

# An overview of the 2023 NISO Plus Conference: *Global conversations/global connections*

Bonnie Lawlor\*

*Guest Editor, NFAIS Honorary Fellow, Radnor, PA, USA*

**Abstract.** This paper offers an overview of some of the highlights of the 2023 NISO Plus Annual Conference that was held virtually from February 14–February 16, 2023. This was the fourth such conference and the third to be held in a completely virtual format due to the global impact of COVID-19. These conferences have resulted from the merger of NISO and the National Federation of Abstracting and Information Services (NFAIS) in June 2019, replacing the NFAIS Annual Conferences and offering a new, more interactive format. As with prior years, there was no general topical theme (although the importance of metadata was a common thread throughout), but there were topics of interest for everyone working in the information ecosystem - from the practical subjects of standards, archiving, digital preservation, Open Science, and Open Access. With speakers and attendees from around the world and across multiple time zones and continents, it truly was a global conversation!

Keywords: NISO, metadata, open science, open data, open access, information accessibility, artificial intelligence, FAIR data, indigenous knowledge, digital preservation, archiving, knowledgebases, information sharing, information discovery, miles conrad lecture, scholarly communication, visual-meta, taxonomy, DOIs, persistent identifiers, african research, federated access, seamlessness, helix analytics, TCC Africa, practice research, PID alliance, Non-traditional research outputs, CARE principles, IEEE, COUNTER usage reports

## 1. Introduction

In February 2020 NISO held the first NISO Plus Annual Conference in Baltimore, MD, USA. It replaced what would have been the 62nd Annual NFAIS conference, but with the merger of NISO and NFAIS in June 2019 the conference was renamed NISO Plus and adopted a new format. The inaugural conference was labeled a “Grand Experiment” by Todd Carpenter, NISO Executive Director, in his opening remarks. When he closed the conference, all agreed that the experiment had been a success (myself included), but that lessons had been learned, and in 2021 the experiment would continue. It did, but due to the pandemic the experiment became more complicated as the 2021 conference was held for the first time in a totally virtual format and for me it was the best virtual meeting that I had attended up until that time.

Fast forward to 2023 and the fourth NISO Plus Annual Conference was also held in a completely virtual format in order to accommodate a truly global conversation - and NISO outdid themselves. The general theme from 2021/22 continued - “global conversations/global connections” and again speakers were

---

\*E-mail: [chescot@aol.com](mailto:chescot@aol.com).

recruited from around the world, many of whom we might never have had the opportunity to learn from at a conference due to the location of their home base and travel restrictions. Attendance for the conference remained strong (if slightly down over last year), with six hundred and four attendees versus roughly six hundred and thirty in 2022. (*Note: the first NISO Plus conference was a traditional pre-pandemic meeting with in-person attendance only and it attracted just over two hundred people.*)

NISO's new virtual approach is truly attracting a global audience. The 2023 geographical breakdown was in line with the 2022 conference reflecting NISO's ongoing commitment to increasing international engagement. Attendees came from thirty countries in 2023 vs. twenty-nine in 2022, and twenty-six percent of all attendees were from outside the U.S.A. compared to twenty-seven percent last year. A true tribute to NISO is that a post-conference survey indicated that most of the attendees who completed the survey had attended a prior NISO Plus conference and returned because of the relevance and usefulness of the program. Attendees were a representative sample of the information community - librarians, publishers, system vendors, product managers, technical staff, etc., from all market segments - government, academia, industry, both for-profit and non-profit. There were approximately forty-five sessions, most of which were recorded and are now freely-available for viewing [1].

As in prior years, Todd Carpenter, NISO Executive Director, noted in his welcoming remarks that it was important to lay out NISO's vision for the conference. He noted that many attendees might be new to this concept, and he wanted everyone to understand the conference goal, its format, and why NISO is building on the success of the past three years - they simply want to keep the momentum going. He emphasized that the attendees themselves are integral to making the event special because this meeting is not purely an educational event, it is meant to be a sharing event - a place where participants can interactively discuss and brainstorm ideas.

He went on to say that the goal of the conference since it was reshaped in 2020 is to generate ideas and capture as many of them as possible. NISO designed the conference so that everyone can benefit from the experience and knowledge that *all* participants bring to the topics that will be discussed over the three days of the meeting. The goal is to identify practical ideas that are solutions to real world problems. The problems may not be ones facing everyone today, but ones that are foreseen to be coming and for which we need to prepare. The ideas should produce results that are measurable and that can improve some aspect of information creation, curation, discovery, distribution, or preservation. In other words, the ideas need to have a positive impact - improve our work, our efficiency, and our results. He again said that he would like to have attendees look at the ideas that are generated over the next three days and ask how those ideas could make a difference in their own organization or community and how they themselves might want to be involved. He made it clear that NISO is delighted to have a lineup of brilliant speakers who have agreed to share their knowledge, but that the goal of the conference is not simply to take wisdom from what he called the "sages on the stage". He asked that everyone indulge him in his belief that *everyone* participating in this conference is brilliant and that he would like to hear from each and every one because the diverse reactions to the speakers and the ideas are what will make the event a success. NISO wants to foster global conversations and global connections about the issues that are facing everyone in the Information Community, and this cannot be accomplished by simply listening to the speakers.

Carpenter went on to say that the structure of the conference was designed to foster discussions and at least half of the time in each of the non-plenary sessions would be devoted to discussion. Each session was assigned a moderator and staff that will help encourage and record the conversations. And for the design to work all participants need to engage in the process. He added that if this NISO Plus conference is similar to its predecessors, lots of ideas will be generated, some of which will be great, some interesting, some will not take-off, some will sprout and perhaps a few will turn into giant ideas that have the potential to

transform the information landscape. He was also forthright in saying that NISO cannot make all of this happen. They simply lack the resources to manage dozens of projects. As in the past, they will settle on three or four ideas and perhaps the other ideas will find homes in other organizations who are interested in nurturing the idea and have the resources to do so.

In closing, Carpenter said that on a larger scale the NISO Plus conference is not about what happens over the next three days, although he hoped that everyone enjoys the experience. What is important about the conference is what happens on the days, weeks, and months that follow. It is what is done with the ideas that are generated and where they are taken. Whether the ideas are nurtured by NISO or by another organization does not matter - what matters is that the participants take something out of these three days and that everyone does something with the time that is spent together in the global conversation.

Certainly, most of the sessions to which I listened were interesting, had in-depth discussions, and a few did generate ideas. Be aware that for each time slot on the program there were multiple sessions running in parallel, and while I am fairly good at multi-tasking, even I cannot attend more than one session at a time. Also, I did not, after the conference, listen to every recorded discussion. In fact, not every talk was recorded - recording was done at the speakers' discretion. Therefore, this overview does not cover all the sessions, but it will provide a glimpse of the diversity of the topics that were covered and, hopefully, motivate you to attend next year's meeting which is being developed as I write this. That is my personal goal with this brief summary, because in my opinion, this conference is worthy of the time and attention of *all* members of the information community.

## 2. Opening keynote

The Opening Keynote was given by Dr. David Weinberger, an American author and technologist with a PhD in philosophy. Dr. Weinberger has been deeply affiliated with the Harvard Berkman Klein Center since the early 2000. He was co-director of the Harvard Library Innovation Lab from 2010 to 2014 and he also has been a journalism fellow at Harvard's Shorenstein Center and advisor to high tech companies. He currently edits an open book series for MIT Press about the effects of digital technology and is an independent, part time contributor to Google's moral imagination group, but made it quite clear that does not speak for Google in any regard.

Dr. Weinberger's talk was entitled "Unanticipated Metadata in the Age of the Internet and the age of Artificial Intelligence (AI)", and it was one of the more thought-provoking talks of those that I sat in on. He discussed the history of metadata, how traditional metadata has had to anticipate the uses to which it will be put - who will use it, how users will navigate it, how it will be encoded, and how much metadata is enough. He then asserted that the Internet has changed that approach because what is distinctive about the Internet is that its strategy is *unanticipation* and that this has had important effects on metadata.

He demonstrated his point by comparing the process of product design before and after the Internet became a force using Ford and Dropbox as the examples. Traditionally, products were designed by anticipating market needs. For example, in 1908, Henry Ford and a handful of engineers designed the Model T, anticipating market needs so well that fifteen million of them were sold over the course of nineteen years with only modest of changes [2].

Today's approach is different. Companies often choose to create minimal viable products (MVPs) [3] and let the market dictate features and use. For example, in 2008, rather than anticipate what people would want to do with its cloud backup technology, Dropbox launched a product with the most minimal feature set it could and then waited to see how people used it, what they wanted from it, what they were saying to

one another on the Internet, and so forth [4]. This validated the market desire for the product and helped Dropbox figure out what features to add to its subsequent releases.

According to Weinberger, unanticipation is the strategy behind the Internet itself which was designed to support as many applications and new protocols as the world cared to make, and he raised the question - why anticipate when you can learn along the way?

He then went on to talk about the three types of metadata. Metadata of the first kind is where the metadata, such as a label, is attached to the object. He pointed out that while the metadata is helpful, space limits the amount of information that can be provided. Indeed, how much information can be put on the spine of a book? This type of metadata is designed for use in a system where the objects themselves are physical and can only be put in one place at a time.

It reinforces a sort of rigidity of thinking about the nature of things.

The second type of metadata emerged when we realized that we could separate the metadata from the object itself and make a new physical object - for example, card catalogs for books. This type of metadata provides the capacity for more information than the object itself, and we can have multiple ways of organizing the data - by author, by title, by subject, etc. This was a break-through that enabled new possibilities for finding and putting together knowledge.

He said that the third type of data emerged in the age of the Internet in which everything is digital - the content is digital as is the metadata. And this new world enables the rise of new types of metadata. For one thing, there can be an unlimited amount of metadata because the internet, in practical terms, has infinite capacity. This opens the possibility of users adding their own metadata to digital objects and this changes the fundamental relationship of data and metadata.

He went on to say that in the current environment, Artificial Intelligence (AI) has come along to complicate things further as there are multiple ways om which we can use AI/Machine Learning - as a metadata tool, for example:

- AI can do at least some of the work of classifying items into established, structured metadata categories.
- AI can discover its own categories, and not just sort items into categories that already exist.
- Machine Learning enables a rich form of discovery.

He provided an example for the final bullet point. He prompted chatGPT [5] with: “Dante postulated an afterlife with three levels. What are ten artworks that also show something in three parts?” He said that it came back with an excellent list, with the tripartite nature of each work clearly explained. He noted that many of these new roles for metadata are possible because Machine Learning is able to deal with complexities that computers cannot explain. This enables them to take in more than we humans can of the ultimate black box: the world.

Weinberger said that now that this new technology enables us to wring benefit from patterns in data that are unintelligible to us, we are better able to acknowledge and appreciate the overwhelming and chaotic nature of our situation in the world. To gain advantage from such a universe - and perhaps even to survive it - we need to demand more of our metadata and in closing he provided a few examples of what needs to be done, e.g., metadata must become more contextually aware, both of the subject matter and of the needs and capabilities of the person (or system) using the metadata.

