

An overview of the 2022 NISO plus conference: *Global conversations/Global Connections*

Bonnie Lawlor*

NFAIS Honorary Fellow, 276 Upper Gulph Road, Radnor, PA, USA

Abstract. This paper offers an overview of some of the highlights of the 2022 NISO Plus Annual Conference that was held virtually from February 15 - February 18, 2022. This was the third such conference and the second to be held in a completely virtual format due to the Pandemic. 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 last year, there was no general topical theme, but there were topics of interest for everyone working in the information ecosystem - from the practical subjects of standards and metadata quality to preprints, Wikidata, archiving and digital preservation, Open Science and Open Access, and ultimately Globalization of the Information Infrastructure, the Metaverse, and Visions of the Future. With speakers and attendees from around the world and across multiple time zones and continents, it truly was a global conversation!

Keywords: NISO, metaverse, metadata, open science, open data, information accessibility, wikidata, knowledge graphs, virtual reality, sustainable libraries, scholarly ethics, climate change knowledge cooperative, sustainable development goals, artificial intelligence, machine learning, discovery services, FAIR data, metrics, open access, indigenous knowledge, digital preservation, archiving, knowledgebases, information sharing, information discovery, miles conrad lecture, scholarly communication, frame, information preservation policy, COPIM W7, NLM, visual-meta, PDF problems, authoritative information, zooniverse, co-competition, dynamic semantic publishing, data lakes, pluriverse

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 took on 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 one year and the third NISO Plus Annual Conference was also held in a completely virtual format. The general theme from 2021 continued - “global conversations/global connections” and again speakers were 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. The conference

*E-mail: chescot@aol.com.

attracted six hundred and thirty attendees - an increase of 270% over 2020. Participants came from twenty-eight different countries (two more than last year), nearly every inhabited time zone, and from the vast majority of continents. They 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. One hundred and forty speakers participated in forty-six sessions, most of which were recorded and are now freely-available [1] for viewing - if you have more than thirty hours in which to watch nine hundred and seventy-four gigabytes of video!

Todd Carpenter 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 two 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 just 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 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 and ask how those ideas could make a difference in their own organization or community and how they 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 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 has been 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 others will find homes in other organizations who are interested in nurturing the idea and have the resources to do so. From the 2021 Conference NISO launched three specific projects and several other ideas are now germinating in other communities as well.

He went on to say that it is important to note that the process that he has described is not easy - it takes a lot of effort - and NISO has many dedicated volunteers who work very hard to make things happen. They care about improving the world - to bringing ideas to life - through a consensus process that takes time and dedication.

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.

I can attest that at least for the sessions to which I listened the discussions were interesting, in-depth, and some did generate ideas. Be forewarned - as I noted last year, I am fairly good at multi-tasking, but even I cannot attend four meetings simultaneously and I did not, after the fact, 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. That is my personal goal with this 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. Siva Vaidhyanathan who is the Robertson Professor of Media Studies and Director of the Center for Media and Citizenship at the University of Virginia. He is also the author of *Antisocial Media: How Facebook Disconnects Us and Undermines Democracy* (Oxford University Press, 2018), and *The Googlization of Everything – and Why We Should Worry* (University of California Press, 2011). Vaidhyanathan said that when Facebook changed its corporate name to “Meta” in November 2021, it planted a flag as a leader in a widespread movement to generate a new sense of human-computer interaction - a new sense of human consciousness. The promise of this radical, new vision that combines artificial intelligence, virtual reality, augmented reality, and cryptography is immense. So are its potential threats to human flourishing. His presentation walked through what are from his perspective the major elements and controversies of the movement towards a “metaverse”, and he outlined the potential benefits of living with such technology (Warning: I am a Metaverse fan).

He made it very clear that when Mark Zuckerberg changed the name of Facebook to Meta [2] that Zuckerberg did not, in his opinion, clearly articulate his vision for a “Metaverse [3]”, which can briefly be defined as a virtual-reality space in which users can interact with a computer-generated environment and other users. Vaidhyanathan felt that Zuckerberg failed to link his vision to a wider discussion of metaverse in a broader community. A community of people who have been involved in virtual reality research, who have been involved in game development, and who have been involved in almost utopian visions of what technological change might do for us - all of which Vaidhyanathan believes is crucial to understanding the next decade or even the next twenty-five or thirty years.

He went on to say that we owe Zuckerberg a big “thank you” for changing the name of his company because it has generated so much buzz and discussion. It has forced people to find out what a Metaverse is and its possible implications for the future. He asked what it would mean to take such a vision seriously? Is Metaverse just really good virtual reality? He does not think so. It has been difficult to establish widespread consumer interest in virtual reality beyond the gaming niche. And beyond specialized training, it has not been adopted commercially very well. He believes that Metaverse is something much bigger and much different. He encouraged everyone to take a look at augmented reality - an interactive experience combining a real-world environment and computer-generated content. He asked that the audience think about the companies that are investing significantly in augmented reality and virtual reality and understand

that they are using the same teams to do the research and product development. He believes that at some point there will be a fusion between virtual and augmented reality and wonders what that will mean to daily life? What will that mean to Human consciousness?

He went on to talk about wearable sensors and smart clothing. He said that he has had students at the University who are on the football team and who wear “Smart” clothing while they go through practice. All of their body signals are monitored, tracked, and graphed. He asked them what they think about it, and they say it is all about optimal training, e.g., keeping them from overtraining. He noted that the concept of self-tracking and wearable has become increasingly popular. The Fitbit was the first successful consumer product that does this sort of thing. Now, all of these services are directly connected to servers and data is being collected. Basically, the Human body is being tagged and monitored, Vaidhyathan added that things in which human bodies exist are also being monitored and used the Tesla car as a prime example. It is an automobile that is completely connected, constantly monitoring things, and reporting back to a central power. This is in addition to other such monitors that we invite into our homes through thermostats, appliances, alarm systems, etc.

To my surprise he then talked about cryptocurrencies and said that this is the most interesting layer of the metaverse vision. In the current metaverse there are economies developing in order to purchase capabilities, items, identities, etc. and a currency is needed that can be accepted across borders. So various forms of cryptocurrency are being selected to fulfill that requirement. This makes it possible to have a simulated economy and what is emerging is the use of simulated currency to run a simulated economy in a simulated world. And layered on top of that is the blockchain-enabled practice of non-fungible tokens, NFTs, which have been getting a lot of notoriety and are currently being used to fund scientific research! [4] (*Note: This hit a button with me. For the past three years I have been part of a team developing a White Paper for the International Union of Pure and Applied Chemistry (IUPAC) on the use of Blockchain Technology along the scientific research workflow and during the past year NFTs and the metaverse have been creeping into the conversation!*). He went on to say that cryptocurrency and NFTs are becoming a very big part of the metaverse vision and is one of the reasons why Facebook is getting into this business. They are looking for a way to facilitate economic exchange.

