records for harvest, is there someone with expertise in XML who could build a static repository for your MODS records?" All respondents except one answered yes to this question. Question 30 asked the repositories to identify the level of proficiency their staff has with the OAI protocol. All respondents report library and technical staff with a high (58%) or medium (42%) level of proficiency with the OAI-PMH.

***Commitment to Sharing Metadata.*** There were also a series of questions designed to gauge an institution's ongoing interest or commitment to sharing metadata via the OAI protocol. The responses to all of these questions were very positive. All 12 respondents report active and ongoing support for the OAI-PMH, with all having growing collections, most of which are also available via the web.

To Question 28, "What is your priority for sharing metadata via OAI in general," 75% responded high and 25% responded medium. However, the response to Question 29, "What is your priority for sharing MODS records via OAI in particular," was not as encouraging with 33% high, 42% medium, and 25% low.

### Obstacles and Solutions

The good news is that the ongoing interest, commitment, and technical proficiency for sharing metadata via the OAI-PMH are widespread. However, there are clearly some technical limitations that impact an institution's ability to provide data to the Aquifer aggregation. The primary limitation being some tools' inability to natively provide metadata in the MODS format. Apparently there is a MODS patch for DSpace version 1.3.2, and MODS support is built into DSpace version 1.4, but not everyone is aware of or willing to adopt these patches or versions. Michigan is actively pursuing the addition of MODS support for their DLXS tool, but programming support for this enhancement seems to be limited. Future support for MODS in CONTENTdm is currently not known. Presumably many of the locally developed OAI implementations also lack support for MODS, but there appears to be some willingness to correct this or to provide an OAI static repository with MODS metadata.

The task group proposes the following solutions:

1) Reach out to tool vendors, such as for CONTENTdm, to encourage support of MODS.
2) Educate data providers on the capabilities of certain tools and encourage them to utilize these capabilities as appropriate, such as using DSpace's MODS support.
3) Sponsor or encourage the development of additional crosswalks and XSLT transformations from different formats into MODS, for example from the native CONTENTdm export into MODS.
4) Sponsor or encourage the development and deployment of tools to assist with the creation of OAI repositories both static and dynamic, such as Emory's Metadata Migrator Tool (http://metacluster.library.emory.edu/mosc/), UIUC's FileMakerPro Gateway and Z39.50 Gateway (http://zmarco.sourceforge.net/), or tools similar to CDL's 7train METS Generation Tool (http://seventrain.sourceforge.net/) for CONTENTdm (except for OAI instead of METS.)

## *Section 4. Rights*

The final section of the survey sought to gauge whether repositories believed that they had sufficient rights and permissions to contribute metadata and digital objects to Aquifer and to allow Aquifer to transform those files as needed. One of the respondents did not answer any of the questions in this section.

### Discussion

Questions 32 and 33 asked whether the repository had the authority to contribute the records and to allow Aquifer to transform those files to meet Aquifer standards. Four-fifths of the respondents were confident that their repository had the necessary rights in both areas (10 of 12.) The remaining two repositories were not certain if they had sufficient rights. The confidence expressed by the majority probably stems from the fact that the repositories went through a rights

review process prior to exposing the metadata and digital objects to end-users via a searchable/browseable interface on the web. Eight of the ten repositories had made their collections available via the web. Of these repositories, one was not sure about sharing the data and two were not sure that they had the rights to allow Aquifer to transform it.

Question 35 asked whether the repositories would be able to commit librarians and/or legal counsel to resolving rights issues. Most respondents were unsure (6 of 11.) Only one respondent indicated that their repository could offer support in this area.

Question 36 solicited additional comments on rights issues. One respondent stated that his repository would not contribute records that it did not feel it had the rights to contribute. Another respondent noted that his repository might need to secure additional rights to contribute some records. A third respondent asked, "As for [question] 35,[1] if OAISTER and the UIUC project didn't meet with these challenges, Aquifer will? Why?"