Dr. Weinberger has written an article based upon his presentation that appears elsewhere in this issue of *Information Services and Use*. I highly recommend that you read it.

### 3. Telling a story with metadata or always drink upstream from the herd

#### 3.1.

The first speaker in this session was Julie Zhu, Senior Manager of Discovery Services, at IEEE. She opened with a brief overview of IEEE (it is the world's largest technical membership association with more than four hundred thousand members in one hundred and sixty countries!).

She said that metadata is an essential part of publication and products. It is needed to display the key components of articles in publications such as author-related information, publication, information, article information, access information, funding information, and more. It is needed for discoverability, linking, access, and to track usage. Metadata flows through multiple systems in various pipelines from multiple systems - publishing systems and platforms, indexing systems, link resolving systems, authentication systems, library systems, user checking systems, etc. and forms a huge metadata ecosystem.

She pointed out that metadata problems can occur at any time during any stage of the process - during metadata creation, enrichment, transfer, and configuration. Problems in the metadata upstream will impact the functions downstream and fixing metadata downstream may not resolve problems upstream. She published a post in the *Scholarly Kitchen* in 2019 entitled "Building Pipes and Fixing Leaks: Demystifying and Decoding Scholarly Information, Discovery and Interchange" [6] which describes the complexity of this piping system and reminds us that content providers, discovery service providers, and libraries are all data plumbers. We need to continuously build data pipes and fix data leaks to ensure a better end user experience.

She went on to talk about a few of the many different types of metadata that exist today - especially author-provided metadata that is at the very beginning of the pipeline. Misspellings, special characters in titles and names, undefined abbreviations, etc. can all trigger problems as the information flows downstream resulting in challenges to information access and discovery, e.g., misspellings or a special character in the names may cause Google Scholar to drop the author names from article search results. She then went on to discuss some of the standards that exist to help metadata governance during the various stages of information creation and curation.

She said that metadata is everywhere in all kinds of systems, and internally at IEEE they need to work closely across the various teams and units. They have separate editorial and publication teams for different content types, such as journals, conference proceedings, standards, eBooks, and e-learning courses. When one team creates metadata correctly, it does not mean that the other teams also do it correctly. So, they monitor different teams and encourage exchange among them. Also, when metadata is created correctly, it does not necessarily mean that that data will be stored correctly in all the systems, in all the formats or the databases. They constantly monitor and troubleshoot even when they handle all the metadata currently internally. They also have to ensure that their external data partners currently pick up, store, and index their metadata.

In closing, she said that it is a never-ending process. We will always have to build better data pipes and fix the leaks.

Zhu has written an article based upon her presentation that appears elsewhere in this issue of *Information Services and Use*. She has written two articles for this issue. The title of her paper for this session is *Unlocking Potential: Harnessing the Power of Metadata for Discoverability and Accessibility*.

### 3.2. What if your metadata isn't properly represented in the stream?

The second and final speaker in this session was Jenny Evans, Research Environment and Scholarly Communications Lead, University of Westminster, UK. She is based at the University of Westminster in central London. The University has four campuses around the area and more than nineteen thousand students. They are very much a research-engaged University, with a primary focus on the arts, humanities, and social sciences, although they do have some scientific research.

She spoke about something of which I had never heard - practice research - research that covers diverse disciplines and has outputs beyond traditional text-based scholarly work. She said that the type of research to which Julie Zhu referred is traditional research for which the output is usually a text-based article or book. Practice research can have text-based output, but usually the outputs are not. It can represent performance or other live events and she provided the following definition:

“Practice Research is an umbrella term that describes all manners of research where practice is the significant method of research conveyed in a research output. This includes numerous discipline-specific formulations of practice research, which have distinct and unique balances of practice, research narrative and complementary methods within their projects.”

Research narrative: “In a practice research output, a research narrative may be conjoined with, or embodied in, practice. A research narrative articulates the research enquiry that emerges in practice.”

I must admit that the definition is somewhat confusing to me, but I thoroughly enjoyed her talk. She noted that the existing infrastructure often overlooks the nuances of practice research, hindering its discoverability and reuse and she went on to talk about an Arts and Humanities Research Council-funded project entitled “Practice Research Voices,” which aims to scope recommendations for enabling practice research across repositories, metadata standards, and community engagement. She discussed the key challenges facing the practice research infrastructure, including the complexity of representing iterative, multi-component outputs. She said that the project demonstrates that lessons learned from practice research disciplines can benefit research more broadly through inclusive and flexible systems.

Fortunately, Evans has written an article based upon her presentation that appears elsewhere in this issue of *Information Services and Use*. If you are interested in the subject it is worth a read – I cannot do it justice.

## 4. Making the business case for investing in metadata

This session focused on the relevance of metadata within the research lifecycle, from research organizations to funders to publishers as well as to researchers themselves, and why it is important to the overall research process. Its premise was that metadata provides context and provenance to raw data and methods, and is essential to both the discovery as well as the validation of scientific research. Metadata is also necessary for finding, using, and properly managing scientific data, and it plays a crucial role in increasing the quality of data within information systems.

### 4.1. Exploring the costs of not re-using metadata effectively

The first speaker was Josh Brown, Co-founder, Research and Strategy, at MoreBrains Cooperative [7], who inferred the value of metadata by exploring the cost of not using it effectively using as an example

some projects that his organization conducted analyzing metadata reuse in the Australian and UK research systems.

He said that they focused on persistent identifiers (PIDs) [8] because they are a keystone form of metadata and they are a source of a huge amount of information in and of themselves. The presence of a PID in a metadata record implies a relationship between the PID and that other thing. But also, there is a huge amount of metadata stored in registries about the things that are identified today.

He went on to say that there are multiple ways you can derive value from metadata, but today he chose to focus on metadata re-use, i.e., where structured metadata that is stored in registries is pulled into other information systems throughout the information ecosystem and flows through some PID registries between funders, researchers, institutions, publishers, and content platforms. He said that not re-using metadata actually comes at a price. Their research has shown that there is a huge amount of duplicated effort which could be eliminated by the more effective use of PIDs. They were able to assign a value to that in time, money, and to a large extent the impact on national economies given the scale of the research sector in the countries that they were looking at. They focused on time, and by extension, the cost associated with the re-keying of metadata into applications such as research management systems or grant reporting platforms.

He said that this is a well-identified and well-documented challenge and it has immediate resonance. The scale of it is pretty significant with estimates of the time that researchers have spent on re-keying have been ranging from 10% to 42% of their time. This is a pretty tangible thing that has been well-researched. But he said that there are other potential benefits that are no less important. And these he discussed later.

The first step his group took was to develop a method for valuing the impact of PIDs on the research information system in the UK with a cost-benefit analysis. They started by calculating the time spent on basic everyday metadata entry tasks and were able to do this based upon previous research. They found that the time taken to enter project or grant information is about ten minutes. This data would be simply a basic description of the grant or the project, not the full document.

He said that there are about thirty-six thousand grants issued annually in the UK, resulting in nearly a quarter of a million publications per year. And that number continues to grow from UK-affiliated institutions. They took into account the number of researchers by using higher education statistics and with the help of a forensic accountant at JISC they estimated the cost of re-keying metadata as well as the costs of a support program for the UK to generate PID adoption along with the cost of the systems required, arriving at savings of about 45.7 million Pounds over a five-year period. He said that this is significant, but not compelling, but it does mean that you do recoup the cost of developing a PID integration system.

Their next step was to do a similar cost-benefit analysis in Australia in order to go a little bit beyond what they were able to do in the UK. The scale of the Australian Research system is significant, about six thousand grants per year from Australian funders alone. The volume of publications per year is getting close to that of the UK and growing even faster. But the difference between the two studies is that they were actually able to find out how many times metadata was reused. In the UK study they assumed it was only re-used once because they wanted to be very conservative and did not have evidence to support a higher number.

They did a survey of repository managers and research managers at universities across Australia and found that data entry tasks were being conducted 3.0 and 3.25 times for grant information typically, and for publication information 3.1 times. This results in twenty-four million Australian dollars being wasted annually - a much more compelling number than they discovered in the UK.

They then chose to revisit their UK project with the extended methodology - no longer assuming that metadata was reused only once, and they found that nearly nineteen million pounds were being wasted annually in staff costs just by multiplying the number of entities by the number of reuses.

He then said that there are other benefits beyond the cost-savings from the elimination of re-keying that should be considered. The first is automation. And this is where simply a PID in a metadata record or an information system triggers an action, and this is much harder for his group to quantify with the evidence that was available to them in the projects he discussed. But he asked everyone to think about the possibilities. For example, Grant DOIs could be associated with raw IDs for institutions and funders, with ORCID IDs for investigators, with PIDs for projects funded, etc. If those things were linked consistently and reliably, the value of automation can go beyond time saved to include harder to quantify things such as having more complete information and more timely information processing. Another example he put forth is aggregation and analysis. There is a huge amount of data held in registries and at the institutional or the national scale. Aggregating this information about entities and the relationships between them can actually provide a range of strategically-crucial analytics and insights.

He said to think about the coverage and a completeness of PID registries continuing to grow and becoming more valuable as a source of increasingly authoritative information - information about grants funded and all the entities associated with the grants (researchers, funders, institutions, etc.). This would increase the likelihood of capturing data about outputs linked to those projects and improve strategic decision making. It would allow you to follow a project through time to see evidence of impact into the future, thereby providing evidence of return on investment on research and innovation expenditure. And it also enables projects and funding to be managed more efficiently and effectively.

In closing, he thanked the organizations that sponsored the projects that he discussed - JISC, the Australian Research Data Commons, and the Australian Access Federation.

#### *4.2. Metadata's role in knowledge management and measuring its value*

The second speaker in this session was Heather Kotula, Vice President, Marketing and Communication, Access Innovations, Inc., who spoke about the role of semantic metadata in Knowledge Management (KM). Specifically, she answered two questions: (1) how does metadata support knowledge management? and (2) how can the return on investment in metadata be expressed?

She began by saying that without really realizing it, we are all surrounded by semantic metadata - think taxonomy. She showed a kitchen drawer with an insert tray that allows you to "file" flatware - knives, forks, and spoon - each in its own space - and said that in reality this is an implementation of a very simple taxonomy. You know exactly where each type of utensil is to be placed. She then showed a "spork" (a hybrid spoon/fork) and said that in your flatware drawer taxonomy, you would have to expand your vocabulary or that insert tray to accommodate this utensil.

She went on to say that with semantic metadata, we label things or concepts with words, and in doing so, we put a handle on them for retrieval. As an example, your earliest memory probably coincides with the time in your life when you learn to speak, when you were able to attach a word or words to some event. You are able to retrieve that memory because it has words attached to it. Metadata's role is to serve as a label that facilitates the discovery and retrieval of information. Words help us to organize our thoughts. They are verbal symbols of our knowledge, and we can organize that knowledge from a random collection of thoughts by using various knowledge organization systems. These range from simple systems, such as a controlled vocabulary to more complex and comprehensive systems such as taxonomies, thesauri, and ontologies. All of these systems have a common purpose. They help us to label and organize our knowledge to make it more useful - to enable effective storage and retrieval of that knowledge information.