He noted that when we talk about the metaverse we are talking about diverse human interactions - being connected to networks and dataflows. Back in the 1990s we talked about logging on computers and eventually logging off. And this is one of the things he wants people to think about. In the past we thought about digital networks, digital communication, and digital interaction under a very different model of daily life and interaction. One in which for a small portion of the day, one would choose to log on to a chat room, interact, and then move away. And there would be a clear distinction between what we might call the real world and what we might call the online world or the cyber world. Now we carry with us, on us, on our skin sometimes, devices that are always on. And therefore, we are always on. There is no in or out. There is no on or off. There is just being. That is the metaverse!

He went on to say that we could look at this in a nostalgic way. We are going back to the vision that we had in the 1990s, where we could create an ideal world, control our identity, and escape from the bonds and troubles of daily life. We tried it and thought that the internet would do that for us. It did not. We carried all of our old baggage into these new environments. Today the data flows through us and there is no distinction between online and offline. We are on all the time via devices and sensors. But now, this vision, at least the virtual reality part, seems to invite us into that 1990s framework of escape, of a new start.

He urged everyone to get beyond the virtual reality notion of metaverse and understand that it is a fully-connected collection of human bodies and minds. This has significant and different implications

than just creating a better virtual reality He has no answers, but he has a lot of questions he plans to think about over the next few years, such as: What are the implications of all these efforts to enhance, embed, infuse virtual reality, augmented reality, haptics, wearable technologies, self-tracking, smart devices and appliances, automobiles, smart cities, and cryptographic assets? What could happen? What useful things could emerge? For people with limited abilities, a lot of these technologies can be tremendous life enhancers. You can imagine that certain virtual and augmented reality technologies can do a lot to enhance the quality of life for someone with limited sight, limited hearing, limited mobility, limited dexterity. Haptics [5] is another great area of research for this. Given the prospects for translation in real time as humans interact, that has tremendous prospects to enhance human well-being as well. Back in the 1990s, we were talking about data and documents. What can we do with data? What can we do with documents? How can we enhance knowledge from data to documents to bodies and minds? Today we no longer talk about data and documents. The data and document game is not of interest to current investors, inventors, entrepreneurs, and corporate leaders. They are interested in bodies and minds.

Vaidhyanathan believes that Zuckerberg is largely responsible for the shift. He believes that Zuckerberg never cared about the things that Google cared about and that this is the major difference between the two companies. Google's founders were very clear about their mission which to this day remains their mission statement - to organize the world's information and make it universally accessible and useful. It has always been about information. But Zuckerberg never cared about information. He cared about people. He cared about bodies and minds and how they might be monitored, managed, monetized, and manipulated. And that is exactly what he has accomplished. Facebook/Meta is all about monitoring, monetizing, and manipulating minds.

The big question *we* should be asking is what can the technologies that comprise the metaverse do to enhance human flourishing? Vaidhyanathan had already mentioned the ability to compensate for limited human capabilities. But are there other ways that human flourishing can be enhanced? Perhaps art in a virtual reality platform can be mind-blowing in ways that have not yet been experienced just as film was an enhancement to the art world, the creative world. Perhaps in the areas of translation and certainly in training. On the downside, there are ways in which the opposite could happen and only the corporate control of our lives and decisions is enhanced.

He went on to say that from about 1995 to 2015 society in general did not pay attention to the big question that he just raised, and did not pay attention to potential negative outcomes. We made the assumption that what was good for Google or Facebook would be good for humans. The same can be said for Amazon and Microsoft (this sentiment was repeated by other speakers throughout the conference). We just let these companies build their systems to their own specifications and to fill their own needs. And we have discovered that there is a huge price to pay as a result. He encouraged everyone not to repeat the same mistake.

In closing, he again said that Zuckerberg has done us a favor. By laying out the vision of a metaverse, he has allowed us to ask the basic questions again. Only now *we* have more knowledge and awareness, and *we* can actually ask harder questions with better information. Who knows - *we* might actually be able to guide the next few major technological decisions in a healthier way than we have over the past couple of decades.

As an aside, this Opening Keynote somewhat echoed that of the 2021 NISO Plus conference that was given by Cory Doctorow, a science fiction writer and journalist. A summary appears in the overview of that conference [6].

3. Working towards a more ethical information community

There were four parallel sessions immediately following the Opening Keynote and I chose to sit in on one that looked at how libraries and publishers are working to become more ethical in their production and distribution of information, especially within the context of the United Nations' seventeen sustainable development goals (SDG) laid out in 2015 [7]. More than one hundred and fifty publishing organizations have signed the SDG Publishers Compact, and the American Library Association (ALA) added sustainability to their core library values in 2019. The speakers in this session discussed the importance of ethics from a variety of viewpoints and shared the steps that they are taking to become more ethical.

3.1. *The just transition*

The first speaker was Rebekkah Smith Aldrich, the Executive Director of the Mid-Hudson Library System in the State of New York, and the current Chair of ALA's Council Committee on Sustainability. She is also the sustainability columnist for the *Library Journal*. She talked about the need for a transition, not only of the economy, but of our mindsets, the work that we do, and how we do that work. She said that as we start to think about how we do our work and the frameworks in which we make decisions, we have to take a more global view than we have in the past about decision making and about how we control the future from our own sphere of influence. The work that Aldrich has been doing both in New York state with ALA and now with the national program, the Sustainable Libraries Initiative [8], is attempting to develop a methodology for working in libraries, whether they are public, academic, or school libraries, in order to create better decision-making processes that result in more ethical decisions by libraries as institutions and that create a more ethical and thriving future for the people that libraries serve.

