

# **RDA/ADHO Workshop: Evaluating Research Data Infrastructure Components and Engaging in their Development**

**July 12, 2016 9:30am-4:00pm**

This workshop took place at the Digital Humanities  
2016 conference in Kraków, Poland, July 11-16.

Report prepared by: Lindsay Poirier, Kim Fortun, and  
Bridget Almas

## Table of Contents

<b>WORKSHOP LEADERS</b>	<b>3</b>
<b>SUMMARY</b>	<b>3</b>
<b>AGENDA</b>	<b>4</b>
<b>PARTICIPANTS</b>	<b>5</b>
<b>BURSARIES</b>	<b>6</b>
<b>EXPOSURE TO RDA</b>	<b>6</b>
<b>OUTCOMES</b>	<b>6</b>
<b>PID TYPES</b>	<b>7</b>
<b>DATA TYPES REGISTRY</b>	<b>8</b>
<b>DISCUSSION</b>	<b>9</b>
<b>NANO</b>	<b>9</b>
<b>DATA</b>	<b>11</b>
<b>TECHNO</b>	<b>12</b>
<b>MICRO</b>	<b>13</b>
<b>MESO</b>	<b>14</b>
<b>MACRO</b>	<b>16</b>
<b>META</b>	<b>17</b>
<b>CONCLUSION</b>	<b>18</b>

## Workshop Leaders

- Bridget Almas, Tufts University, [bridget.almas@tufts.edu](mailto:bridget.almas@tufts.edu);
- Kim Fortun, Rensselaer Polytechnic Institute, [fortuk@rpi.edu](mailto:fortuk@rpi.edu);
- Natalie Harrower, Digital Repository of Ireland, [n.harrower@ria.ie](mailto:n.harrower@ria.ie);
- Eveline Wandl-Vogt, Austrian Academy of Sciences, Dariah-EU, [Eveline.Wandl-Vogt@oeaw.ac.at](mailto:Eveline.Wandl-Vogt@oeaw.ac.at)
- Lindsay Poirier, Rensselaer Polytechnic Institute, [poiril@rpi.edu](mailto:poiril@rpi.edu)

## Summary

The purpose of this workshop was to conduct a meaningful examination of the data fabric and infrastructure components being defined by the Research Data Alliance (RDA), to test their relevance and applicability to the needs of the digital humanities community, and to discuss opportunities for humanities engagement in further standards development.

Much of the infrastructure needed to support data sharing is of great relevance to Digital Humanities projects, where we find ourselves too often developing and reinventing ad-hoc solutions for data management, draining resources that could be put to better use focusing on the domain-specific nature of our problems and driving new research. It's easy, especially when time and resources are constrained, to get locked into thinking that our problems are unique and that we need to design custom solutions, but when we examine the problem from other perspectives, the abstractions begin to rise to the surface. But in order to take advantage of the solutions as they are built, we must be part of the discussion about the requirements, push for our use cases to be considered in their design, and take part in testing, implementing and sustaining the solutions.

RDA is an international initiative to facilitate the development of effective data practices, standards and infrastructure in particular research domains, and across domains. It aims to enhance capacity to archive, preserve, analyze and share data, and for collaboration both within and across research communities. The humanities have an important presence in RDA, and can benefit from the opportunities RDA provides to learn across research communities working to develop digital infrastructure. RDA also brings together diverse types of technical expertise, which is organized to put forward (best practice) "adoption products." Some of these products are starting to be taken up in the humanities, such as the Practical Policy Recommendations mentioned below, and

there is significant potential for further collaborative work between RDA and digital humanities developers in the future.

This was a full day workshop in the format of a hands-on round-table and open discussion. Participants were asked to come prepared to discuss details of their particular use cases, as well as solutions and needs relevant to two of the initial outputs of the RDA: the Persistent Identifier (PID) Types and Data Types Registry (DTR). In advance of the workshop, organizers provided summaries of these outputs and detailed examples of their analysis for use cases in humanities and other relevant domains.