She went on to talk about taxonomies, defining a taxonomy as a controlled vocabulary for a subject area with its terms arranged in a hierarchy. The purpose of a taxonomy is to index or describe the subject

matter of a document or a collection of documents. It is the list of words that we use to label that content. A taxonomy is a central part of a knowledge management system, and it provides the most efficient way for users to access content by using the controlled vocabulary terms of a taxonomy to label concepts in a clear, consistent, and standardized way.

We can represent materials on those concepts and store and retrieve them efficiently. We remove them from a dead-end miscellaneous folder or from other forgotten files and make them available for dissemination and use as knowledge assets. So how is taxonomy used in knowledge management? A taxonomy reflects the concept in a document or collection of documents that are important to stakeholders. The taxonomy is used to describe the subject matter of documents - what they are about. It is the basis for indexing or categorizing the content. Indeed, using a well-designed taxonomy results in more efficient retrieval, leading to better productivity and cutting user frustration in searching. It saves time and money, not to mention the searcher's nerves. And a taxonomy directs the searcher to targeted knowledge.

She again showed the spork and said that the metadata for this utensil should be structured with multiple broader terms such as "fork" and "spoon", so a search for either forks or spoons will also retrieve sporks.

Her presentation then turned to measuring the value of semantic metadata. Traditional return on investment models are calculated using actual historic data about income and expenses (as exemplified in Josh Brown's presentation). Think of things such as infrastructure, hardware, software, furniture, and labor. These are things that leave a financial paper trail. We can get hard numbers for these and numbers tend to inspire confidence. However, we do not always have hard numbers to use in our calculations for metadata. She gave as an example calculating the return on investment for an exercise machine. You spend money to purchase it, but do you get money back when you use it? Does it generate revenue? Maybe if you are the owner of a gym. Possibly the return on investment or the value of that machine is in the improved performance of the user. The user gets stronger faster, perhaps even thinner, but they do not get money back. Yet many people think that the exercise machine is worth their outlay of money.

Another option for expressing value and another alternative to return on investment when building a business case for metadata might be the total economic impact methodology that was developed by Forrester. They suggest using this phrase exactly as it is with the blanks filled in:

"We will be doing \_\_\_\_\_ to make \_\_\_\_\_ better, as measured by \_\_\_\_\_, which is worth \_\_\_\_\_."

But she warned that it can be a challenge to fill in those blanks.

She filled in the blanks to give an example:

"We will be doing **a semantic metadata project** to make **search** better, as measured by **increased per article sales and/or decreased customer complaints**, which is worth **an estimated 80% increase in sales.**"

She went on to say that another way to look at the value of metadata is to measure the opportunity costs. The opportunity cost of a particular activity is the value or benefit given up by engaging that activity relative to engaging in an alternative activity. More effectively, it means if you choose one activity, for example, making an investment in the stock market, you are giving up the opportunity to do a different activity such as buying a car.

She said that opportunity costs are perhaps the most obvious way to measure the value of metadata. These include (1) time saved searching. We think we know that knowledge workers, time and success is of great value and up to 30% (or more) of knowledge workers time is spent searching and those searches are successful less than 50% of that time; (2) time to market; (3) reduction of duplicate effort; and (4) customer satisfaction.

She closed by saying that this was what she wanted to share today and thanked the audience for attending the session.

### 4.3. *The impact of metadata on research outputs*

The next speaker was Michelle Urberg, an Independent Consultant, Data Solve LLC, who is a frequent (and excellent!) speaker at the NISO Plus conferences. She opened by saying that the business case for metadata is elusive as Heather has just demonstrated, but believes that it can be made in part by looking at how metadata affects research outputs from the perspective of a number of participants in the scholarly communications ecosystem as follows:

- End Users: Does the metadata allow them to find information easily? Do the search results pertain to their needs?
- Researchers: Will metadata make their work publishable? Can they measure impact?
- Funder: What data does this study use? Has this data been used by others I want to fund?
- Content Provider: Will this metadata facilitate sales and keep us in business?

She went on to talk about a study sponsored by Crossref that she is working on with Lettie Conrad, Senior Associate, Product Research and Development at NISO. They are looking at the impact of book metadata on end user discovery in Google Scholar. The outcomes to date are that:

- DOIs matter at the title level (book chapter DOIs are less useful).
- Other pieces of metadata matter: titles (paired with subtitles), author/editor surnames and/or field of study.
- Metadata protocols used by Google Scholar are not fully-integrated into our industry's established scholarly information standards bodies.

She also talked about what an organization should consider with metadata and research outputs such as data sets. First, what are the key outputs in your organization and how do they relate to a version of record, be it a book, article or other book chapter, other data set, etc.? Do you need metadata for data sets, articles, books, and videos all at the same time? The reality of creating metadata for different formats and keeping that data linked requires using many different types of relational linkages in the metadata schema and actually in the metadata records. Do your schemas and IDs accommodate this?

Second, where do you find friction in your system that slows the flow of metadata discovery? Is it in creating item level metadata? Is it working with your vendors? Is it getting good information back from the content provider about usage?

Third, what metadata are you providing across channels? Is it sensitive to the needs of each channel while still being robust? What they found with the Google study is that the standards that currently work for traditional means of dissemination do not work with Google Scholar. And Google Scholar is not integrated well with many of the traditional pathways for disseminating information.

Fourth, what information is missing in your pipeline to make better business decisions about investing in metadata? What data do you need to make business decisions? Is it feedback from your clients and/or industry support or other guidance that does not currently exist?

Fifth, what data needs to flow back into your organization? What analytics are actually missing from the data that you currently receive?

In closing, she summarized what she sees as the truths about metadata: In order for a positive impact to be felt with the discovery of research, it is necessary to feed large amounts of good data into the ecosystem and keep it flowing. The experience of good metadata is frictionless and it provides a path of easy discovery to end users. When metadata is doing its job (1) you will not see what it does for end users, researchers,

and funders; and (2) the analytics are robust and usage can be accurately measured. For all those reasons, the ROI of metadata is very important to quantify and figure out, but also very elusive to pin down.

Urberg has written an article based upon her presentation that appears elsewhere in this issue of *Information Services and Use*. As with her submissions in past years, it is very informative and provides more detail than my summary.

#### 4.4. *Building a business case for investing in metadata*

The final speaker in this session was Julia McDonnell, Publishing Director, Oxford University Press. She opened by saying that she would talk about how publishers can play a role in the investment in metadata and provide the publisher perspective on building a business case. She said that a short and glib answer to the question of how can publishers contribute to and benefit from the use of metadata in publishing is how can we not? She thinks that all the prior speakers have made it clear that there is so much at stake when it comes to getting metadata right, and it is a reality that metadata are central to modern publishing. They play a crucial role in providing that contextualization for digital content in their many varied forms. And as we have increasingly disparate digital objects, it is essential that as much as possible, we use consistent and persistent metadata so that we can ensure each piece of content, no matter what it is, no matter how granular it is, is able to retain its context for entities who encounter that content, whether human or computer.

She went on to say she believes that the importance of metadata is only going to increase as we add more varied research outputs and as we see the complexity of the research environment grow (echoes of David Weinberger!). In many ways, she would argue that the production, distribution and maintenance of high quality, well-structured metadata, is one of the most essential roles that publishers have to play in the modern publishing environment.

She added that this is something that is all too easy to overlook, and as previous presenters have said, the better publishers do metadata, the more invisible it is. So, while on the one hand, the business case for metadata largely makes itself - for without metadata, the integrity of the research record cannot be supported, nor can access and discovery be effectively driven. This is an area where the role that publishers play is all too easy to undervalue. And she suspects that a surprising number of people throughout the environment of research and publishing perhaps still view metadata as a somewhat kind of distant prospect from the core publishing process.

She would argue that the business case from a publisher perspective is primarily twofold. One, the role of metadata is to enhance the research experience for authors and readers. And, and another role it can play is navigating the increased complexity when it comes to compliance. As Josh Brown has already demonstrated - there is a huge amount at stake here. So much time is being spent and wasted on re-keying of data. She believes that a critical piece here is looking at how metadata can be collected as far upstream as possible. How can this be done once in a consistent way that uses recognized industry standards, so that metadata can then float all the way through the publication process and downstream to post-publication. And as part of that, how do we effectively use standards to ensure that metadata is being passed accurately from system to system throughout the entire process? How metadata is used has moved on a long way from the days of searching for something through library catalogs. Publishers are working in a much more complex environment in terms of supporting discovery, access, and impact - an environment that involves collaboration with diverse organizations (Libraries, Google Scholar, Abstracting and Indexing services, Discovery services, etc.) in order to make content easily accessible. And publishers must continue to invest in all of those relationships to enhance the research experience for authors and readers.

She said that the second business case is related to compliance as authors are expected to navigate a lot of different requirements. With multi-authored papers there will be requirements from different funding bodies, data management requirements, perhaps publication fees, etc. She said that as an industry we need to challenge ourselves on how we can use metadata to support this compliance - how can we use it to help authors navigate that complexity and do so in a way that is easy and intuitive for researchers, funding agencies, and the institutions who are supporting this content. We need to work together on this.

In closing, she emphasized the need for collaborative efforts - working together, we need to ensure that the metadata that we are producing is of high quality, well-structured, and compliant with FAIR [9] principles. We need to ensure the portability of metadata throughout the research and publication ecosystem.

## 5. Minding the gaps: Bibliometric challenges at the margins of the academy

The speakers in this session were Shenmeng Xu, Librarian for Scholarly Communication, Vanderbilt University, Jean and Alexander Heard Libraries, Clifford B. Anderson, Director of Digital Research, Center of Theological Inquiry and Chief Digital Strategist, Vanderbilt University, and Charlotte Lew, Coordinator of Digital Projects and Collections, Vanderbilt University.

This session addressed challenges in tracking and measuring the research output and impact of minor academic disciplines in the humanities. Scholarly fields such as theology and religious studies exist at the margins of the contemporary academy, and prevailing research information management tools do not accurately capture the range and reach of their academic contributions. They drew on a multi-year investigation at Vanderbilt University as well as a practitioner's perspective at the Center of Theological Inquiry to document the extent to which existing bibliometric tools fail to capture the full scholarly output of scholars in these subdisciplines. They assessed the current state of research information management in minor academic fields, then suggested traditional and non-traditional bibliometric solutions to measure them more comprehensively and fairly. They explored how open-source initiatives such as Wikidata/WikCite [10] and OpenAlex [11] could help to mend the gaps between central and peripheral academic fields. They also explored how emerging techniques such as network analysis help to demonstrate the reach and interdisciplinary impact of scholarship in these disciplines.

Xu, Anderson, and Lew have written a joint article based upon their presentation that appears elsewhere in this issue of *Information Services and Use*.

## 6. Lightning talks

This was a series of very brief talks (five to six minutes) on a variety of topics. Fortunately, all but two of the speakers submitted a manuscript with a bit more detail and I was able to fill in the blanks for one of them.