She said that libraries cannot successfully do their work without recognizing and valuing diversity and pushing forth the idea that success will not come if there is no empathy, respect, and understanding for our neighbors, regardless of their backgrounds and where they come from. She added that no matter where we sit in the information profession, part of our work is to allow everyone to be heard through the work that we do in a variety of formats, industries, and sectors. Indeed, everyone in the information community is well positioned to be a powerful player in this transition to a brighter future for all, which requires an ethical framework for making decisions, both personally and at the institutional level. Sustainability has an urgency tied to it today, to the point where libraries need to embed it into the thinking of the profession and to that of the next generation of leaders. In fact, she noted that the resolution that made sustainability a core value of ALA also mandated that accreditation standards for library schools throughout the USA now teach sustainability in graduate school to facilitate a shift in mindsets. ALA actually adopted the triple bottom line definition of sustainability to help people understand that sustainability is not just talking about going green or the climate as it results and relates to the environment. Sustainability is the balance of environmentally sound choices that are (1) fair to all involved, (2) socially equitable, and (3) economically feasible because we can no longer continue to make decisions that are unaffordable for us both locally and globally. It is the balance of these three items that is the goal in decision making. That balance, she went on to say, will provide a mindset and a framework for making more ethical decisions as professionals, as institutions, and ultimately, as local and global communities. She noted that this may not be a radical idea, but when we understand the radical mind shift that is necessary to truly make change happen in our world, we understand that we are touching on a very delicate topic - economics. Her position is that we are in a transition to an economy and a way of thinking that is necessary for a brighter future for us all.

Aldrich said that the concept of a “just transition” has basis key elements to it that must be considered as we think through the future of our institutions and how we do our work: advancing ecological restoration as the natural world around us is suffering quite a bit; democratizing communities and workplaces so that more voices are heard; driving racial and social equity to make sure that past wrongs are righted; encouraging empathy and the gap bridging that is necessary to have more people involved in the decisions that affect their lives; thinking very heavily about your locale and really speak to the need to grow local economies; and retaining and restoring cultures and traditions as it is important to celebrate the wisdom of the past and to learn from that, either whether they be mistakes or achievements that have been made in the past, but to respect what has come before us (*Note*: this was a common theme among several presentations, especially with regard to indigenous knowledge). She emphasized that the more work that is done to strengthen the fabric of our communities, to help people understand, respect, and have empathy for one another, is going to build the resilience of those communities to withstand disruption.

In closing, she said that our institutions need to: lead by example from the inside out; make policy decisions that are good for all; treat their workers well; run their facilities in a way that is ethical and environmentally conscientious; and serve as a model for others in their community wherever they are geographically located. She suggested that those interested look at the roadmap [9] developed for the Sustainable Libraries Initiative.

3.2. *The American Society of Civil Engineers, the Climate Change Knowledge Cooperative, and the SDGs*

The second speaker in the session was Dana M. Compton, Managing Director and Publisher, the American Society of Civil Engineers (ASCE) who talked about how the ASCE has incorporated mission-based goals such as the United Nations Sustainable Development Goals (SDGs) and the Climate Change Knowledge Cooperative (CCKC) into its publishing strategy. She noted that sustainability is a key strategic area for ASCE as an organization to the point where about a year and a half ago a chief sustainability officer was appointed. They have a committee on sustainability, as well as one focused on adapting to climate change. And they have a sustainability roadmap that frames sustainability as a strategic issue for practicing civil engineers that must be integrated into professional practice. The roadmap is really intended to transform the civil engineering profession with respect to sustainability. The roadmap has four main priorities with the first three inwardly-focused on civil engineering practice, and the fourth focused outward on communication as follows:

- Sustainable Development Project: Do the right project.
- Standards & protocols: Do the project right.
- Building capacity: Expand technical capacity.
- Communicate/advocate: Making the case.

She pointed out that in the fourth priority, she sees that ASCE as an organization assumes a responsibility to communicate the need for transformational change to promote a sustainable infrastructure with all stakeholders - not just their membership, author base, or readership - but also the public, which is different from their priorities in other areas of what they do as a society. And ASCE is stating that its end goal for priority four on the sustainability roadmap is for both the membership and the public to demand a sustainable infrastructure to enable healthy and thriving communities (echoing Aldrich’s sentiment). But for our members and the public to demand something, they need to fully understand it.

Compton then drilled down into ASCE's publication strategy that is organized around six guiding principles, two of which dovetail with the ASCE's sustainability goals. The first is sustainability which for ASCE means positioning themselves as a leading publisher of sustainability and civil engineering content. Some of the goals here include assessing their existing content, developing content collections that raise the profile of the great volume of sustainability-related content that they already have, as well as guiding decisions about future acquisitions, ancillary content opportunities, etc. The other pertinent principle for them is accessibility. And this encompasses not just the typical technical aspects that you might think of, but also seeking opportunities to translate technical content for a less technical or a non-technical audience. This includes things such as synopsis statements, videos, infographics, podcasts, etc. They focus on finding channels to expose that accessible content to audiences outside their typical academic audience.

Compton noted that as with anything, ASCE faces a number of challenges in achieving their goals. They have a huge amount of applicable content. ASCE publishes thirty-five research journals and more than forty-six hundred articles per year. They have a book program with more than fifty front list titles per year and conference proceedings that add more than four thousand papers per year, and they also publish standards. Much of the content has an element of sustainability with sustainable infrastructure in it. Therefore, curating that content and making sure that the right pieces are discoverable to the right users is quite difficult. She said that they have made some progress with more granular taxonomy, primarily aligned around the UN SDGs. But in some cases, even an automated solution raises tens of thousands of search results. A manual, human curation element uncovers the most relevant, up-to-date, highest quality subsections of content, but that puts more burden on a small staff (just over thirty). And she admitted that ASCE staff are neither civil engineers nor subject matter experts, they are publishing professionals. So, the work would fall on their already-overburdened volunteers. They have begun relying on authors and editors to provide some practitioner-focused contextualization, but they are not positioned to translate content for multiple audiences, including the public. So that simply has not been happening. Also, ASCE's reach within publications is primarily to academia. They do have some ASCE channels that reach their membership, which is largely practicing civil engineers. But for the most part, ASCE content is consumed within the profession.

She said that they had an opportunity to participate in the Climate Change Knowledge Co-operative (CCKC) [10]. It is a major, new collaborative initiative to help broaden the discovery and understanding of climate change research and accelerate its application towards a sustainable future. It is co-organized by KUDOS and Impact Science. There is a group of participating publishers and sponsors and the way it works is that CCKC offers several levels of participation at different price points based upon the number of research articles that a publisher wants to showcase. In return for the publisher's investment, Impact Science creates the plain language summaries, the videos, or infographics via accessible content, to post in an overall cross-publisher showcase and in a branded publisher showcase if the publisher so chooses. The CCKC goal is to help publishers, societies, universities, etc. maximize their reach to broader audiences using understandable lay language, to offer publicity beyond academia at an accessible price point due to economies of scale, and to demonstrate how scholarly communication is instrumental in tackling climate change.