## Obstacles and Solutions

There are two considerations in the area of rights. The first concerns the amount of risk the DLF is willing to assume. The second concerns the amount of risk the repositories are willing to assume. (Given that copyright status can be murky, there is often some risk to be assumed.) Aquifer will need to decide how it will address the challenge posed by copyright; managing the risks while taking full advantage of rights to use materials for educational purposes as defined in copyright law.

It appears that from the perspective of the repositories rights issues will not be an obstacle for the first Aquifer harvest. Most of the repositories have already exposed their records to researchers and exposing them through Aquifer as well may pose little additional risk. They believe that they have or can easily acquire the rights to share the records through Aquifer. Those wishing to contribute aggregated records would need to make sure that they had secured the permission of the originating repository. It seems likely that most repositories would not be able to contribute staff time towards working on any rights issues that arise or are uncovered.

The task group proposes the following solutions:

1) Develop a deposit agreement that indemnifies the DLF. If the repositories are willing to sign the agreement, that would place the onus of sorting out the rights issues on the repositories.
2) Develop suites of tools to aid repositories in securing rights as appropriate; including standard procedures for determining rights for a given type of work, resources for determining copyright status, resources for locating copyright holders, sample licensing agreements, and (in the case of copyright orphans) mechanisms for copyright holders to notify repositories of possible infringements on their sites. Such tools would make it easier for repositories to work through rights issues.

[1] "Can your repository commit librarians and/or legal counsel to resolving any rights issues related to adding your metadata and/or digital objects to the Aquifer collection?"

## *Appendix 1: Participating Institutions*

- California Digital Library (University of California Office of the President)
- Emory University
- Indiana University
- University of Chicago Library
- University of Illinois at Urbana-Champaign
- University of Michigan
- University of North Carolina at Chapel Hill
- University of South Carolina
- University of Tennessee
- University of Texas at Austin
- University of Virginia Library
- University of Washington

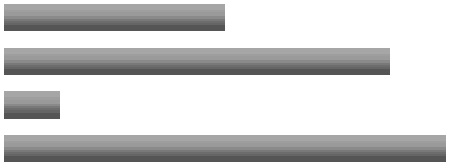
One other institution participated but did not identify itself.

## *Appendix 2: Results Summary from SurveyMonkey*

1. Are your collections grouped into OAI sets?

| | Response Percent | Response Total |
|---|---|---|
| **Yes** | **76.9**% | **10** |
| No | 7.7% | 1 |
| Not sure | 15.4% | 2 |
| **Total Respondents** | | **13** |
| **(skipped this question)** | | **0** |

2. If yes, what are the sets based on? Check all that apply.

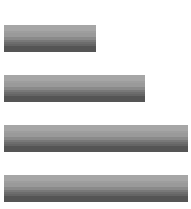


| | Response Percent | Response Total |
|---|---|---|
| Provenance | 36.4% | 4 |
| Topic | 63.6% | 7 |
| Date span | 9.1% | 1 |
| **Other (please specify)** | **72.7**% | **8** |
| **Total Respondents** | | **11** |
| **(skipped this question)** | | **2** |

3. Do only a portion of the sets fall within Aquifer's scope (American history and culture)?

| | Response Percent | Response Total |
|---|---|---|
| **Yes** | **84.6**% | **11** |
| No | 15.4% | 2 |
| Not sure | 0% | 0 |
| **Total Respondents** | | **13** |
| **(skipped this question)** | | **0** |

4. If only a portion is within scope, approximately how much is within scope?



| | Response Percent | Response Total |
|---|---|---|
| <10 % | 0% | 0 |
| 25% | 15.4% | 2 |
| 50% | 23.1% | 3 |
| **75%** | **30.8**% | **4** |
| **100%** | **30.8**% | **4** |
| **Total Respondents** | | **13** |

| | | |
|---|---|---|
| **(skipped this question)** | | **0** |

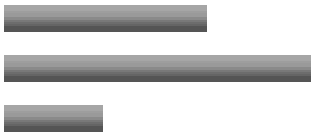
5. Are you an aggregator harvesting records from other repositories?