### 6.1. A tiny library with a million volumes

This presentation had two speakers from the Equinox Open Library Initiative, Andrea Buntz Neiman, Project Manager for software development, and Roger Hamby, Data and Project Analyst and they talked about the future of resource sharing for libraries. Roger opened with a brief description of Equinox

Open Library Initiative [12]. It is a non-profit organization that works with libraries of all kinds - public, academic, government, special - to provide open-source solutions through development, consulting, etc. He said that as a non-profit organization, their guiding principle is to provide a transparent, open software development process, and they release all code developed to publicly-available repositories. He said that one of the things that makes them different is the equinox promise [13]. He said that he and Andrea would talk about an open-source product called Fulfillment. The days of FUD (fear, uncertainty, and doubt) are starting to fade away. Fear, uncertainty, and doubt he said, has been spread by proprietary vendors who have moved to what he termed “open washing” - taking products that are really proprietary and doing the minimum to technically call them open source. Why? Because the market for open-source services is growing dramatically. But Equinox puts their code where their mouth is and you can find Fulfillment on GitHub freely-available.

He put forth a rhetorical question - how can open-source help libraries? - and went on to say that the transparency helps security and privacy, which are so important these days. Also, the lack of licensing costs is a big deal and it allows direct community participation by all levels of libraries. He said that people instinctively understand that small libraries really benefit from the larger collections of large-scale resource sharing. However, large libraries have large populations which often have more specialized interests and really take advantage of the collections at small libraries that may have obscure materials. So large libraries often end up as net borrowers. He told the audience that if they want to read more about that concept, he has published about it [14].

Andrea followed with a description of Fulfillment which is Equinox’s software for open-source resource sharing. She said that it provides all full request management, Union catalog record loading, and matching streamlined interfaces. It provides a lot of loading scripts for SeamlessAccess (a topic mentioned by several speakers throughout the conference), and is standards compliant. They plan to release a new version of Fulfillment this month.

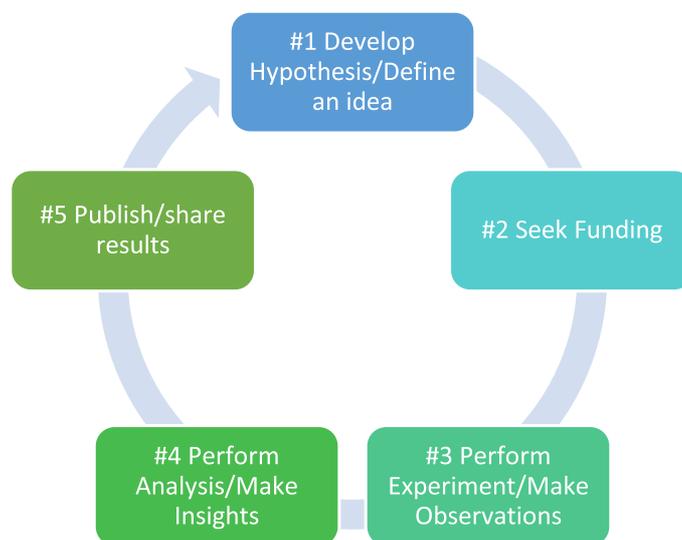
Hamby and Buntz Neiman have written an article based upon their presentation that appears elsewhere in this issue of *Information Services and Use*.

## 6.2. The use of blockchain technology along the scientific research workflow

I myself gave the next talk. I spoke about a project that I led that was sponsored by the International Union of Pure and Applied Chemistry (IUPAC) to identify how blockchain technology is being used along the scientific research workflow. Time restrictions did not permit me to go into the level of detail covered by the full paper - this was just a teaser to motivate conference attendees to read the article when it is published at the end of 2023 or early 2024.

So first - what is blockchain technology? “Blockchains are a type of immutable distributed digital ledger systems (i.e., systems without a central repository) and usually without a central authority. They enable a community of users to record transactions in a ledger public to that community such that no transaction can be fraudulently changed once published [15]”. Blockchain technology is *NOT* Bitcoin - it is the engine “under the hood” of Bitcoin - an engine that can be used for diverse purposes of which cryptocurrency is just one and which predates Bitcoin by almost twenty years.

My project team defined the scientific research workflow as the five-step process you see below and we held about a dozen in-depth interviews with major global players across diverse disciplines who are successfully using blockchain technology for a variety of purposes in research.



What the three-year study found is that the technology is indeed being used in almost all the steps in the scientific research workflow - from hypothesis development through to publication - by commercial organizations as well as by non-profits and across all market sectors, even governments. For example, the U.S. Department of Health and Human Services (HHS) uses blockchain technology in a pilot program, the Grant-recipient Digital Dossier (GDD), to manage their grant program more efficiently. As of July 2021, GDD had reduced the time required to complete grant assessment tasks from four-plus-hours to a fifteen-minute process. In my presentation I was able to very briefly summarize the findings, discuss the pros/cons of the technology, and provide a glimpse of how the technology is impacting the future of scientific research (remember: I had five to six minutes to cover the content of a sixty-two page paper!).

More details can be found in the manuscript that I submitted to this journal and it appears elsewhere in this issue. Feel free to contact me if you have any questions ([chescot@aol.com](mailto:chescot@aol.com)).

### 6.3. *Moving from piecemeal to systematic: Reprioritizing how academic libraries approach research information management*

This presentation had two speakers, both of them from Indiana University Purdue University Indianapolis (IUPUI) - Olivia MacIsaac, Research Information Management Librarian, and Ted Polley, Director of the Center for Digital Scholarship and Digital Publishing Librarian.

Ted began by saying that libraries are increasingly engaging in research information management work (RIM). At IUPUI, RIM-related work has been historically distributed across different librarians and library services, including data services, research metrics, consultations, and open access policy implementation. They are attempting to take a more systematic approach because in 2024 their campus will change. They will no longer be IUPUI university, but we will become IU Indianapolis, and Purdue University will leave the campus to form its own campus. So, the departure of Purdue programs, researchers, and funding will undoubtedly impact the research output of our campus. Understanding how this change will impact their research footprint will help them to better understand and assess their research output and will be useful in strategic planning.

Olivia said that with research, information management, and open infrastructure, they already see tools, systems, and services that can be utilized. But these services and systems are not being leveraged by institutions or researchers in an equal way. So, to look at RIM, they do turn to different pieces to understand how they can fit together a landscape that will work with people such as librarians, researchers, research officers, and information technology staff. Sometimes they work independently from each other, which hinders open research and hinders a more complete research story being told. So, this begs the question of how academic libraries can engage with other people to support a more rapid sharing of knowledge about research outputs through the implementation of these different systems and services in a scalable way.

Ted said that they have been focusing on the people part of RIM work, specifically in building partnerships. They have partnered with the Office of the Vice Chancellor for research and academic affairs on a shared ORCID membership through Lyris [16]. Sharing this expense has helped them forge new connections with other units at the University. In the future, they plan to continue to look for these kinds of partnerships with organizations or offices that have a shared, overlapping mission with the library. And as their campus undergoes its change, the units that support faculty research will likely be centralized, at least to some extent. Moving forward, they will explore new opportunities for collaboration across the University. At the moment, they are just starting out the process of developing a systematic approach to research information work, and promoting the researchers control over their own data to get a better picture of the research activities that matter the most to are changing.

MacIsaac and Polley have written an article based upon their presentation that appears elsewhere in this issue of *Information Services and Use*.

#### 6.4. Federated access – data insight around the world

The next speaker was John Bentley, Commercial Director, OpenAthens [17]. He said that OpenAthens exists to connect people to information wherever they are in the world and whatever time of day it is. OpenAthens takes them on that journey to the insight that they need at that moment in time. And when they pass someone over to digital platforms or applications, they call that a transfer. OpenAthens sees around 250 million transfers to publishers each year, and although they do not have customers in every country, they do see transfers from users in nearly every country in the world. This data shows research and learning is no longer tied to specific locations.

He went on to say that single sign on and federated access management is not the only way of logging into digital platforms. Other methods include a username and password issued by the platform itself. IP recognition is used to read a user's IP address and this enables access to the platform. This remains, by far, the most common method used around the world. And this is often based on location. Federated access allows the users to access their resources wherever they are, and it takes the users to their institutional log in systems and allows them to authenticate using their familiar institutional credentials. That then carries a message confirming the user's identity back to the publisher, who is then able to authorize that user into that platform.

He said that authentication is done by the institution. Authorization is done by the platform. This aspect of Federated access allows privacy preservation. No personal data is required to be passed back to secure access. And he can see that Federated access is growing. It is a slow process because both publishers and libraries need to move together. But there is a recognition that single sign on does provide a secure access experience when everything aligns.

The institutions and publishers joining the OpenAthens Federation is growing and alongside that, the number of transfers that are made each month is growing as well. And as the transfers are carried out, they

are able to see trends and patterns regarding the technology used. For example, the aggregation of the data enables them to see which operating systems are used in different regions of the world. Looking at these operating systems, they see that Microsoft is still very dominant. If they look at the browsers around the world, they see the importance of Google Chrome if it is used effectively. This kind of insight allows them to inform and improve the decisions that they make to support their users. He said that he hopes people are aware that the browsers supported by the underlying browser engines are making changes designed to improve user privacy and prevent unwanted tracking. This will have an impact on everybody who uses digital platforms. But looking at the data, OpenAthens can tailor their efforts to ensure that the messages from their community land in the right place and where those messages will have the greatest impact.

He went on to say that there is further detail about devices. For example, they can drill down and see those users who are using desktops versus those users who are using mobiles or tablets. And they can see patterns around the world where certain regions are more reliant on that mobile experience. He added that it is important that learners or researchers, wherever they are in the world, can get access to the information that they need. Yes, this means equity of data, but OpenAthens also works to improve the user experience. For example, when recently developing a new service they worked hard during the development cycle to make sure that the experience is as good for a mobile user as it is for someone accessing research from their desktop.

He closed by thanking everyone for their attention and inviting them to reach out to him if they have any questions.

### 6.5. *Visual-meta*

The final speaker was Frode Alexander Hegland, Director, the Augmented Text Company. He spoke about Visual-Meta at the 2022 NISO Plus conference, but gave much more information at that time. At that meeting he was supported by Vint Cerf, Vice President and Chief Internet Evangelist for Google (yes, *that* Vint Cerf!) who is also an evangelist about Visual-Meta. That was one of my favorite talks of that conference, if not my favorite, so I am using my summary from last year's conference to give you a snapshot of Visual-Meta.

This is a new service from Hegland's company (he is also the product designer). The product addresses the problem that documents, particularly published documents, such as academic PDFs, lack access to digital features beyond the basic web link. The goal with Visual-Meta is to provide those documents with metadata to enable rich interaction, flexible views, and easy citing in a robust way. The approach is to write information at the back of a document. He said that it sounds simple and ridiculous, and it is.