ASCE invested in this initiative knowing that it is a "product" that generates zero revenue. They justify their participation because it supports their publishing strategy. Compton went on to say that they measure success using metrics supplied by KUDOS such as social media mentions, engagement exposure, the press, etc. She noted that usage statistics are quite telling. Initially they put five articles into the initiative. They looked at overall usage in the first month after online publication and then compared that to the first month after the article summaries were added to the CCKC showcase. She pointed out that the articles

are getting much more usage post-CCKC inclusion than they had upon initial release. She admitted that there could be factors other than CCKC, but she finds this compelling at least in terms of looking at the one factor

Article	1st month post e-Pub	1st month post-CCKC
1	31	132
2	68	185
3	21	105
4	48	86
5	12	171

that she knows has changed for these articles and that is the CCKC posting. She said that the first two articles were both about three, to four months old at the time of being added to CCKC. And those articles have now been in there for about the same amount of time. So, what can be seen is that in the three or four months before being added to CCKC accounts for about 23% of the usage, and in the three to four months after, 77% of the usage. ASCE plans to add fifteen articles over time in batches of five. All have summaries and some will have videos or infographics.

In closing, Compton said that looking further ahead, her team has been creating curated collections that align to four additional UN SDGs. They plan to create landing pages for those on their own site, but she said that if anyone from Kudos or Impact Science is listening, ASCE would love for the CCKC initiative to continue. She also invited listeners to reach out to her if they want to learn more and provided her contact information (dcompton@asce.org).

3.3. *Embedding SDGs in strategies and products*

The third speaker in the session was Andy Robinson, Managing Director, Publishing at CABI (see: <https://www.cabi.org>). He said that CABI is an international not-for-profit organization that improves people's lives by providing information and applying scientific expertise to solve problems in agriculture and the environment. They help farmers grow more and lose less of what they produce, combating threats to agriculture and the environment from pests and diseases, and improving access to scientific knowledge. Robinson admitted that CABI is not a typical publisher. It is a not-for-profit owned by forty-nine member governments and is dedicated to solving problems in agriculture. He said that as a mission-driven organization, it is fair to say that everything CABI does should align with the United Nations' SDGs and they do focus on eight of the seventeen SDGs in particular: #1, Eliminate poverty; #2, Eliminate hunger; #4, Provide quality education; #5, Gender equality; #12, Responsible Consumption and Production; #13, Climate action; #15, Life on land; and #17, Partnerships for the goals.

CABI's first goal is to improve food security and the livelihoods of smallholder farming communities. He said that globally there are approximately five hundred million smallholder farmers who face many barriers to selling their produce, which restricts their earning power and keeps them in poverty, whether that be from production and post-harvest losses, to compliance with food quality standards, to accessing finance and credit. He added that nearly a billion people go hungry every day, while 30% to 40% of

crop production is lost each year to pests and diseases. If crop losses can be reduced by 1%, millions more people could be fed. CABI contributes to this by strengthening national plant health systems. For example, they provide information resources, such as the Plantwise Knowledge Bank [11]. He said that farmers around the world will have to feed an estimated global population of nine to ten billion people by 2050, which means that they need to produce more food from the available land and water and reduce food loss and waste. He noted that food production is already responsible for a quarter of greenhouse gas emissions and that proportion will only increase.

Meanwhile, smallholder farmers are already experiencing climate disruptions due to weather patterns and, as climate changes, there are far-reaching implications for what crops can be grown where, along with the migration and spread of crop pests. In the face of shifting pest outbreaks, an example of how CABI supports farmers is the Pest Risk Information Service (PRISE) [12], which is funded by the UK Space Agency, where CABI delivers text alerts to farmers on pest outbreaks based on geospatial data and the modeling of climatic conditions and pest epidemiology.

Robinson noted that CABI also focuses on reducing inequality for women and young people. He said that women make up 43% of the global agricultural workforce. In some areas, farming is almost entirely carried out by women as men migrate to cities to find work. But women's production levels are about 20% to 30% lower than men's because of lack of access and control over land, labor, credit, knowledge, and market opportunities. Also, young people do not see agriculture as an attractive career option for many of the same reasons. CABI now explicitly builds gender considerations into all of their development projects. More specifically, they tailor information, training, and communication channels for women and young people according to local custom and practice.

Robinson went on to talk about a case study to show just how CABI has started to address one of the big SDG challenges via their products - the problem of invasive species, which cost the world \$1.4 trillion dollars every year. The Fall Armyworm, for example, hopped on a boat from the Americas and decimated maize crops as it ate its way across Africa and Asia. Invasive species undermine food security and economic growth, contribute to population migration, and cause massive loss of biodiversity. Therefore, it is not surprising that invasive species are a specific SDG target as part of SDG 15, life on land. But they impact almost all the SDGs. Parthenium, for example, also known as famine weed, causes dermatitis and respiratory problems in humans. Seventy percent of schoolchildren leave school during peak weeding times to control invasive plants. And in Africa, one hundred million women spend twenty billion hours weeding every year. Invasive species undermine water supplies and energy production. They de-oxygenate aquatic environments. And the challenge is only growing due to increased trade, travel, and climate change.

To address this, CABI has launched an Action on Invasives Program [13] in 2018. And one of the enduring outputs from that project was the Invasive Species Compendium (ISC) [14]. This is an encyclopedic resource that supports decision making in invasive species management, and it focuses on the global invasive species that have the greatest impacts on livelihoods and the environment. It contains twelve thousand curated data sheets on their distribution, pathways of entry, their natural enemies, and five thousand practical guides to identification and management, which means it is useful to researchers, to risk assessors, land managers, and plant protection officers. The species distribution data has been abstracted from thousands of research articles over decades and that distribution data powers the horizon scanning tool. It is a decision support aid that helps to identify species that might enter a particular country or area from another country. Literally, it helps to predict what invasive species might be just over the horizon.

In closing, Robinson said that in recognition of SDG 17, Partnerships for the goals, that he would like to thank the many supporters of the ISC who made its creation possible. He said that he could have mentioned many other CABI knowledge initiatives and said that some of them are highlighted on the website of the

International Publishers Association (IPA) [15]. He also encouraged the publishers in the audience to sign up to the SDG Publishers Compact [16] that was launched by the United Nations and the International Publishers Association in 2020, and to embed action into their organizational and product development strategies.

3.4. Sustainability, scholarly ethics, and North-South divide: A search for connection

The final speaker in this excellent session was Dr. Md. Haseeb Irfanullah, an Independent Consultant on the Environment, Climate Change and Research Systems and a visiting research fellow of the Center for Sustainable Development at the University of Liberal Arts Bangladesh. He spoke about sustainability, research, and scholarly ethics, and how these relate to what he called the north and south divide - basically the divide between institutions enjoying healthy economies (the North) and those in countries that are poor (the South). He noted that if we look at the world around us, we can see that there are many societal and environmental challenges, social inequity, etc. We see poverty and food and water insecurity. He said that these challenges are not only for now, but for the future and have a transgenerational dimension component. That is why when the world came together in 2015 to develop and agree upon United Nation Sustainable Development Goals (SDGs), it bound us together in all the different dimensions, - the social, economic, and environmental.