This workshop was a complement to the panel by Dr. Natalie Harrower et. al. entitled *“Digital data sharing: the opportunities and challenges of opening research.”* The panel presented particular challenges in humanities research data management, and aimed to generate a discussion around the uniqueness and challenges inherent in humanities research data. This workshop, on the other hand, was a hands-on effort to work with real humanities data use-cases, provided by participants, to understand how to best shape RDA outputs to enable better data sharing and management in the humanities.

## Agenda

### **9:30-10:15 (Kim)**

Introduction to RDA, RDA IG-DPHE, and why it’s been helpful for PECE

### **10:15-10:30**

On half sheets of paper, participants:

- Mapped out whom they would turn to for technical support in implementing RDA outputs.
- Outlined digital objects in their systems that they believe deserve PIDs.

### **10:30-10:50**

Coffee Break

### **10:50-11:50**

Presentations describing RDA outputs:

- PID types, DTR and Collections (Bridget)
- Practical Policies (Lindsay)

### **11:50-1:00**

Each participant responded to the following questions:

- Who are you, and where are you from?
- What is your interest in this session?

- Which, if any, of the outputs just discussed do you think are relevant for your use case?

### **1:00-2:30**

Lunch

### **2:30-3:15**

Bridget demonstrated how to implement the PID Types and Data Type Registries outputs in a gazetteer.

### **3:15-3:30**

Coffee Break

### **3:30-4:00**

Discussion

- What are the barriers to participating in RDA and implementing RDA outputs in the humanities?
- What issues do we need to address in order to enable data sharing, management, and long-term preservation in the humanities?

## **Participants**

- Smiljana Antonijevic, Penn State University
- Francesco Beretta, LARHRA, CNRS / Université de Lyon
- Dawn Childress, UCLA
- Tom Gheldof, KU Leuven / DARIAH-BE
- Rebecca Grant, Digital Repository of Ireland / Royal Irish Academy
- Leif Isaksen, Lancaster University
- Ian Johnson, University of Sydney
- Neven Jovanovic, University of Zagreb, Faculty of Humanities and Social Sciences
- Tibor Kalman, GWDG, DARIAH, epic
- Anastaisia Khaminova, Tomsk State University, Department of Humanitarian Problems of Informatics
- Asanobu Kitamoto, National Institute of Informatics
- Brandon Locke, Michigan State University
- Matthew Munson, DH University of Leipzig
- Matteo Romanello, German Archeological Institute / Swiss Federal Institute of Technology in Lausanne
- Sofia Papastamkou, Université Paris 1 – Panthéon-Sorbonne
- Silvia Piccini, Institute for Computational Linguistics, CNR, Pisa
- Thomas Padilla, University of California, Santa Barbara

- Matthias Schloegl, Austrian Centre for Digital Humanities
- Jakub Szport, Interdisciplinary Centre for Computational and Mathematical Modelling at the University of Warsaw
- Danah Tonne, Karlsruhe Institute of Technology
- André van Kooij, Utrecht University, Netherlands
- Zdenko Vozar, Charles University – CRHEC UPE Paris
- Mara Wade, University of Illinois
- John Wallance, Research Computing, Dartmouth College
- Hsiang-An Wang, Institute of Information Science, Academia Sinica
- Tariq Yousef, DH University of Leipzig

## Bursaries

RDA EU3 supported two bursaries of €500 each for participant travel to the workshop. These bursaries were awarded to Zdenko Vozar and Neven Jovanovic. To apply for these bursaries, applicants submitted a Statement of Interest and a Use Case. Applications were evaluated based on the relevance of the statement of interest to the workshop aims, the quality and relevance of the applicants' use cases, the relevance and importance of the applicants' projects, and the career status of the applicant (with preference for early career scholars).

## Exposure to RDA

Of the participants that responded to the workshop survey:

- 78% had heard of RDA prior to the workshop call
- 28% were members of RDA interest or working groups
- 28% had attended an RDA meeting
- 28% had adopted an RDA output<sup>1</sup>

## Outcomes

During the workshop we used a frequently shared digital humanities data type, a linked data record from the Pleiades Gazetteer, to guide us through a hands-on analysis of the applicability and adoptability of two RDA outputs: the Persistent Identifier Types and Data Types Registry. This was an interesting use case for the participants because linked

---

<sup>1</sup> Notably, while the same overall percentage of respondents were members of RDA groups, had attended meetings, or had implemented outputs, only 11% of respondents answered affirmatively to all three.

data scenarios using URLs as persistent identifiers are very prevalent across digital humanities projects.