He went on to say that in a normal paper book, one of the first few pages has information about the publisher, the title, etc. and that is metadata. It basically is the metadata that you need to cite the document. While PDFs currently could have metadata, it is too complicated to do it. What his company has done is take the metadata from the front of the book and move it to the back of the document. He put on screen the proceedings of the Association of Computing Machinery (ACM) Hypertext Conference 2021. At the end of each document, there is Visual-Meta. The formatting is inspired by BibTeX [18]. He said that not everyone knows what BibTeX is, but it is an academic format, part of LaTeX [19]. He then highlighted the BibTeX piece and pointed out (because the screen was cluttered with information) the term "author" equals, in curly brackets, the name of the author. The article title equals, in curly brackets, the title, etc. His approach is all based upon wrappers. There are start and end tags for Visual-Meta. Within those tags there is a header which basically says what version of the software is being used. Then there are the self-citation bits, what his group calls the actual BibTeX, because that is what someone would use to cite

the document. Importantly, there is an introduction in plain text saying what the information entails. His company has grand goals of a document using Visual-Meta being readable for hundreds of years in the future.

Last year Vint Cerf talked about the implications of what Hegland has been able to do - at least as of February 2022. His first observation was that by making Visual-Meta simply text at the end of the PDF, he has preserved the document's utility. Over long periods of time, in theory, it could be printed physically. It could be scanned and character - recognized, etc., so users are not trapped into a specific computing representation of the material because it is in this fungible text form. The second thing is related to URL references. He said that as we all know, if a domain name is no longer registered, then a URL that contains that domain name may not resolve, in which case, the reference is not useful. Therefore, the reference information that is incorporated into Visual-Meta is a much richer and probably a more reliable and resilient form of reference. The third thing Cerf observed is that Hegland has designed this to be extensible which is extremely important because it anticipates that there will be other document types such as programs or a Virtual Reality space that will require references. The extensibility of this design is also extremely important and its resilience as well.

In closing at this year's conference Hegland said that Visual-Meta is open and freely-available [20].

## **7. Collaborating to implement seamlessaccess: What publishers, libraries, and service providers can do to ensure better end user access to scholarly content**

In the past two years, more and more publishers have implemented SeamlessAccess [21], resulting in a better user experience and increased usage. However, to ensure that more users benefit from federated authentication and seamless access, more collaborations among publishers, libraries and SeamlessAccess are needed. This panel included: (1) an update from SeamlessAccess on their roadmap and development; (2) a publisher perspective on comparing publisher implementations of SeamlessAccess, reasons behind the variations and flavors, the challenges of adding IDPs to platforms, and the need for better library-publisher communication and collaboration; and (3) a library perspective on the challenges to implement federated access and benefits for the end-users in their research experience that come with federated access and how SeamlessAccess makes that even better.

The Library perspective was given by John W. Felts, Jr., Head of Information Technology and Collections, University Libraries, Coastal Carolina University. He said that in the past two years an ever-increasing number of publishers have implemented SeamlessAccess resulting in a better user experience and increased usage. However, to ensure that more users benefit from federated authentication and seamless access more collaborations among publishers, libraries, and SeamlessAccess are needed. He gave the library perspective on the challenges to implementing federated access, the benefits that federated access brings to libraries and to their end-users in their research experience, and a brief walk-through of the SeamlessAccess experience that demonstrates how it enhances federated access.

Felts has written an article based upon his presentation that appears elsewhere in this issue of *Information Services and Use*.

The publisher perspective was given by Julie Zhu, Senior Manager of Discovery Services, at IEEE. She provided insights into a publisher's perspective on implementing and evaluating SeamlessAccess, a collaborative initiative that facilitates seamless access to scholarly resources subscribed by libraries. She outlined the motivations behind adopting SeamlessAccess, including the surge in federated authentication usage during the pandemic, its adoption by other publishers, and considerations regarding security and

privacy. She discussed IEEE's decision to employ a hybrid approach, combining the advantages of SeamlessAccess with its own solutions to maximize flexibility. She also highlighted pilot projects focused on incorporating Identity Providers (IDPs) and EZproxy login URLs into the discovery process, along with strategies employed to monitor and assess the implementation of SeamlessAccess. By utilizing data, she demonstrated the overall success of the SeamlessAccess implementation project, while acknowledging potential obstacles that publishers and libraries may encounter during integration. Furthermore, she emphasized the importance of increased collaboration between publishers and libraries to fully realize the vision of federated authentication as a genuinely seamless experience for researchers.

Zhu has written an article based upon her presentation that appears elsewhere in this issue of *Information Services and Use*. She has written two articles for this issue. The title of her paper for this session is *Implementing and Assessing SeamlessAccess: A Publisher's Experience*.

## 8. Creating digital collections with and for indigenous communities

NISO Plus conferences have been addressing the issue of indigenous knowledge since 2021 when Margaret Sraku-Lartey, PhD., Principal Librarian, CSIR-Forestry Research Institute of Ghana, gave a talk entitled "Connecting the World through Local Indigenous Knowledge" [22] that totally fascinated me. In 2022, Joy Owango, Executive Director, Training Centre in Communication (TCC Africa) at the University of Nairobi and a Board Member for AfricArXiv [23], spoke on indigenous knowledge, standards, and knowledge management, and she returned this year. The session was one of the first on the second day of the conference.

Indigenous knowledge is also referred to as traditional knowledge or traditional ecological knowledge. It is the systematic body of knowledge acquired by local communities through the accumulation of experiences, insights, and understanding of the environment in a given culture. It can be in the form of oral and written observations, knowledge, practices, and beliefs. Indigenous knowledge infrastructure has different elements and in a very simplified way it is represented by FAIR [24] as the technical, policy standards and practices - these elements of the infrastructure. And equally important as represented by CARE [25], is the value driven and sociocultural elements such as trust and ethics.

The title of Joy's talk this year was *Increasing African Indigenous Knowledge and Research Output through Digitization*. She opened by saying that the Training Centre in Communication known as TCC Africa is an award-winning trust established in 2006, and it is housed at the University of Nairobi. The Centre provides capacity support in improving researchers' output and visibility through training in scholarly and science communication. The Centre's mission is to contribute to the increase in profile, locally and internationally, of African science and how it can impact the lives of Africans by improving skills in technical communication in all forms, in academia and other relevant forums in Africa.

Most of their projects are done with partners - they are collaborative, supportive, and goal-oriented, and work closely with their partners to ensure that they are able to achieve their various goals and objectives. She emphasized that they train, support, and empower researchers, government, and institutes. They have trained fifteen thousand six-hundred and forty-five researchers and more than eighty institutes, have supported researchers in more than forty countries, have more than nine hundred mentors, and have won three awards.

She went on to say that they are not taking over what universities do, rather they are being complementary to the services that universities provide, especially when it comes to supporting researchers. Now when it comes to the partnerships, their institutional capacity is based on supporting the research

life cycle of the higher education sector. So that means working with grant councils, research councils, universities, and libraries.

As they were going through their activities, one of the questions that came up very clearly, especially when it comes to Open Science, was how do you support Open Science for industry? This led TCC Africa to create a startup with partners, called Helix Analytics, whose objective is to promote data for impact (D4 Impact) by leveraging Open Science and modern data infrastructures to enhance public insights. This is an entrepreneurial commercial entity with a cyclic economy. The Centre is working with data scientists who also happen to be students. And the revenue that is being generated not only supports these data scientists, but most importantly it supports the activities that her team is launching for their projects out of Helix Analytics, which will be supporting Open Science for industry.

The industry sectors that will benefit from this project are research institutions, the health care industry, agriculture, logistics and transport, manufacturing, and the higher education sector. The healthcare, agriculture, logistics and transport, and manufacturing will fall under the business optimization research division, and the higher education and research institutions will fall under the higher education division, which leads to our latest project coming out of this, which is actually going to connect literally governments, industry, and also higher education.

The Centre's flagship project out of Helix Analytics is the Africa PID alliance. This alliance is a community of Persistent Identifier (PID) enthusiasts in and from Africa, aiming to lead and realize a FAIR [26] sharing of access and data through the use of PIDs in innovation, research, and technology within the cultural, scientific, and cross-industry ecosystems. And what is exciting them about this project is that they already have commitments from continental partners for this project: (1) the Africa Academy of sciences (the umbrella body of all National Academies in the continent); (2) the Association of African universities, (the umbrella body of all African universities); (3) the African library and information associations institutions (the umbrella body of all libraries, national, and government and private within Africa); and (4) the Kenya education research network, which is a national research network that will be providing the tech and cyber support for the Alliance's activities). And the community engagement drivers will be TCC Africa and also the Helix Analytics team.

She said that the African PID Alliance workflow will focus on knowledge collection and that they are looking specifically at three areas: patent information, indigenous knowledge, and African research outputs. They are looking at the curation of this information as well as the registration and consolidation of this information and data accessibility.

She said that they are excited about this project, but at the same time they are scared because from a Global South perspective, that is the region that has suffered when it comes to the loss of indigenous knowledge because they have not quite figured out how to protect it and how far back they are able to protect it. Also, when it comes to patent information that exists in databases, it is assigned to the research output. So, you can go do a search and find the funder, as well as the research output that was produced and the innovations, hence the patents. But then when you critically look at some of the patent agencies, not all of those patents come out of research and as a result, you do not get the full picture of the patent and innovation landscape coming out of a region.

She went on to say that Africa is now being perceived as the Silicon Savannah because of the large-scale startups that are emerging and the innovation landscape that is building up, but they still do not have the true picture of the innovation landscape because it is not based on the entire patent landscape that is filed within the patent agencies. Her group is trying to remedy that. And also most importantly, they have noticed that not all the patent agencies have digitized their information. So, they are trying to see how they can support that entire process to make this information accessible. They want to ensure that everyone

can have a true picture of what is happening in Africa (this issue was re-iterated by Caleb Kibet later in the program).

She said that the Centre does podcasts and are now on TikTok because researchers are getting younger and younger. In fact, they are using various social media platforms to make sure that they are reaching young researchers. Their goal is to ensure that African research output becomes much more visible and that everyone sees the true picture of what is coming out of the continent in terms of not only research, but also patent information. Most importantly, they want to protect and preserve their indigenous knowledge.

The other speakers in this session were Cindy Hohl from the Santee Sioux Nation, Director of policy analysis and operational support at the Kansas City public library, and Erica Valenti, Executive Vice President, North America, Emerald Group Publishing. I only listened to parts of their talks, but they were both great. If you are interested in the preservation of indigenous knowledge you should listen to the recording of this session.

## **9. Data and software citations: What you don't know CAN hurt you**

The premise of this session is the following. When we read a published scholarly article we rarely, if ever, ask to see the machine-actionable version of the text. And yet this hidden version is used to enable much of the downstream services such as automated attribution and credit. When it comes to data and software citations in the reference section, recently the probability of an accurate machine-readable version was very low. For some journals, even zero!! Why? The citation looks just fine in both the online version and the downloadable PDF - what could possibly have gone wrong? Well, there is a plethora of challenges to uncover. First, data and software citations require different validation steps during the production process. Because of this, the machine-readable text is often not analyzed correctly, and some text might be altered such that the citation is no longer actionable. Further, Crossref requirements are also different for these types of citations causing those citations sent improperly to land on the cutting room floor. This session detailed the differences in the production process and provided specific guidance to make the necessary corrections.