He went on to say that to achieve the SDGs we need research and access to information, and he then summarized the research workflow. A researcher needs to access past research through journal articles, theses, books, databases, etc. They then design their proposed research project; seek funding; perform experiments; collect, analyze, present, and interpret data; and then draw conclusions. Finally, they communicate their findings via research reports, PhD/Master's Theses, journal articles, peer-review, etc. Then their work becomes part of the body of past research for others to build upon. He asked that if we keep this framework in mind, where does the ethical dimension fit in? The obvious areas are in research communication. There are ethical concerns when we talk about authorship, editorial editorship, editorial board makeup, peer review, risk reduction investigation, etc. But what about ethical concerns when it comes to the North/South Divide?

He said that some progress has been made and as an example he mentioned the Research4Life [17] initiative that provides institutions in lower income countries with online access to academic and professional peer-reviewed content. The initiative aims to improve teaching, research, and policy-making in health, agriculture, the environment, and the life, physical, and social sciences. He went on to say that when we talk about publishing research there is some criticism regarding Open Access publishing because sometimes there are very high Article Processing Charges (APCs). However, publishers can (and do) waive fees, or offer discounts to the researchers from the global south. And when we talk about publishers supporting the global equity there are regional Open Access journals that are enhancing global or geographical equity among different geographical regions. But there are some situations which might not be quite that ethical.

As an example, he mentioned that publishers ask peer reviewers to offer their service for free while the publisher, in parallel, charges thousands of dollars for institutions to subscribe in order to read those peer reviewed articles. The same goes for authors (who are often peer reviewers). When a researcher wants to publish their article in a good peer-reviewed, Open Access journal, they must pay thousands of dollars. He said that he himself often serves as a peer reviewer from the global south, but if his institution does

not have access to that particular journal, he is unable to read the final version of the paper that he peer-reviewed for free. Similarly, if there are pieces of research done on his country and it has been published in a journal which is not Open Access, his country as a whole will not be able to access that research.

The final scenario that he presented is the following. He said that the lower- and middle-income countries are progressing economically. His country is a lower-income country that hopes to become a lower middle-income country in a few years. As a citizen of a lower-income country he enjoys access to thousands of journals for free through the Research4Life for program. But when his country graduates from a lower-income country to a lower middle-income country that few access will stop. This is problematic because while his country's economy is improving, their investment in research systems is not progressing at the same pace. Suddenly stopping support from an initiative such as Research4Life will seriously impact science in his country.

He went on to ask two broad questions: (1) Can we have an ethical information or research community without reducing the North-South divide while meeting the SDGs? and (2) Are we contextualizing the ethical considerations of the North for the South? He did not answer them, but went on to ask some specific questions to get the audience to think about the issues. Here are three of the questions:

- If researchers in lower-income countries have free access to global research, but will publish only in local journals, can you say that they are being unethical because they are exploiting free access to the global literature?
- If researchers in lower-income countries say that as a result of Article Processing Charges (APCs) being so high that they cannot afford them, they will only publish in the freely-accessible local journals to which they have access. Can you blame them? Also, certain institutions demand that their scientists publish a certain percentage of their research locally.
- Can we expect Southern researchers to abide by the ethical guidelines that are articulated by the North without the North helping them to change their systems?

He went on to say that if we really want to make our scholarly society more ethically aware and ethically active, we need to do three things.

- (1) Create a collective narrative of scholarly ethics with regional and international partners to contextualize the South's issues and concerns.
- (2) Work with the South to support its understanding of scholarly standards, the need for structural changes, and evidence-guided actions for the SDGs.
- (3) Look beyond our current networks in the North. We need to learn from, capitalize on, and harness a diverse pool of regional and international perspectives and expertise.

In closing, he said that we need to change the way we narrate ethics in scholarly communication. But that change will not be made by the North. The North should not call out the South to please join them to change the system. Of course, the South should not wait for a call from the North to make those changes. We, the North and the South, need to come together, meet in the middle, collaborate, and create a more fiscally-responsible global scholarly system and information community.

Note: There was a parallel session in this time slot entitled, "The Importance of Investment in Open Research Infrastructure" that I wanted to attend. I tried to access the recording afterwards, but to this day I keep getting blocked. One of the speakers, Ana Heredia, an independent consultant, gave an overview of Latin America's current scientific information infrastructure, highlighting its key role in the adoption of Open Access and Open Science in the region. Ana submitted a manuscript based on her presentation,

“A tradition of open, academy-owned, and non-profit research infrastructure in Latin America” that appears elsewhere in this issue of *Information Services and Use*. I am sorry that I did not get to hear her speak, but I thoroughly enjoyed her informative paper - worth a read!

4. Indigenous knowledge, standards, and knowledge management

This session (as did a plenary talk last year) highlighted the fact that after many years of being overlooked and marginalized, there is now a growing awareness of the importance of Indigenous Knowledge, and the need for information systems and standards that can support it. Developing these in ways that are respectful of the context - cultural, historical, and more - as well as the ownership of this information, is essential. Numerous conversations about these issues are taking place around the world and the session made it clear that it is time to move from words to action.

4.1. Increasing visibility of Indigenous knowledge from Africa

The first speaker was Joy Owango, Executive Director, Training Centre in Communication (TCC Africa) at the University of Nairobi and a Board Member for AfricArXiv [18], a community-led platform for African scientists of any discipline to present their research findings and connect with other researchers. The goal of her presentation was to demonstrate how the visibility of Indigenous knowledge from Africa is being increased through the collaborative activities of AfricArXiv and TCC Africa [19]. TCC Africa is a fifteen-year-old research trust. It is an award-winning organization that has trained more than ten thousand early-career researchers. They have worked in more than forty African countries and have over eighty institutes. And they have a research mentorship group of nine-hundred plus researchers from two continents who they support in their research lifecycle from idea, through research to publishing.

In October 2021, TCC Africa formalized their partnership with AfricArXiv, with the objective of providing a financial and legal framework for the organization such that it could increase its support and build its community in the continent. In order not to duplicate effort they have partnered with six repository partners - Open Science Framework, ScienceOpen, Figshare, Zenodo, PubPub, and Qeios. African researchers submit their research results through AfricArXiv, and in turn, their output is shared in any of these platforms, thus increasing discoverability. Owango went on to explain how this supports Indigenous knowledge.