## PID Types

For the analysis of the PID Types output, we identified a list of possible properties a PID Record for a linked data object identified by URL might need to include:

Property	Required
License	No
Formats Available	Yes
Access Method	Yes
Schema	No
Checksum	No
Creator	Yes
Created At	Yes
Previous Version	No
Publisher	Yes
Publication Date	Yes
Expiration Date	No
Reliability Measure	No
Data Type	Yes

Some of the identified properties themselves would likely to require definition of complex data types in a data types registry. For example, we could envision that the Formats Available property might have values identifying specific formats, such as XML, JSON, CSV, etc., which might need to be described further with additional properties including schema, etc. And the Data Type property might point at a complex data type, as in this case, a Gazetteer Record.

While the exercise did succeed in surfacing a number of metadata properties which would be essential to reuse of such data, it was a challenge for participants to fully appreciate the value of the PID Record and its relationship to the PIT API in this offline context. In order to fully identify the cost-benefit of applying such an approach to persistent data identifiers, it would be useful to have a testbed environment to which PID records of various data types could be added and then accessed and visualized through a common toolset via the PIT API.

## Data Types Registry

For the analysis of the Data Types Registry, we described the properties and purpose of a Pleiades Gazetteer record according to the suggested DTR schema. A sample of these details expressed in JSON, with temporary handle prefix identifiers is provided below:

```
{
  "identifier" : "20.5000.347/a2d1cdae060d561e877f",
  "name" : "Place Gazetteer Record",
  "description" : "Data record that disambiguates a conceptual place.",
  "standards" : [ {
    "issuer" : "DTR",
    "name" : "Pelagios Gazetteer Interconnection Format",
    "natureOfApplicability" : "specifies"
  } ],
  "provenance" : {
    "contributors" : [ {
      "identifiedUsing" : "Text",
      "name" : "RDA/ADHO DH 2016 Workshop Participants",
      "details" : "Collaborative effort by participants at the RDA/ADHO
Workshop at DH 2016 in Krakow."
    } ],
    "creationDate" : "2016-08-31T14:04:17.750Z",
    "lastModificationDate" : "2016-08-31T14:04:17.755Z"
  },
  "expectedUses" : [ "Used for providing a persistent, machine-readable,
interoperable identifier for a historical conceptual place." ],
  "representationsAndSemantics" : [ {
    "expression" : "Format",
    "value" : "GeoJSON"
  }, {
    "expression" : "Format",
    "value" : "KML"
  }, {
    "expression" : "Format",
    "value" : "CSV"
  }, {
    "expression" : "Format",
    "value" : "JSON-LD"
  } ],
  "properties" : [ {
    "name" : "location",
    "identifier" : "20.5000.347/6469ac84a89748bb67b9"
  }, {
```



```
"name" : "name variant",  
"identifier" : "20.5000.347/7677d6e6760296584931"  
} ]  
}
```

The need for a Data Types Registry resonated fairly well with the participants in the workshop. A few participants who were responsible for existing humanities infrastructure expressed confidence that their data types could be represented in this mode and that they might be able to provide some data to help seed such a registry. There were some concerns about how to avoid proliferation of types, but the need for, and value of for such an infrastructure component was largely undisputed.

One reasonable next step from this portion of the workshop would be to follow-up with those participants who expressed interest in providing data types to try to get some momentum going around establishment of some seed repositories.

## Discussion

During the closing discussion, we used the following image to describe the layers/scales at which there are challenges to enabling both data sharing and data infrastructure implementation in the digital humanities. In what follows, I use this chart to organize several points, discussed throughout the day, on why it is difficult to adopt RDA outputs in the humanities domain. I begin by describing the issues, questions, and challenges that each layer represents. For each layer, I then go on to outline points, discussed over the course of the day, that pertain to the relevant issues, questions, and challenges.