This session had two speakers, Shelley Stall, who is the Vice President of Open Science Leadership at the American Geophysical Union (AGU) and Patricia Feeney, Head of Metadata at Crossref. These two participated in an almost identical session entitled “Linked Data and the Future of Information Sharing” [27], that was held at the 2021 NISO Plus at a time when Crossref was updating their guidelines.

Shelley gave the first presentation. She said that she was sure that the topic of this session is very provocative for the conference attendees, and it does demonstrate that there were a number of clarifying moments when publishers came together to figure out how to ensure that data and software citations make it all the way through the publishing process. She said that we consume data citations and software citations in two different ways - one is consumed by humans by using their eyes and one is consumed by computers using the bits that need to be set up correctly. Humans have access to references and style guides and there is a lot of understanding within the research community on exactly how to navigate those reference materials. And we even have persistent identifiers that humans can use to clip and search. But it is the machine-readable version of the citation that makes it much easier to connect, link, find, and access citations.

She said that there are three problems that she wants the audience to consider:

- (1) Data and software citations have different criteria for quality checks than the traditional type of journal article references. These quality checks have changed and improved over the past two years since Patricia and Shelley presented on this topic in 2021. And most publishers are unaware of the changes.
- (2) Journal policies are generally poor when encouraging/requiring data and software citations from their authors. This results in ethical issues when proper credit and attribution is not given.
- (3) Even when journal policies support data and software citations there is little guidance to publishers for how to ensure that the downstream services are activated for automated credit and attribution. This results in frustration and assumptions, in finger-pointing, and a decrease in thrust in the publishing community.

She stressed that both human- and machine-readable citations are essential and that it is the machine-readable versions that are essential to support automated credit and attribution.

She went on to say that as a publisher you must look at every step in the process: Policy; Publishing staff, editors, and reviewers; the peer review process; copy-editing of citations; production mark-up of citations; and content hosting/publishing. She highly recommended that publishers use the research data policy framework to develop and/or revise their journal policies and suggested the one provided by the Research Data Alliance (RDA)[28]. It not only describes what is needed, but also why it is needed. With regard to publishing staff, editors, and reviewers she said that you must clarify what is required versus what is encouraged since not all researchers have access to the infrastructure that is needed to support leading practices. Give them examples and FAQs. Give them training. She said that anyone can use the checklist [29] that AGU has developed for this purpose. With regard to the peer review process there should be a checklist of the elements that are to be reviewed and reviewers should have the option to examine the data and/or the software. She emphasized that the peer reviewers *must* be able to validate that the data and software supports the science itself and the visualizations, and that the data and software citations are accurate, are deposited by the author, and are made available by the repository.

When she moved on to copy-editing of citations she reinforced that data and software citations are different from each other and from journal citations. They have to be handled differently downstream. They need to be correctly marked so that they can be correctly handled in production and publishing. The citation needs to be connected/coded to references within the text and availability statements. She said that there are three clues to correctly identifying the citation type [30]:

- (1) The dataset or software is identified in the availability statement prompting you to look for a citation.
- (2) The “type” of DOI is “data” or “software.”
- (3) The citations have bracketed descriptions to indicate “dataset,” “software,” “collection,” or “computational notebook.”

With regard to production mark-up of citations this entails validation and mark-ups and ensuring that the mark-ups are correct. This is where data in-text citations and software in-text citations get marked correctly. To find these she said to look in the following:

- (1) Methods Section: Check that in-text citations and/or text in the Availability Statement link correctly to citations.
- (2) Availability Statement: This text should include availability statements for all datasets and software that support the research. Any statement that includes a DOI should have an in-text citation linking to the reference section.

- (3) Citations: Journals should review and update the markup requirements for dataset citation (NISO, 2020) and software citation (NISO, 2021). The persistent identifier or URL should be an active link in the paper to the reference. Dataset and software research outputs should use the same journal citation style and treatment of the persistent identifier with slight adjustments to include software version information and bracketed descriptions.

The next and final step was content hosting and publication. Here she said that is essential that you follow updated guidance from downstream services such as Crossref to ensure automated attribution and credit. What to check is as follows:

- (1) Register the paper with Crossref and ensure that the metadata is sent to Crossref, including the full reference list. Ensure that all citations are included in the file going to Crossref, and not being removed inadvertently.
- (2) Display the human-readable citation correctly.
- (3) Provide the machine-readable citation correctly to downstream services (e.g., Crossref).

She told the audience that the *Journal Production Guidance for Software and Data Citations* has been updated. She presented the correct information in her talk, but the citation on her slide will soon be out of date. (As an FYI, the updated material was published in *Scientific Data* in September of 2023 [31].)

Patricia Feeney picked up after Shelly and talked about the changes that Crossref has made with regards to data and software citations. She explored the common practices in making data citations machine-readable within journal article citations, focusing on the markup of data citations in the Journal Article Tag Suite (JATS) and the use of Crossref as a data citation metadata endpoint.

Feeney has written an article based upon her presentation that appears elsewhere in this issue of *Information Services and Use*.

## 10. Miles Conrad lecture

A significant highlight of the former NFAIS Annual Conference was the Miles Conrad Memorial Lecture, named in honor of one of the key individuals responsible for the founding of NFAIS, G. Miles Conrad (1911–1964). His leadership contributions to the information community were such that, following his death in 1964, the NFAIS Board of Directors determined that an annual lecture series named in his honor would be central to the annual conference program. It was NFAIS' highest award, and the list of Awardees reads like the Who's Who of the Information community [32].

When NISO and NFAIS became a single organization in June 2019, it was agreed that the tradition of the Miles Conrad Award and Lecture would continue and the first award was given in 2020 to James G. Neal, University Librarian Emeritus, Columbia University. In 2021 the award went to Heather Joseph, Executive Director of the Scholarly Publishing and Academic Resources Coalition (SPARC). In 2022 the year the award was presented to Dr. Patricia Flatley Brennan, Director of the U.S. National Library of Medicine (NLM). And this year it went to Dr. Safiya Umoja Noble.

Dr. Noble is an internet studies scholar and Professor of Gender Studies, African American Studies and Information Studies at the University of California, Los Angeles (UCLA) where she serves as the Faculty Director of the Center on Race & Digital Justice and Co-Director of the Minderoo Initiative on Tech & Power at the UCLA Center for Critical Internet Inquiry (C2i2). She is a Research Associate at the Oxford Internet Institute at the University of Oxford where she was a Commissioner on the Oxford

Commission on AI & Good Governance (OxCAIGG). Her academic research focuses on the Internet and its impact on society. Her work is both sociological and interdisciplinary, marking the ways that digital media intersects with issues of race, gender, culture, power, and technology. She is author of the book, *Algorithms of Oppression: How Search Engines Reinforce Racism* [33], published in 2018. You can find more information on her website [34].

She opened her presentation by saying that her talk is not going to be a lecture, but rather a provocation for our field entitled “decolonizing standards,” and it truly is a provocation. She asked everyone to go on a journey with her and that this is a journey that she has been on over the past few years, ever since the writing of her book, *Algorithms of Oppression*. She thanked everyone who may have read the book and for taking her work seriously. She said that the Miles Conrad Award “recognizes the contributions of those whose lifetime achievements have moved our community forward.” She said that she is currently fifty-three, the age that Miles Conrad was when he passed. Ever since she learned that was to receive the award she has been thinking about what it means to reach certain milestones in one’s career when you think that you actually are just hitting your stride, and she said that she is sure that Miles Conrad himself thought he was just getting started in his work and in his contributions when he passed.

She said it is important to note that we are working in a field that has not made its political commitments entirely clear at all times. For example, we have ideas about access to information and knowledge, and yet we also have institutions, efforts, and histories that have limited scientific information to very specific groups of people and nation-states at the expense of and in a direct effort to preclude others from gaining access to that knowledge. She gave an example of attempts to ensure that scientists in what was then called the Third World or the Non-Aligned Movement [35] did not have access to certain military advancements.

She said that today we continue to deal with the effects of colonization and occupation of Indigenous Peoples’ lands (shades of Joy Owango!). As a field, we have struggled to name our place and figure out our way to contribute to reimagining a more fair and just world. This is at the heart of what it means when we talk about things such as diversity, equity, and inclusion. There are people around the world who are trying to reject colonization and occupation, who are trying to reclaim control over their natural resources, and who are trying to assert independence and liberation and have a fair and full footing with more powerful nation-states, governments, and other peoples around the world.

She talked about the “standardizing” of people - using cranial measurements, measuring melanin in one’s skin, or studying the width of one’s face or brain size - in order to categorize people into racist hierarchies. In closing she said that the field of information standards has much to offer and we should be at the forefront- not passively allowing our work to bolster colonial projects of the past, even if we did not know it was happening. Once we know, we have a responsibility to work with the knowledge to which we are introduced and grapple with it.

Dr. Noble has written an article based upon her presentation that appears elsewhere in this issue of *Information Services and Use*.

## 11. Visualizing institutional research activity using persistent identifier metadata

The speakers in this session were Negeen Aghassibake, Data Visualization Librarian, University of Washington Libraries, Olivia Given Castello, Head of Business, Social Sciences, and Education, Temple University Libraries, Paolo Gujilde ORCID US Community Specialist, Lyrasis, and Sheila Rabun, Program Leader for Persistent Identifier Communities, Lyrasis.

This session investigated the opportunities and current challenges involved in using persistent identifier (PID) metadata to understand institutional research activity. It was based on a 2022 data visualization project led by the ORCID US Community (administered by Lyris) in partnership with two fellows from the Drexel University LEADING program. The fellows created an R script that was used to retrieve information about publishing collaborations between researchers at a home organization and other organizations across the globe, based on metadata from researchers' ORCID records and DOI metadata. The resulting dataset can be imported into a Tableau Public dashboard template, resulting in a data visualization that can be shared with multiple stakeholders to emphasize how PIDs can be used to visualize researcher activity and impact. However, multiple gaps in the ORCID and DOI metadata, such as authors with no ORCID iD or ORCID records with no institution or works data, and missing co-author information in DOIs, indicate that the industry still has a long way to go before these PIDs can be used to demonstrate a more complete picture of research activity.

Aghassibake, Castello, Gujilde, and Raban have submitted a joint manuscript based on their presentation that appears elsewhere in this issue of *Information Services and Use*.

## **12. Unlocking open science in Africa: Mentorship and grassroots community building (the EMEA [36] Keynote)**

The speaker for this session was Dr. Caleb Kibet. He is a bioinformatics researcher, a lecturer, an Open Science advocate, and a mentor. He has a Ph.D. in Bioinformatics from Rhodes University, South Africa. In addition to teaching bioinformatics at Pwani University, Dr. Kibet is a Postdoc at The International Centre of Insect Physiology and Ecology (icipe) in Nairobi. As a 2019–20 Mozilla Open Science fellow, he developed a research data management framework for resource-constrained regions. He is also a member of the Dryad Scientific Advisory Board and a board member of the open Bioinformatics Foundation.