She said that Africa has fifty-four countries and more than two-thousand languages. As a result, the African Union mandated the use of integration languages. And the integration languages that somewhat unify regions in Africa are French, Arabic, Swahili and, to a certain degree, English. However, in Africa there are countries in which their Indigenous Language is also their national and business language, and this means that their research output is produced in that Indigenous Language. If a research output cannot be seen in these traditional languages, society is missing the diversity of the output that is being produced by the continent.

As a result, TCC Africa and AfricaArXiv are trying to create an awareness of the language diversity in science. She said that they understand that English is the language of science, but noted that we live in a diverse society, and it is good to recognize and acknowledge the research output that results from the diversity within the society. They know that there is a need to have a common language to connect, and that is why so far English has been used, but this is at the detriment of loss of diversity. She went on to say that while we need a common language to connect, we also need to share our diversity - we need to have a bit of a balance of both, and they have found that technology can help them find the balance. They are

using technology to help increase the visibility of the diversity of the research output coming out of the continent, particularly focusing on research output that is produced in Indigenous Languages. Their efforts have increased the digital discoverability of African research, built scientometrics based on measurable outputs, and by working with established repositories they can provide long-term digital storage. This has been achieved through a partnership with Masakhane [20], a grassroots organization whose mission is to strengthen and spur research in African languages, for Africans, by Africans. Their main objective is to translate Indigenous Language into the various integration African Union languages, which would in time increase the visibility of the research output. This type of partnership has never existed in Africa before, because it was simply assumed that research output that has been done in Indigenous Language would remain at the national or local level. That assumption can no longer exist.

Through the adoption of technology by the partnership of TCC Africa, AfricaArXiv, and Masakhane, Indigenous Languages are translated into the African Union integration language, which helps to increase the visibility of the research output coming out of the continent. AfricaArXiv has been in existence since 2018 and it is also an award-winning organization with more than thirty-six partners. The most exciting thing about AfricaArXiv is that since its inception it had received submissions from thirty-three African countries by the end of 2021. This is the highest submission that any repository has ever had in the continent which is why Owango proudly says that they are a continental platform supporting research results coming out of the continent. In addition, there are more than ten countries outside of Africa who have submitted material - researchers who are doing research on Africa, but who are not based in Africa.

In closing, Owango said that she views this initiative as a major steppingstone. So far there have been more than six hundred submissions and with the AfricaArXiv partnership in building the community, she believes that they can triple or even quadruple the numbers, because it is free for individuals (*Note: institutions have to pay a community fee*).

Owango gave a more detailed presentation on the preprint services provided by AfricaArXiv at the 2021 NISO Plus Conferences and her slides are available on gishare [21].

4.2. *Indigenous knowledge and information systems from a publisher perspective*

The other speaker in the session was Darcy Cullen, the founder of RavenSpace [22] and an acquiring editor and the head of acquisitions department at the University of British Columbia Press (UBC Press), which is a long-standing scholarly publisher of Indigenous studies in the humanities and social sciences. It is located on the Point Grey Campus in Vancouver on the unceded ancestral territory of the Musqueam First Nation.

Cullen's focus was on the work that they are doing at RavenSpace, which was founded at UBC Press with global partners from publishing, Indigenous technology, museums, and libraries, with financial support from the Andrew Mellon Foundation. RavenSpace is a platform and a model of publishing that challenges scholarly communication to develop ways of integrating support for Indigenous voices and authority in the presentation and publication of Indigenous scholarship and Knowledge.

Cullen said that as academia is undertaking changes to address colonial legacies, the structures of knowledge sharing and dissemination need to reflect these changes and empower communities to locate and reclaim their voices in the academic and public record. She said that while there is a robust body of resources for community-engaged research practices, there is little to guide the way the results of that research are then shared and made available for the benefit of Indigenous partners and communities in today's traditional formal modes of publishing. Publishers, librarians, and other professionals in scholarly communications are grappling with questions that are both ethical and methodological about

how Indigenous representations and information are handled in their production, presentation, circulation, and through systems of review, cataloging, discoverability, access, and use. In turn, readers and audiences are seeking trustworthy sources and guidance in how to engage and interact respectfully with Indigenous cultural heritage and knowledge. Cullen said that at UBC Press they asked themselves how publishers can support multiple modes of expression instead of just text. How can they support collaborative authorship and extend the relationships of trust that are formed through the research activities? And how, then, can they make these works widely and openly accessible while respecting Indigenous protocols for heritage and knowledge? For UBC Press RavenSpace was the answer.

The platform supports various media formats and features text, music, digital art, video, maps, animations, interactive mapping, annotations, etc., and can stream content from media-sharing resources around the world. With its built-in and customizable tools, Indigenous creators can assert how their cultural heritage and intellectual property can be accessed and shared. The various contents can be shaped in different ways to appeal to and meet the needs of different audience groups. And Indigenous languages are supported, both on the authoring side and on the audience side, with a keyboard for searching in different orthographies. And with a process for community consultation and co-creation between knowledge holders and authors, as well as artists, media producers, web developers, and others, RavenSpace is seeking to extend the research relationships and productive creation relationships. It also seeks to represent Indigenous worldviews, to highlight the marginalized voices and experiences, and to re-contextualize archival materials on Indigenous peoples and challenge the materials' colonial underpinnings. To do this, they draw on the best practices guidelines of the Association of University Presses for peer review, which she said is the gold standard in academic book publishing. But they have also expanded the definition of peer review to recognize the expertise that resides in communities.

Another facet of their work is around the engagement of their audiences. While direct audience engagement is part of their strategy to bring dynamic works to Indigenous source communities and audiences around the world, part of their work now is also to link up with the other nodes in the research lifecycle, namely with libraries and distributors. For example, they have succeeded in obtaining a cataloging and publication data record for the publication, and a process for doing these functions. More complicated seems to be categorizing the work, which is at once a long form peer-reviewed publication and also a dynamic online resource, and does not fit neatly into existing categories.

They are also finding that metadata is a useful, but complex tool in this work. They are dealing with information about information, information about content, and in this case, invaluable and Indigenous digital co-creations. They have integrated Indigenous knowledge fields in order to power different tools in the publication so that Indigenous authors can provide valuable context and information about their material and where they can assert their rights in cultural property or knowledge and raise awareness with readers about how to interact respectfully with that content.

Cullen went on to say that they hope to include libraries - post-secondary, tribal, and public - as partners in their work to identify ways of bringing these web-based, community-driven publications into the discovery and access systems in ways that will preserve their unique features. There are questions worth exploring between libraries and publishers about the roles that we all play in the presentation, access, and circulation of Indigenous knowledge and doing so in tandem with authors and source communities. She then showed a short video that provided an overview of RavenSpace in a visual form (worth a look because it provides concrete visual and audible examples of their important work) [23].