| | Response Percent | Response Total |
|---|---|---|
| Yes | 38.5% | 5 |
| **No** | **61.5%** | **8** |
| Not sure | 0% | 0 |
| **Total Respondents** | | **13** |
| **(skipped this question)** | | **0** |

6. If you are an aggregator, do you plan to include any of those records in your contribution to Aquifer?

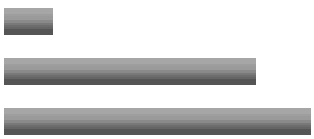
| | Response Percent | Response Total |
|---|---|---|
| Yes | 14.3% | 1 |
| **No** | **85.7%** | **6** |
| Not sure | 0% | 0 |
| **Total Respondents** | | **7** |
| **(skipped this question)** | | **6** |

7. If you plan to contribute aggregated records to Aquifer, are you adding a provenance statement so that those records might be identified as potential duplicates?



| | Response Percent | Response Total |
|---|---|---|
| Yes | 33.3% | 2 |
| **No** | **50%** | **3** |
| Not sure | 16.7% | 1 |
| **Total Respondents** | | **6** |
| **(skipped this question)** | | **7** |

8. Can your repository commit catalogers and programmers towards filtering the records held in your repository before they are harvested by Aquifer? (e.g., to remove records that are out of scope or that may be exposed by more than one repository)



| | Response Percent | Response Total |
|---|---|---|
| Yes | 8.3% | 1 |
| No | 41.7% | 5 |
| **Not sure** | **50%** | **6** |
| **Total Respondents** | | **12** |

| | |
|---|---|
| **(skipped this question)** | **1** |

9. Are there any other obstacles or challenges related to content that the working group should consider?

| | |
|---|---|
| **Total Respondents** | **6** |
| **(skipped this question)** | **7** |

10. Do you create collection- or set-level metadata?

| | Response Percent | Response Total |
|---|---|---|
| **Yes** | **100**% | **12** |
| No | 0% | 0 |
| Not sure | 0% | 0 |
| **Total Respondents** | | **12** |
| **(skipped this question)** | | **1** |

11. If so, which schema is used (Dublin Core, EAD, etc.)? Check all that apply.

| | Response Percent | Response Total |
|---|---|---|
| Dublin core | 50% | 6 |
| EAD | 25% | 3 |
| METS | 0% | 0 |
| MODS | 8.3% | 1 |
| TEI | 16.7% | 2 |
| **Other (please specify)** | **58.3**% | **7** |
| **Total Respondents** | | **12** |
| **(skipped this question)** | | **1** |

12. Do you use controlled vocabularies?

| | Response Percent | Response Total |
|---|---|---|
| **Yes** | **91.7**% | **11** |
| No | 8.3% | 1 |
| Not sure | 0% | 0 |
| **Total Respondents** | | **12** |
| **(skipped this question)** | | **1** |

13. If so, which ones? Check all that apply.

| | Response Percent | Response Total |
|---|---|---|
| AAT | 45.5% | 5 |
| LCNAF | 36.4% | 4 |
| **LCSH** | **90.9%** | **10** |
| Thesaurus for Graphic Materials | 54.5% | 6 |
| Thesaurus for Geographic Names | 27.3% | 3 |
| Other (please specify) | 18.2% | 2 |
| | **Total Respondents** | **11** |
| | **(skipped this question)** | **2** |

14. Are the controlled vocabularies indicated within the markup?

| | Response Percent | Response Total |
|---|---|---|
| **Yes** | **63.6%** | **7** |
| No | 36.4% | 4 |
| Not sure | 0% | 0 |
| | **Total Respondents** | **11** |
| | **(skipped this question)** | **2** |

15. Which descriptive metadata formats do you use at the item-level? Check all that apply.

| | Response Percent | Response Total |
|---|---|---|
| **Dublin Core** | **83.3%** | **10** |
| MARC | 33.3% | 4 |
| MEP | 0% | 0 |
| MOA II | 0% | 0 |
| MODS | 41.7% | 5 |
| TEI | 41.7% | 5 |
| VRA Core | 16.7% | 2 |
| Other (please specify) | 33.3% | 4 |
| | **Total Respondents** | **12** |
| | **(skipped this question)** | **1** |