### Nano

*Nano-level analysis focuses attention on the frames, paradigms, and thought styles through which researchers carry out their work. The cultures of different disciplinary communities, the language ideologies that guide their work, and assumptions that they bring to their work all implicate the nano level.*

*Questions/Challenges that we must address at the nano-level include:*

- *How do we convince humanists that data-sharing is a worth-while endeavor, particularly in fields that are technology-averse or tend to work in isolation?*
- *What educational programs would orient the next generation of DH researchers towards seeing data sharing as, not only a valuable endeavor, but also an imperative one?*
- *How can researchers best deal with the paradoxical need for research data infrastructure customized in keeping with their own research tradition, yet with enough standardization to allow research data sharing among different communities?*

Towards the beginning of the workshop, we asked participants to draw out with detailed granularity all of the different units in their digital systems to which they would like to attach unique identifiers. As we suspected, despite all working towards the digital humanities, there was considerable diversity in the data types researchers identified.

Notably, participants in this workshop represented diverse humanities fields and traditions (e.g.

ethnographers, lexicographers, historians, archivists, literary scholars, librarians, etc.). While there is some overlap in the data such disciplines work with and the methods they use to curate and analyze the data, there are also considerable differences, which implicate how RDA outputs are implemented.

For instance, in discussing the data types registry RDA

output, there was a question of when differences between the way researchers understand their data types make enough of a difference to warrant creating a new data type. While recognizing the need to respect diverse research traditions, there was also concern that constantly evolving data types would not actually enable interdisciplinary data sharing. It would be helpful to digital humanities communities if such tricky considerations were addressed explicitly in the outputs, along with general guidelines or recommendations for how to proceed bearing in mind the data sharing trade-offs.

#### Other Nano-Level Data Challenges

One participant noted that often metadata has a very different purpose for humanists and librarians/archivists – with the former aiming to contextualize an object of knowledge, and the latter aiming to preserve it. Often when humanists access and curate data from a library or archive, they add additional metadata to the objects. However, they then find it difficult to reconcile different data standards in ways that enable putting the curated data back into the library or archive. Thus, they end up building layers above the archive for storing the new curated collections.

This anecdote, of course, points to much broader issues of concern for RDA (and for which I will not attempt to tackle in this report). How do we respect disciplinary differences, while also enabling and facilitating data sharing and long-term data preservation? To RDA's great credit, such epistemological issues have been on the leadership's radar for a long-time. As Mark Parson writes in 2014:

...infrastructure is never truly planned or designed up front. It evolves through a staged process that can involve complex dynamics, unanticipated consequences, and significant friction between individuals, organizations, and systems... **The Research Data Alliance embraces this complex dynamic.** It has no overarching plan or architecture, but it provides core principles and a "neutral place" that provide **enough alignment to move forward while still recognizing the value of friction.**

Mark Parsons, "Learning from Past Infrastructure to Embrace Friction and Create the Research Data Alliance" (American Geophysical Union, San Francisco, CA, December 16, 2014), <http://www.slideshare.net/ResearchDataAlliance/parsons-rda-agu2014>.

In a more general sense, RDA working groups have been particularly attentive to considering the humanities when defining their outputs, often coming to the Digital Practices in History and Ethnography Interest Group in search of a use case. However, it is important to bear in mind that there are many differences in data practice *within* the humanities and that one humanities use case should not stand in for all of the humanities.

## Data

*Data-level analysis focuses attention on data architecture and configuration – how information is organized (highlighting some things and sidelining others). As Kim Fortun noted, “At least two kinds of humanities work needs to be done at this level. Given the way data architecture and configurability implicates how meaning and value is produced, humanities scholars need to analyze data architectures and configurations as they have long studied texts and performance. Further, to enroll humanities scholars in the use of digital research infrastructure, such infrastructure will need to render embedded epistemological assumptions explicit so that researchers can select infrastructure in keeping with their theoretical commitments.”*

*Questions at this level include:*

- *How do we design data architecture to protect the diversity of analytic modes and thought styles across research traditions?*
- *How do we design and structure data architecture and configurations to fairly represent information?*

The data architecture level was addressed only briefly towards the end of the workshop. Specifically, one participant noted that while applied archaeologies tend to not express concern over how the design of data architectures reorder research practice, it is of great concern for theoretical archaeologists.