She said that digital infrastructures carry the bias of all the other systematic frameworks around them. And they found that when Indigenous names, protocols, rules, and permissions are added, there were no

data fields for that information. They had to create new fields and new labels. It will take a long time to correct the missing provenance information, but at least they are creating space for that to happen.

In closing, Cullen said that as they have moved the project into different spaces, they came to understand that researchers and institutions each need their own specific tools. The work of both the researcher and institution is to ensure that Indigenous interests are transparent to other users. And that is where the use of the labels and the ability of the communities themselves to approve those labels and say what they mean to them comes in. She said that one of the most powerful things that they have experienced as a community is the gifts that have come back from their ancestors (*Note: the conference closing keynote referred to and praised this work*).

I very much enjoyed this session. It reminded me of a keynote given at the 2021 NISO Plus by Margaret Sraku-Lartey, Principal Librarian, CSIR-Forestry Research Institute of Ghana. She talked about the importance and value of local Indigenous Knowledge and how it is being threatened in today's modern world rather than being leveraged by the global information community to catalyze development. An article based upon that presentation appears on *Information Services and Use* [24] and is worth a read.

5. Wikidata and knowledge graphs in practice: Using semantic SEO to create discoverable, accessible, machine-readable definitions of the people, places, and services in global information community institutions and organizations

The premise of this session was that the very definition of libraries is static, outdated, and misleading. Why? Because search engines and indexing software agents have limited knowledge of the dynamic nature of libraries - the people who make the library happen, the services that they provide, and the resources that they procure. Yet as a part of the global information community, libraries provide content and education that expands the access and visibility of data and research in support of an informed public.

The speakers were Doralyn Rossmann, Professor, Head of Digital Library Initiatives at Montana State University (MSU); Jason Clark, Professor, Lead for Research Informatics, MSU; and Helen Williams, Metadata Manager, Digital Scholarship and Innovation Group, London School of Economics Library (LSE). Their consensus is that at its core this lack of knowledge about libraries is a metadata problem. LSE and MSU have each recognized the problem, but each has approached the problem in their own unique way. MSU used local structured data in a knowledge graph model and LSE used “inside-out” definitions in Semantic Web endpoints such as Wikidata [25]. I admit that I did not know anything about Wikidata and if you are like me, I found the introduction [26] to it very helpful.

MSU Library ultimately found that implementing a “Knowledge Graph” linked data model within HTML markup leads to improved discovery and interpretation by the bots and search engines that index and describe what libraries are, what they do, and their scholarly content. In contrast, LSE Library has found that contributing to a collaborative and global metadata source, such as Wikidata, is a means to extend reach and engagement with libraries and how they are understood.

The technical details of their efforts were explained during the session: they demonstrated how Wikidata can be used as a tool to push out data beyond organizational silos; the technical details of knowledge graph markup and semantic Search Engine Optimization (SEO) were discussed; they worked through questions about how metadata can represent an institution/organization equitably; and ultimately explained how this work improves the accessibility and reach of global information communities.

In the end, both approaches had positive results. As noted earlier MSU found that implementing a “Knowledge Graph” linked data model leads to improved discovery and interpretation by the bots and search engines that index libraries and their content. A comparison of users and sessions in Google

Analytics, pre- and post- optimization showed a 20% increase in users, including a 29% increase in new users, and an 8.6% increase in sessions. Overall traffic patterns from search engines showed growth as well - a 10% increase in organic search result referrals from Google, and a 34% increase in organic search result referrals from Bing.

In contrast, LSE found that contributing to Wikidata, a collaborative and global metadata source, can increase understanding of libraries and extend their reach and engagement. Indeed, LSE did an analysis of eighty titles that they added to Wikidata and found that downloads in the six months after adding titles to Wikidata were on average 47% higher than in the six months before their addition to Wikidata.

Rossmann, Clark, and Williams have written an article based upon their presentation that appears elsewhere in this issue of *Information Services and Use*. I highly recommend that you read it.

6. Accessibility in the scholarly information space

Accessibility of information resources is a topic that is broadly understood, but details of the barriers to accessibility are sometimes maddeningly complicated. This session focused on some of the less-discussed problems in making published information accessible, e.g., license restrictions, and offered possible solutions to these issues.

6.1. Perspectives from a blind disability rights scholar

The first speaker was Anna Lawson, a Professor of Law and Joint Director of the Centre for Disability Studies [27] at the University of Leeds in the UK. Lawson described her experience in college when she was in the process of losing her sight and reading print had become almost impossible for her. At that stage, she accessed all of her studying materials through cassette tape. The Transcription Centre at Leeds organized volunteers to record material for her and for other visually-impaired and print-impaired students. She said that this was transformational to her learning experience and had a positive impact on her grades. After doing graduate work elsewhere she returned to Leeds as a member of staff and all of her reading today is done electronically via a screen-reader. The Transcription Centre still transcribes books for her, but she also has an assistant who helps her locate other electronic recordings, which are already available in accessible formats.

Lawson went on to talk about a new journal, the *International Journal of Disability and Social Justice* [28], which, along with her co-director of the Centre for Disability Studies, she co-founded as an editor. Launched in December 2021, the goal of the journal is to tackle the barriers that they know exist to opening up access and inclusion to ensure inclusion. These challenges include the following:

- Tackling financial barriers - for readers and for authors
- Communicating content beyond academia - with plain English summaries, videos
- Writing style - e.g., academic versus plain English
- Providing an alternative electronic format (PDF and HTML)
- Integrating accessibility into general systems and processes
- Inviting individual requests for support and assistance

She then went on to talk about the problems that she faces when attempting to access and use information and these are related to PDFs, e-books, e-journals, and behind-the-scenes scholarly spaces. PDFs do not scan well, and this often results in messy documents that are difficult to navigate. E-books are not as readily-available as one would assume, and they are not easy to read, especially if they have images that

are not described - also annotating and making notes is difficult. With e-journals the problems are with page numbers when she needs to cite something. Page numbers are not always made apparent in a text-based way. Her “behind-the-scenes” obstacle relates to the fact that she believes that there is often an assumption, not only by publishers, but also by many people in academia, that accessibility is important to students, but that staff in different capacities will not be disabled. So, although some things that are geared at students might be accessible, if there are things that only staff are expected to access, such as grant reviewing systems or online marking systems, accessibility is often just not there.