16. Are you adhering to particular data content standards?

| | Response Percent | Response Total |
|---|---|---|
| **Yes** | **66.7%** | **8** |
| No | 25% | 3 |

| | | |
|---|---|---|
| Not sure | 8.3% | 1 |
| **Total Respondents** | | **12** |
| **(skipped this question)** | | **1** |

17. If so, which ones?

| | Response Percent | Response Total |
|---|---|---|
| **AACR2** | **87.5%** | **7** |
| APPM | 25% | 2 |
| CCO | 12.5% | 1 |
| DACS | 37.5% | 3 |
| Other (please specify) | 25% | 2 |
| **Total Respondents** | | **8** |
| **(skipped this question)** | | **5** |

18. What impact, if any, will the proposed Aquifer MODS guidelines have on your repository's development of local MODS implementation?

| | |
|---|---|
| **Total Respondents** | **11** |
| **(skipped this question)** | **2** |

19. If your metadata is not already in MODS, can your repository commit metadata librarians, catalogers and programmers towards transforming it into MODS?

| | Response Percent | Response Total |
|---|---|---|
| Yes | 36.4% | 4 |
| No | 18.2% | 2 |
| Not applicable | 0% | 0 |
| **Not sure** | **45.5%** | **5** |
| **Total Respondents** | | **11** |
| **(skipped this question)** | | **2** |

20. Are there any other obstacles or challenges related to metadata that the working group should consider?

| | |
|---|---|
| **Total Respondents** | **6** |
| **(skipped this question)** | **7** |

21. What tools are you using to implement your OAI data provider? Check all that apply.

| | Response Percent | Response Total |
|---|---|---|
| ContentDM | 25% | 3 |

| | | |
|---|---|---|
| DSpace | 8.3% | 1 |
| **Locally developed** | **58.3%** | **7** |
| Other (please specify) | 33.3% | 4 |
| | **Total Respondents** | **12** |
| | **(skipped this question)** | **1** |

22. Is the OAI provider part of an active program or was it created as part of a project that has concluded?

| | Response Percent | Response Total |
|---|---|---|
| **Yes** | **100%** | **12** |
| No | 0% | 0 |
| Not sure | 0% | 0 |
| | **Total Respondents** | **12** |
| | **(skipped this question)** | **1** |

23. Is staff available to maintain the OAI provider on a ongoing basis?

| | Response Percent | Response Total |
|---|---|---|
| **Yes** | **100%** | **12** |
| No | 0% | 0 |
| Not sure | 0% | 0 |
| | **Total Respondents** | **12** |
| | **(skipped this question)** | **1** |

24. If your current system does not allow you to expose MODS records for harvest, is there someone with expertise in XML who could build a static repository for your MODS records?

| | Response Percent | Response Total |
|---|---|---|
| **Yes** | **81.8%** | **9** |
| No | 18.2% | 2 |
| Not sure | 0% | 0 |
| | **Total Respondents** | **11** |
| | **(skipped this question)** | **2** |

25. Are your OAI records accumulating or growing regularly?

| | Response Percent | Response Total |
|---|---|---|
| **Yes** | **100%** | **12** |

| | | |
|---|---|---|
| No | 0% | 0 |
| Not sure | 0% | 0 |
| **Total Respondents** | | **12** |
| **(skipped this question)** | | **1** |

26. Are the collections currently exposed to end-users via a searchable/browseable interface on the web?

| | Response Percent | Response Total |
|---|---|---|
| **Yes** | **83.3**% | **10** |
| No | 16.7% | 2 |
| Not sure | 0% | 0 |
| **Total Respondents** | | **12** |
| **(skipped this question)** | | **1** |