He gave an excellent presentation on the importance of African Research. He referred to a presentation that Joy Owango gave at a NISO Plus conference where she said that it has been stated that Africa accounts for only 0.1% of the world's scholarly content and she went on to say that the figure is incorrect and the reason is that African research is simply not easily discoverable for the following reasons:

- There is a regional bias in Western journals' editorial teams.
- African scholars are inclined to list their Western partner institutions rather than their African home institution.
- A large proportion of present and historical African output is in print form.
- Bottlenecks in infrastructural networks and internet connectivity.

Kibet concurs with Owango and went on to talk about what is being done to improve the discoverability of African research (Owango's passion as well). He mentioned H3ABioNet, an organization focused on developing a bioinformatics capacity in Africa and specially enable genomics analysis by H3African researchers across the continent. He is seeing increased funding for African-driven publishing and highlighted a few of the initiatives. He said that they are also seeing an increase in Open Access Publishing [37], but a low adoption of preprints.

He talked extensively about Open Science and its various definitions and the importance of building communities that collaborate and share information and talked about the difficulties in achieving this because of funding and infrastructure issues. He also talked about the importance of mentorship. He said that he has learned a lot on his journey towards Open Science:

- Greater progress comes from community collaboration.
- Be open by design not by default.
- Be inclusive and supportive: challenges exist, but we can all reap the benefits of Open Science if we work together and support each other.
- Be an ally – create the paths for others to follow.
- Change the culture.

He closed by saying he hopes that “future generations will look at the term “Open Science” as a tautology - a throwback from an era before science woke up. “Open science” will simply become known as science, and the closed, secretive practices that define our current culture will seem as primitive to them as alchemy is to us [38].”

Dr. Kibet has written an article based upon his presentation that appears elsewhere in this issue of *Information Services and Use*.

### 13. Metadata’s greatest hits: Music thru the ages!

This session on the final day of the conference was unbelievable - a combination of humor and education - one of the most original sessions of this conference (or any conference). Marjorie Hlava, President and Founder of Access Innovations, Inc., and Heather Kotula, Vice President, Marketing and Communication, Access Innovations, Inc., were participants in this session and they both kindly provided me with some background on its development. They said that to bring home the topic of metadata in all its forms and contributions to the scholarly records, a group of like-minded people put together a program for the NISO meeting called “Metadata’s Greatest Hits.” They said that something becomes a final and mature form of thought or industry when it can be either spoofed or added to song. That is, the kinks have been worked out and people can make fun of it without breaking it. The goal of this session was to do both by putting the concepts of metadata to song and verse. The lyrics are set to popular songs and are meant to convey the concepts of metadata in a different way. Each presentation was separately written by the person or persons who sang the song.

Hlava and Kotula said that they presented the final song *Amazing PID’s* (to the tune of *Amazing Grace* published in 1779 with words written in 1772 by English Anglican clergyman and poet John Newton). Their hope was that this conveys the notion that identifiers such as the ORCID will provide the author with a handle and traceability far beyond the publication of a single paper. Below are the lyrics of their song. Play *Amazing Grace* in your head as you read.

Amazing PIDs! How sweet the identity,  
That cited an author like me!

I once was anonymous, but now am Id’ed,  
Was unknown, but now I have ID.

‘Twas LAUREN\* that assigned my ORCID ID,  
And CHRIS\* my obscurity relieved;

How precious did that ORCID appear  
The hour I first believed!

The registrar hath promised unambiguity to me,  
Their word my ID secures;

They will my identity and profile be,  
As long as research endures.

When we've been there publishing  
Bright shining as the sun,

We've no less days to peer review  
Than when we first began.

\*Hlava said Lauren Haak and Chris Shillum were very active in ORCID at the beginning, Lauren as its Executive Director, and Chris as early implementor of the technology behind the operation of the ORCID. Many other people also played significant roles, and this is not to diminish their contributions, but only to emphasize that people made the ORCID happen and all authors now appreciate it.

This was a fun and informative session - you must watch the recording to appreciate it.

#### **14. OA usage reporting: Understanding stakeholder needs and advancing trust through shared infrastructure**

The speakers for this session were Tim Loyd, CEO, LibLynx, Tricia Miller, Marketing Manager for Sales, Partnerships, & Initiatives, *Annual Reviews*, Christina Drummond, Executive Director for OA Book Usage Data Trust, University of North Texas Libraries, and Jennifer Kemp, Strategies for Open Science (Stratos).

The speakers in the session noted that the complexity of usage reporting for Open Access content continues to grow, particularly with content syndication to organizations like ScienceDirect and ResearchGate, which deliver content across multiple platforms at an unprecedented scale. They strived to answer the following questions: What kind of usage data do diverse stakeholders (including libraries, publishers, authors, and editors) need? Can the work done to support OA book usage data analytics use cases inform OA article and data use cases? What standards and policies are required to ensure the usage data is accurate and meaningful? What infrastructure is needed to collect and disseminate this data effectively and efficiently?

The speakers brought together different perspectives to consider these questions: an OA publisher, a research infrastructure, an emerging usage data trust, and a usage analytics service provider. They walked through what is known, and then unpacked the questions for which they do not yet have answers. Their goal was to inform community understanding of the challenges ahead and, hopefully, start to lay the groundwork for constructive policies and shared solutions.

Lloyd, Miller, Drummond, and Kemp have written a joint article based upon their presentation that appears elsewhere in this issue of *Information Services and Use*.

#### **15. Additional highlights**

I want to highlight a few additional presentations from sessions that I did not fully attend from speakers who did submit manuscripts that appear elsewhere in this issue of *Information Services and Use*. These are:

*NISO's Content Profile/Linked Document Standard: A Research Communication Format for Today's Scholarly Ecosystem*

*Speaker:* Bill Kasdorf, Principal, Kasdorf & Associates, LLC, Co-Founder, [Publishing Technology Partners](#)

Kasdorf discussed the purpose and importance of the NISO Content Profile/Linked Document standard. The need for the standard is obvious: users demand the delivery of contextualized, targeted content delivered as a natural part of their workflow and publishers aspire to produce machine-actionable FAIR (Findable, Accessible, Interoperable, Reusable) materials, but many publishing workflows are complicated in order to enable print and digital outputs. This standard is an application of HTML5 and JSON-LD to create semantic relationships between data elements in scholarly publishing workflows and express machine actionable content, to ease reuse and interchange of scholarly research information. The format description defines a set of rules that outline the minimal characteristics of documents (Linked Documents) that conform to the standard and a mechanism to define more detailed Content Profiles that extend and refine the rules for specific use cases.

*Open Access and COUNTER usage: Hybrid OA impact on a Private Liberal Arts College*

*Speaker:* Yuimi Hlasten, E-Resources and Scholarly Communication Librarian, Denison University

The popularity of open access (OA) publications has increased in recent years. This situation leads to several pertinent questions for academic libraries: Is the rapid expansion of OA availability negatively affecting their COUNTER usage reports? And, as a consequence, is the increasing accessibility of OA publications prompting academic libraries to reconsider their subscriptions to traditional, subscription-based resources? Hlasten's primary objective was to investigate the impact of hybrid OA growth on Denison University's subscribed e-resource usage and to illuminate the potential effects of OA on academic library subscriptions.

*Recognizing and Harnessing the Transformational power of Persistent Identifiers (PIDs) for Publicly-Engaged Scholars*

*Speakers:* Kath Burton, Portfolio Development Director (Humanities), Routledge, Taylor & Francis, Catherine Cocks, Michigan State University Press, and Bonnie Russell, Product Manager, Humanities Commons

The premise of the talk is that while more recognizable publishing models continue to be favored across academic research and publishing systems, the diverse forms of knowledge emerging from publicly-engaged projects derived from working directly with and for communities requires novel and more dynamic publishing solutions. This presentation considered how the appropriate application of metadata and persistent identifiers to the processes and outputs of engaged scholarship are required to support the goals of the publicly-engaged humanities, and potentially lead to faster and more effective forms of impact for the people and places involved.

*Closing the Gap: Addressing Missing Standards in Small Academic Libraries through the Implementation of the ANSI/NISO Z39.87-2006 (R2017) Data Dictionary.*

*Speaker:* Russell Michalak, Library Director, Goldey-Beacom College.

Michalak presented a case study highlighting the significance of adopting the ANSI/NISO Z39.87-2006 (R2017) Data Dictionary standard to small academic libraries, using Goldey-Beacom College Library in Wilmington, Delaware, as an example. The study focuses on the impact of the standard's absence

on the institution's archival collection and emphasizes the benefits of implementing the standard for small libraries with similar digital collections. Additionally, the paper addresses the challenges faced by small libraries in adopting standards and provides recommendations for overcoming these challenges. The findings emphasize the need for increased awareness and the advantages of adopting the Data Dictionary standard to improve access and management of digital assets.

#### *Sharing, Curation and Metadata as Essential Components of the Data Management Plan*

*Speaker:* Jennifer Gibson, Executive Director, Dryad

Gibson stressed the critical importance that data management plans actually be data management and sharing plans, that encourage data-sharing at the early stages of research; that indicate that data must be made as open as possible (and as closed as necessary); and that specify a minimum level of curation and collection of essential metadata. She also stressed that fact that it is no longer sufficient to post material openly and hopes that others will be able to make use of it. The use of metadata and careful data curation of the data are essential to facilitate the discovery and successful re-use of data.

## **16. Closing**

In his closing comments, Todd Carpenter, NISO Executive Director, noted that the success of this program was a direct result from setting out the global team in advance who would be willing to - and did - engage an international audience. He thanked the thirty-three sponsors who made it possible and allowed the global audience to participate from thirty countries at an affordable rate (*note:* this was one more country than in 2022!). He added that in 2020, two hundred and forty people attended the conference in Baltimore, which was, from his perspective, a great turnout. But he had no reason to believe that the 2023 conference would bring together six hundred and four people to engage in a what was truly a global conversation. He went on to say that NISO could not fulfill its mission and do what it does without the talented and dedicated volunteers who give so much of their time, talents, and expertise.

In closing, Carpenter said that while we are at the end of a three-day journey, even more work will begin tomorrow as NISO assesses all the ideas that have emerged - will they make an impact? Can they transform our world? He asked if any of the ideas struck home with anyone to please send him an email and state what idea(s) are of interest and why.

Note that as of this writing NISO Plus 2024 is being planned and is scheduled for February 14th–13th *So mark your calendars!!*

## **17. Conclusion**

As you can see from this overview, there was no major theme to the conference other than it being a global conversation. Having said that, there were common themes/issues raised throughout and some of them resonated even with the topics of the prior years' conferences.

- Open Science and sharing, citing, and reusing datasets require a cultural and behavioral shift among researchers. The global research community is not there yet and the guidelines that facilitate these activities such as the creation of data and software citations continue to evolve.
- Creating rich metadata is essential to facilitate information discovery and preservation (also a theme in 2021 and 2022). Actually, this conference was pretty heavy on sessions focused on metadata, actually "singing" its praises!.

- Respect for, and the preservation of, Indigenous Knowledge is of global importance.
- Using standards is essential to the global sharing of data and scholarly information (always a theme at any NISO meeting!).