Notably, there is a great deal of literature in the digital humanities community itself that addresses issues at the data level – considering how data architectures harbor assumptions about social and scientific orderings and thus shape particular (and biased) digital representations of phenomena. A particularly interesting challenge for the RDA may be to consider how such research and writing could interface with working group outputs – perhaps by extending invitations to critical digital humanists to write blog posts for the RDA or by facilitating *book/journal article clubs* for RDA members interested in integrating critical perspectives into their working/interest group discussions.

## Techno

*Techno-level analysis focuses attention on technology design, functionality, skill, and sustainability. Many different types of technologies are needed to support data sharing – both material technologies (servers, databases, platforms) and non-material technologies (standards, classifications). Further, to make data sharing work, all these technologies need to be appropriately networked. This requires time, skill, funding, and space to store material infrastructures.*

*Questions at this level include:*

- How do we select which standards and technologies get enlisted in shared data infrastructure, or in other words, which become “best practice”?*
- How do we network these technologies in ways that work for interdisciplinary communities?*
- How do we factor the time to build technologies into research projects?*
- How do we ensure there are (affordable) places to put new technologies and data?*
- How do we ensure that the infrastructure (Internet access, bandwidth, and processing power) needed to support data sharing technologies are accessible to humanists?*

In a broad sense, the technical issues experienced by the digital humanities community are the same issues experienced by all disciplinary communities. How do we select “best practice” technologies? Where are we going to store our data? How do we appropriately network technologies to sustain data sharing?

I would argue that the particular challenges to the humanities community addressed in the workshop come at the interface of the technical level and several other levels. For instance, with depleting financial support for humanities research (macro level), who is going to fund the technologies that store data or the technologies that store services? What incentives do humanities scholars have to put time into both learning and maintaining digital infrastructure when it is so far divorced from the work that leads to professional advancement (meso level)? I draw out these issues as they arose in the workshop in other sections of this report.

## Micro

*Micro-level analysis focuses attention on practice – both data practice and research practice. Work in Science and Technology Studies has shown that attempts to create standards for data sharing inevitably creates “friction” between interdisciplinary communities. This is often because different communities follow different research workflows. Thus, In order to build infrastructure for data sharing, different research communities need to make tacit workflows explicit. This is particularly challenging for humanists, who have often never thought through their research workflows in computational terms.*

Questions at this level include:

- *How does data sharing change the research workflow?*
- *How do we build data workflows that align with diverse research methods and practices?*
- *How do we prepare researchers to work with data infrastructure – to integrate data best practice into their day-to-day work?*

Early in the day, participants were asked to map out whom they would turn to for technical support in implementing RDA outputs. The most notable finding from collecting and compiling this survey is that the degree and type of technical support available to digital humanists is extremely diverse across institutions and geographies.

For instance, at some institutions technical support is available through the research data centers in libraries, while at others, they would be more apt to turn to their IT departments for technical assistance. In the EU, many participants noted getting technical help from large institutions such as DARIAH and CLARIN. Still, at some institutions, technical support for implementing data infrastructure is largely not available and must either be self-taught or outsourced to programmers and developers.

This is a particularly significant challenge when considering the experience of some in actually working with RDA outputs. As one participant, well acquainted with RDA, stated:

...I find it hard to advocate for adopting the outputs because I might understand conceptually what they're for, but that doesn't mean I understand what needs to happen for them to be adopted. I think that a lot of the RDA outputs seem very technical. And I'm not from a technical background, so I don't care to bridge that gap, but it's a problem for research generally. Maybe not so much for the digital humanities; I'm not from a digital humanities background. But, like software, things like APIs and profiles for PID types, most archivists would not understand that at all. ... So I think that's something that's kind of missing from

RDA outputs documentation outputs a lot of the time. Like I can look at the information that's given and maybe understand it, but what's motivating me to do anything else?