27. Are there any hardware or software limitations that are preventing you from exposing MODS?

| | Response Percent | Response Total |
|---|---|---|
| Yes | 25% | 3 |
| **No** | **50**% | **6** |
| Not sure | 16.7% | 2 |
| If yes, please describe. | 41.7% | 5 |
| **Total Respondents** | | **12** |
| **(skipped this question)** | | **1** |

28. What is your priority for sharing metadata via OAI in general?

| | Response Percent | Response Total |
|---|---|---|
| **High** | **75**% | **9** |
| Medium | 25% | 3 |
| Low | 0% | 0 |
| **Total Respondents** | | **12** |
| **(skipped this question)** | | **1** |

29. What is your priority for sharing MODS records via OAI in particular?

| | Response Percent | Response Total |
|---|---|---|
| High | 33.3% | 4 |
| **Medium** | **41.7**% | **5** |

| | | |
|---|---|---|
| Low | 25% | 3 |
| **Total Respondents** | | **12** |
| **(skipped this question)** | | **1** |

30. What level of proficiency do your librarians or IT staff assigned to digital repositories have with the OAI protocol?

| | Response Percent | Response Total |
|---|---|---|
| **High** | **58.3%** | **7** |
| Medium | 41.7% | 5 |
| Low | 0% | 0 |
| **Total Respondents** | | **12** |
| **(skipped this question)** | | **1** |

31. Are there any other obstacles or challenges related to technical infrastructure that the working group should consider?

| | |
|---|---|
| **Total Respondents** | **4** |
| **(skipped this question)** | **9** |

32. Can your repository certify that it has the authority to deposit the metadata records and digital objects in a publicly available DLF Aquifer collections catalog?

| | Response Percent | Response Total |
|---|---|---|
| **Yes** | **83.3%** | **10** |
| No | 0% | 0 |
| Not sure | 16.7% | 2 |
| **Total Respondents** | | **12** |
| **(skipped this question)** | | **1** |

33. Does your repository have the authority to allow Aquifer to transform the records to meet Aquifer standards?

| | Response Percent | Response Total |
|---|---|---|
| **Yes** | **83.3%** | **10** |
| No | 0% | 0 |
| Not sure | 16.7% | 2 |
| **Total Respondents** | | **12** |
| **(skipped this question)** | | **1** |

34. What might your repository need to do to acquire the needed rights to deposit the metadata and/or digital objects?

Check all that apply.

| | Response Percent | Response Total |
|---|---|---|
| **Seek permission from the rights holder** | **71.4%** | **5** |
| Seek permission from the repository holding the originals | 28.6% | 2 |
| Other (please specify) | 0% | 0 |
| **Total Respondents** | | **7** |
| **(skipped this question)** | | **6** |

35. Can your repository commit librarians and/or legal counsel to resolving any rights issues related to adding your metadata and/or digital objects to the Aquifer collection?

| | Response Percent | Response Total |
|---|---|---|
| Yes | 18.2% | 2 |
| No | 27.3% | 3 |
| **Not sure** | **54.5%** | **6** |
| **Total Respondents** | | **11** |
| **(skipped this question)** | | **2** |

36. Are there any other obstacles or challenges related to rights that the working group should consider?

| | |
|---|---|
| **Total Respondents** | **4** |
| **(skipped this question)** | **9** |

37. Name of your institution:

| | |
|---|---|
| **Total Respondents** | **12** |
| **(skipped this question)** | **1** |

38. May we contact you for further information? If so, please enter your name and contact information below.

| | Response Percent | Response Total |
|---|---|---|
| **Name:** | **100%** | **12** |
| **Position/Title:** | **100%** | **12** |
| **Email:** | **100%** | **12** |
| **Phone:** | **100%** | **12** |
| **Total Respondents** | | **12** |
| **(skipped this question)** | | **1** |

39. Would you like your institution’s responses to be anonymous in the results reported back to the participating repositories?

| | Response Percent | Response Total |
|---|---|---|
| Yes | 25% | 3 |
| **No** | **75%** | **9** |
| **Total Respondents** | | **12** |
| **(skipped this question)** | | **1** |