The majority of the presentations that I “attended” were excellent. I thoroughly enjoyed Weinberger’s opening keynote - it made me think about how the combination of the Internet and the digitization of information has changed the information industry since I first entered it more than fifty years ago. The same for Noble’s Miles Conrad Lecture - thought provoking. I always enjoy Shelly Stall’s talks - she can make even a talk about data and software citations engaging due to her energy and sense of humor. And certainly, the session on Metadata’s Greatest Hits was clever, original, factual, and fun - kudos to that group!!!

I always like it when I walk away from a conference with new knowledge. Last year I was blown away by a technology of which I was unaware - Visual-Meta. This year it was back in the program, but there really was no new technology focus this year, which I found disappointing. Others were disappointed as well because respondents to the survey that NISO sent out to get feedback on the conference indicated that topics such as new technologies and new business models, among others, should have been included [39]. However, I was totally unaware of “Practice Research” - so I did walk away with a new concept, and I always like hearing about new scientific activities in Africa so I found Caleb Kibet’s and Joy Owango’s talks very interesting. In fact, the aforementioned survey indicated that the most popular presentations were those by Dr. Kibet and Dr. Noble. Among the most watched recordings are Noble’s Miles Conrad Lecture and Weinberger’s opening keynote.

At the first NISO Plus meeting in 2020 Todd Carpenter called the conference a “Grand Experiment.” When writing the conclusion of my conference overview I honestly said the experiment was successful. I also said that, as a chemist, I am quite familiar with experiments and am used to tweaking them to improve results. And as successful as that first meeting was, in my opinion it needed tweaking. To some extent the 2021 conference reflected positive modifications, but even then, I said that there needs to be more of the information industry thought-leadership that NFAIS conferences offered, and I still hold fast to that opinion. It is interesting that some of the most popular talks this year - Weinberger’s and Noble’s - were about thought leadership. But perhaps I am being unfair. I will repeat what I said last year. In the term “NISO Plus” NISO comes first and when I think of NISO I think of standards and all of the every-day practical details that go into the creation and dissemination of information. I do not instinctively look to NISO to answer strategic questions such as what new business models are emerging? Are there new legislative policies in the works that will impact my business? What is the next new technology that could be disruptive? I had hoped that those questions would be answered to a certain extent in the “Plus” part of the conference title, but to date the “Plus” part has been a much smaller portion of the conference symposia. Perhaps next year it will expand. I hope so!

Having said that, I congratulate the NISO team and their conference planning committee on pulling together an excellent virtual conference. From my perspective, the NISO virtual conferences have consistently continued to be the best that I have attended throughout the Pandemic - technically flawless and well-executed from an attendee perspective. Perhaps NISO should publish a Best Practice on virtual conferences and make it a global standard!

My congratulations to Todd and his team for a job well done!!

For more information on NISO note that Todd Carpenter and Jason Griffey have submitted an article that appears elsewhere in this issue of *Information Services and Use*.

### **Additional information**

The NISO 2024 Conference will take place in-person from February 13-14, 2024 at the Hyatt Regency Baltimore Inner Harbor, and registration [40] is now open.

If permission was given to post them, the speaker slides that were used during the 2023 NISO Plus Conference are freely-accessible in the NISO repository on figshare [41]. If permission was given to record and post a session, the recordings are freely-available for viewing on the NISO website [42]. The complete program is there as well. I do not know how long they will be available, but it appears that the recordings from 2022 are still available.

### **About the author**

Bonnie Lawlor served from 2002–2013 as the Executive Director of the National Federation of Advanced Information Services (NFAIS), an international membership organization comprised of the world's leading content and information technology providers. She is currently an NFAIS Honorary Fellow. She is also a Fellow and active member of the American Chemical Society (ACS) and an active member of the International Union of Pure and Applied Chemistry (IUPAC) for which she chairs the Subcommittee on Publications and serves as the Vice Chair for the U.S. National Committee for IUPAC. Lawlor is also on the Boards of the Chemical Structure Association Trust and the Philosopher's Information Center, the producer of the *Philosopher's Index*, and serves as a member of the Editorial Advisory Board for *Information Services and Use*. This year she was elected to the ACS Board of Directors effective January 1, 2024 and to the IUPAC Executive Committee, effective January 1, 2024.

### **About NISO**

NISO, the National Information Standards Organization, is a non-profit association accredited by the American National Standards Institute (ANSI). It identifies, develops, maintains, and publishes technical standards and recommended practices to manage information in today's continually changing digital environment. NISO standards apply to both traditional and new technologies and to information across its whole lifecycle, from creation through documentation, use, repurposing, storage, metadata, and preservation.

Founded in 1939, incorporated as a not-for-profit education association in 1983, and assuming its current name the following year, NISO draws its support from the communities that it serves. The leaders of about one hundred organizations in the fields of publishing, libraries, IT, and media serve as its Voting Members. More than five hundred experts and practitioners from across the information community serve on NISO working groups, committees, and as officers of the association.

Throughout the year NISO offers a cutting-edge educational program focused on current standards issues and workshops on emerging topics, which often lead to the formation of committees to develop new standards. NISO recognizes that standards must reflect global needs and that our community is increasingly interconnected and international. Designated by ANSI to represent U.S. interests as the Technical Advisory Group (TAG) to the International Organization for Standardization's (ISO) Technical Committee 46 on Information and Documentation. NISO also serves as the Secretariat for [Subcommittee 9 on Identification and Description](#), with its Executive Director, Todd Carpenter, serving as the SC 9 Secretary.

## References

- [1] See: <https://niso.cadmoremedia.com/Category/8e837242-5e4d-4d0e-ac1c-421a695ee1c6>, accessed October 5, 2023.
- [2] History.com editors, “Model-T”. The History Channel, Apr. 26, 2010, <https://www.history.com/topics/inventions/model-t>, accessed October 23, 2023.
- [3] See: [https://en.wikipedia.org/wiki/Mimal\\_viable\\_product](https://en.wikipedia.org/wiki/Mimal_viable_product), accessed October 23, 2023.
- [4] A. Baus, Examples of successful apps that were MVPs first, *Decode* (2022), <https://decode.agency/article/app-mvp-examples>, accessed October 23, 2023.
- [5] See: <https://openai.com/blog/chatgpt>, accessed October 23, 2023.
- [6] See: <https://scholarlykitchen.sspnet.org/building-pipes-and-fixing-leaks-in-scholarly-content-discovery-and-access/>, accessed October 26, 2023.
- [7] See: <https://www.morebrains.coop>, accessed October 23, 2023.
- [8] See: [https://en.wikipedia.com/wiki/Persistent\\_identifier](https://en.wikipedia.com/wiki/Persistent_identifier), accessed October 23, 2023.
- [9] See: <https://www.go-fair.org/fair-principles/>, accessed October 26, 2023.
- [10] WikiCite - Meta [Internet]. [cited 2023 May 30]. Available from: <https://meta.wikimedia.org/wiki/WikiCite>, accessed September 23, 2023.
- [11] See: <https://openalex.org>, accessed October 30, 2023.
- [12] See: <https://www.equinoxoli.org>, accessed September 27, 2023.
- [13] See: <https://www.equinoxoli.org/about/#History>, accessed September 27, 2023.
- [14] R. Molyneux and T. Hamby, The consortial effect in detail: The scends experience. in: *Library and Book Trade Annual*, 57th ed. 2012.
- [15] D. Yaga, P. Mell, N. Roby and K. Scarfone, NISTIR 8202 Blockchain Technology Overview, National Institute of Standards and Technology, U.S. Department of Commerce, 2018. <https://nvlpubs.nist.gov/nistpubs/ir/2018/nist.ir.8202.pdf>, accessed September 11, 2023.
- [16] See: <https://www.lyrasis.org>, accessed October 28, 2023.
- [17] See: <https://www.openathens.net>, accessed October 28, 2023.
- [18] See: <https://www.bibtex.com/g/bibtex-format>, accessed October 28, 2023.
- [19] See: <https://www.latex-project.org>, accessed October 28, 2023.
- [20] See: <https://visual-meta.info/visual-meta-2/>, accessed October 28, 2023.
- [21] See: <https://seamlessaccess.org>, accessed October 30, 2023.
- [22] M. Sraaku-Lartey, Connecting the world through local indigenous knowledge, *Information Services & Use* 41(1-2) (2021), 43–51, Available from: <https://content.iospress.com/journals/information-services-and-use/41/1-2?start=0>, accessed October 28, 2023.
- [23] See: <https://info.africarxiv.org>, accessed October 28, 2023.
- [24] See: <https://www.go-fair.org/fair-principles/>, accessed October 28, 2023.
- [25] See: <https://www.gida-global.org/care>, accessed October 28, 2023.
- [26] See: reference 9.
- [27] B. Lawlor, An Overview of the 2021 NISO Plus Conference Global Connections and Global Conversations, *Information Services and Use* 41(1–2) (2021), 1–37, see specifically page 5-9, available from: <https://content.iospress.com/journals/information-services-and-use/41/1-2?start=0>, accessed October 29, 2023.
- [28] See: <https://datascience.codata.org/article/10.5334/dsj-2020-005/>, accessed October 29, 2023.
- [29] See: <https://data.agu.org/resources/availability-citation-checklist-for-authors>, accessed October 29, 2023.
- [30] S. Stall, G. Bider, M. Cannon et al., Journal production guidance for software and data citations, *Scientific Data* (2023). doi:10.1038/s41597-023-02491-7, accessed October 29, 2023.
- [31] See reference 27.
- [32] See: <https://www.niso.org/node/25942>, accessed October 29, 2023.
- [33] See: [https://en.wikipedia.org/wiki/Algorithms\\_of\\_Oppression](https://en.wikipedia.org/wiki/Algorithms_of_Oppression), accessed October 29, 2023.
- [34] See: <https://safiyaunoble.com/bio-cv/>, accessed October 29, 2023.
- [35] See: [https://en.wikipedia.org/wiki/Non-aligned\\_movement](https://en.wikipedia.org/wiki/Non-aligned_movement), accessed September 20, 2023.
- [36] EMEA stands for Europe, the Middle East, and Africa.
- [37] K.W. Mwangi, N. Mainye, D.O. Ouso, K. Esoh, A.W. Muraya, C.K. Mwangi et al., Open science in Kenya: Where are we? *Front Res Metr Anal* (2021), [cited 2022 Sep 16]; 6. Available from: <https://www.frontiersin.org/articles/10.3389/frma.2021.669675>, accessed October 30, 2023 [Internet].

- [38] M. Watson, When will ‘open science’ become simply ‘science’? *Genome Biol* **16**(1) (2015), 1–3, <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0669-2>, accessed October 30, 2023.
- [39] See: <https://niso.plus/2023/03/niso-plus-2023-post-event-survey-the-results-are-in/>, accessed October 30, 2023.
- [40] See: <https://www.niso.org/events/niso-plus-baltimore-2024>, accessed October 30, 2023.
- [41] See: <https://nisoplus.figshare.com>, accessed October 30, 2023.
- [42] See: <https://niso.cadmoremedia.com/Category/8e837242-5e4d-4d0e-ac1c-421a695ee1c6>, accessed October 30, 2023.