Further, a participant noted that, it often takes her research team a long time to track down the right expertise to implement outputs, largely because they don't quite understand how to convey to developers what needs to be implemented.

Several suggestions were put forward for dealing with lack of technical support:

- One participant proposed that, for each output, RDA instate one traveling consultant, responsible for actively working with research groups and helping to translate the technical aspects of the output into terms that could be more broadly understood.
- Another suggestion was that RDA could help develop a registry with a vetted pool of technologists that digital humanists could leverage whenever they got a bit of money. This registry could model DHCommons<sup>2</sup> – a site listing available digital humanities collaborators, along with the skills they can bring to a collaboration.
- Finally, there was a suggestion to develop more user-driven tools for implementing RDA outputs, such as a tool for automatically implementing PIDs. (For example, the Platform for Experimental Collaborative Ethnography has used a bit of grant money to build a tool for automatically implementing the RDA output for PIDs, using the Archive Resource Key service, in Drupal.)<sup>3</sup>

## Meso

---

<sup>2</sup> <http://dhcommons.org/>

<sup>3</sup> <https://www.drupal.org/project/ark>

*Meso-level analysis focuses attention on organizations, and the way I see it there are organizational challenges for opening the empirical humanities at multiple scales. In one sense, building and leveraging data sharing infrastructure often requires new forms of both disciplinary and interdisciplinary collaboration. These new collaborations will require that data sharing communities evaluate how different researchers with different skill sets will work together and hold each other accountable, particularly when such work is often unpaid. To enable such collaborations, there also need to be organizational structures in place that incentivize both data sharing and participation in collaborative work. This requires much broader organizational evolution in what counts as “scholarly work.” Finally, recognizing that building and maintaining the infrastructure for global and interdisciplinary data sharing will require many different organizations and research communities working together, we need organizational structures that can network the appropriate expertise, support, and outputs.*

*Questions at this level include:*

- How do we ensure that appropriate incentives are in place to encourage researchers to participate in data sharing and infrastructure development?*
- What governance structures best support collaborative research projects?*
- What sorts of support systems and organizations are needed to both enable and encourage collaborative work?*
- How do we network diverse national and interdisciplinary efforts to build data infrastructure with each other and local data sharing communities?*

In a questionnaire sent out to researchers prior to the workshop, we asked, “What are your top three research data infrastructure needs that are NOT already being met?” The two most common answers to this question were (1) institutional repositories for long-term data storage, particularly at many stages of the research workflow (vs. only finalized data) and (2) institutional services for assigning persistent identifiers. These responses are notable in that they do not point to a lack of technical functionality, nor do they point to a lack of particular expertise. Instead, they represent a gap in organizational support for implementing RDA outputs.

It became clear during the workshop that digital humanists working in the European context felt particularly well-supported technically through organizations like DARIAH (which offers a PID service called EPIC, producing HANDLE identifiers). However, in other contexts, this middle organizational layer connecting data services, institutions, and the digital humanities community largely does not exist.<sup>4</sup> While certain participants had Digital Humanities Centers at their institutions supporting digital humanities

---

<sup>4</sup> Notably, this was a primary goal of Project Bamboo in the US context. However, after much planning, coordination, and deliberation, the project never moved into the implementation phase. Quinn Dombrowski, “What Ever Happened to Project Bamboo?,” *Literary and Linguistic Computing*, June 16, 2014, fqu026, doi:10.1093/lc/fqu026.

infrastructure and services, many other institutions offered little to no digital infrastructure support to the humanities. Thus, practitioners expressed concern over both having to pay for PID services and/or having to develop services for assigning PIDs.

Bearing in mind that there is uneven organizational support for implementing RDA outputs in the digital humanities domain, one recommendation to the RDA working groups would be to produce appendices to all outputs listing and describing supporting services and organizations in different geographic contexts. Another feasible suggestion would be to create a on the Web platform where organizations that have implemented outputs can post implementation notes, describing services they either used or developed.

For instance, as noted above, the Platform for Experimental Collaborative Ethnography project team built a tool for automatically implementing Archive Resources Keys in Drupal, a service that is now accessible to other research communities that would like to use ARK as a PID service. However, currently there is nowhere on the RDA website to advertise this.

## Macro

*Macro-level analysis focuses attention on the financial and legal structures needed to support the continued work of data sharing communities and organizations. Enabling the work described in each of the other levels will require that funding be made available to support infrastructure development and collaborative research. It also will require legal support in a landscape where what it means to own, manipulate, and distribute content is rapidly change.*

*Questions at this level include:*

- *How do we ensure that data sharing practices work within the parameters of existing laws – such as copyright and requirements for human subjects research?*
- *How do we ensure that appropriate financial structures are in place to support and sustain the development and implementation of data infrastructure?*

In the concluding discussion, a participant expressed concern about being able to continue deploying services she built beyond her currently funded project. This observation represents a challenge to deploying data sharing infrastructure in the humanities at many levels. Without guarantee that there are sustainable funding models for hosting data services in the digital humanities, there is little incentive for practitioners to invest the time in building them (a meso level issue).



Another participant suggested that maybe a community repository for services could address such an issue. However, as the initial participant noted:

Although that opens up a lot of other questions like who's responsible for maintaining and keeping that running? What level of guarantee do I have if I put myself in a shared repository? But it's not just about repositories of data right? It's about who's going to host these tools that I've written that implement these APIs. So it's also platforms - you know who's paying for the Amazon cloud instances that they're running? And who's actually keeping those services up and monitoring them?

One suggestion for addressing this particular macro-scale issue was to look at the business models that are forming to support the evolving role of scholarly societies – particularly as journals move to open access publishing. To support this form of publishing, societies are rethinking membership models, with an explicit aim to sustain journals over time. Similar models could likely be applied for the development and maintenance of shared service repositories, hosted and maintained by organizations that would aim to cross-institutional and possibly even disciplinary boundaries.

## Meta

*Meta-level analysis is concerned with discourse about data sharing and open science. At a time when many humanities communities do not see it as priority to invest in data infrastructure, we are challenged with developing language to communicate the value and imperative of data sharing and collaboration. Such a global culture shift will be necessary to gain support at the macro level (where funding is secured and legal support gained).*

*Questions at this level include:*

- How do we build discourse around the promise of open science – particularly at a time when global support for research itself is under attack in many settings?*
- How do we build discourse that encourages investment in the development and maintenance of research infrastructure?*

One critical point described during the workshop is that discourse in support of digital humanities research infrastructure has developed unevenly. A participant and digital humanities practitioner from Croatia noted that, at the national level, there is a depleted language for talking about the importance of digital humanities infrastructure. Yet, he also noted that support on the European level is much more robust, and thus collaborating with people outside of Croatian is often the only option for engaging in data sharing. Another participant went on to note that this is not just an issue at the national level. She noted that it was particularly challenging conveying to

administrators at her institution – a technical institution – that humanities deserve research infrastructure.

A suggestion to address this issue was to develop model documents with language that could be framed for governments, administrators, etc to convey the value of building digital infrastructure for the humanities. The RDA could vet such documents.

## Conclusion

Sharing RDA outputs with the digital humanities community in a conference workshop setting served a dual purpose for the RDA. In one sense, presenting outputs at such a workshop helped to disseminate RDA outputs to a community that has largely had little exposure to the RDA. It was particularly useful to have Bridget Almas available to introduce the outputs, workshop particular use cases, and answer questions posed by the group. Overall, sponsoring conference workshops directed at certain domains appears to be a promising model for enrolling more engagement with RDA and for kick-starting implementation of outputs in diverse projects.

In another sense, the workshop created a space for domain practitioners to present particular barriers to output adoption faced by their communities. Through an initial survey and discussion during the workshop, the workshop organizers were able to record several concerns expressed by workshop participants that can be brought back to the RDA community and help to orient how future outputs are composed and disseminated, as well as how their implementation gets supported. Since the challenge of building digital infrastructure for data sharing is exacerbated by the diversity of disciplinary communities involved, recording such concerns in many different communities will be essential for moving RDA and its outputs forward. I hope that the scale model presented here can serve as a template for recording such barriers in other domains, allowing RDA leadership to juxtapose issues faced by different communities at certain scales.