In closing, Lawson said that when we think about Open Access it is important to remember to whom we are opening up access and that must include disabled people. It is also important to remember that different disabled people will have different access needs and preferences not only depending on their impairment type, but also depending on their disciplinary expertise. It is important to integrate and embed accessibility at all phases and all stages, from procurement right to the final end; to envisage the people who need accessibility not only as being students, but also as being disabled staff in all capacities; and finally, we are never going to get it completely right, so we need to make sure that there are always options for people to ask questions, to let us know of problems, and request help and assistance to access the information that they need.

6.2. *Providing access to instructional and reference materials in post-secondary education*

The second speaker in the session was Jamie Axelrod, Director, Disability Resources, Northern Arizona University. Axelrod basically echoed all of Anna Lawson’s comments, for example - not all materials are commercially-available in an accessible format. Also, there is an assumption that because something is in an electronic format, that it is accessible. And that is simply not the case. The same type of formatting requirements will exist in an e-book that are necessary in an accessible version of what otherwise might be print material or an electronic file. “Electronic” does not mean “accessible”. What makes materials accessible are really important - things such as the markup of headers and other formatting features, alternative text to describe images and graphs, notification of page breaks, marginalia, etc. - these are details that we do not always think about when we are creating an accessible version or access to the primary text. The accessible file must replicate the structure of the document as well as the content of the document itself.

Like Anna before him, Axelrod said that his number one problem is with PDF-based documents. And one thing that he often finds, as he and his team are preparing materials for their students, faculty, and staff, is that they will receive electronic versions of materials as one large PDF. It is not broken up into any way in which someone can navigate around it quickly and easily. And they must take that very large PDF and break it up into chapters or sections to make it more usable for the person who needs it. The document is oftentimes not appropriately structured or tagged, so it may only be readable by someone who is a skilled screen-reader user. When a document is not appropriately structured or tagged, for those who do not have a lot of technical know-how, it can be very difficult to navigate. And while it may be readable, it is very difficult to move around without those features.

As Anna noted, pagination is also a problem. Page numbers can be missing. Or what often happens in those large PDF files if books that he receives from a publisher, page one is the cover because this is a file. And therefore, the page numbers do not match up to the text itself. And when a student is being told that they need to read a certain page range this is problematic. He went on to mention other issues such as in STEM fields the symbols, e.g., mathematical symbols, Greek symbols, section symbols for the law, etc., are often not represented appropriately and as a result cannot be read and accessed appropriately.

Axelrod brought up another serious issue. He noted that once his team has remediated a document to be usable by a visually- or otherwise-impaired student - a not insignificant task - it cannot always be shared. The U.S. copyright law through the Chafee Amendment permits the distribution of remediated accessible files to qualified users. And he believes that the Marrakesh Treaty [29] does that internationally as well. However, when institutions such as his receive files directly from a publisher or a third-party service, the contracts to which they need to agree to in order to receive those files prohibit them from sharing any of those remediated files with other institutions who need them for their disabled students. And in fact, Axelrod said that prohibition against sharing requires him to request the same book again if another student needs it even in the same year or term or potentially the next term. The contract requires him to make a new request and not use the previously-provided file until express permission is provided. If he wants to request material from a publishing house or a third party, it can take some time before the material is received - weeks or even months - far too long for him to be able to provide the material to the student in a timely manner. Often, he needs to re-purchase the same materials.

In closing, Axelrod showed a photo of their Alternative Format Production Office where there are multiple desks, multiple workstations, and where a crew of student workers continuously and regularly work to make materials usable by disabled students which once it is completed cannot be shared. He went on to say that most universities around the United States have a similar production facility doing the same work and, as we all heard the prior speaker, they are doing the same work at her institution. He said that in post-secondary education, much of the materials are the same from institution to institution. Not all, but many. So why we are all repeatedly doing the same work is a real conundrum and question for us.

6.3. *FRAME: Federating repositories of accessible materials for higher ed: A new collaborative framework*

The speaker in this session, Bill Kasdorf, Founder of Kasdorf and Associates, a consulting firm, segued nicely from Axelrod's presentation because he talked about an initiative that will hopefully resolve the duplication of effort around the world to make the same materials accessible to and usable by disabled students. Kasdorf described the multi-institutional project at the University of Virginia that was established to address the problem described by the prior speaker. With a two-year grant from The Andrew W. Mellon Foundation, the University initiated an effort to create a web-based infrastructure allowing Disability Service Officers to share remediated texts, to reduce the nationwide duplication of effort, and thereby make it possible for the staff in these offices to achieve better outcomes for students in higher education. The initiative is known as FRAME: Federating Repositories of Accessible Materials for Higher Education. And its mission is to eliminate as much as possible of wasteful, redundant work by enabling remediated resources to be discovered and shared between responsible parties.

Kasdorf said that The FRAME project was never intended to benefit only the seven participating universities. Once Phase Two has been completed thanks to the Mellon grant, they need to have a membership model in place. Their goal is to establish a dues structure that recognizes the differences in resources between schools of different sizes, as well as the contributions of remediated content that they provide. Their hope is that FRAME and the repository being built to store the remediated material can be self-sustaining with a dues structure that comes close to being offset by the savings a participating university realizes by having access to already-remediated resources. Ideally, their goal is for members to come out ahead in the long run as the number of participating institutions increases and the corpus of remediated content grows.

Kasdorf 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 if you want to learn the details of this worthy project.

6.4. Accessibility in the scholarly information space: A technical perspective

The final speaker in this session was Oliver Rickard, Product Director for HighWire Press. Rickard said that it is very important to HighWire Press that content is delivered in an accessible way as possible. From a technical perspective, the key guideline is the Web Content Accessibility Guidelines (WCAG) [30]. Version 1.0 was adopted in 1990, 2.0 in 2008, 2.1 in 2018 and version 2.2 is expected in December 2022. Over the years laws regarding accessibility have changed, hence the constant evolution of the Guidelines. He said that basically the Guidelines is a list of success criteria that online sites strive to meet. And every one of these success criteria is mapped to a certain conformance level A, AA, or AAA, with AA being the highest. So, if your accessibility is absolutely perfect you will have hit AAA. He said that in reality it is extremely difficult to reach AAA. And fortunately, most of the laws understand that it is difficult, and he has yet to run across a law that says you must meet the AAA level.

To demonstrate the differences between the levels he provided some criteria that are related to color. Level A one states that color cannot be the only means of conveying information. The AA criteria is concerned about the contrast of color, so it provides a contrast ratio that needs to be hit. The AAA one is an even higher contrast ratio. So, these are the kind of things that are included in the WCAG standard. Pickard said that as a person responsible for website development, he cares about the whole world and that means that there are a lot of different laws and countries that he has to consider. Fortunately, most of the laws tie back to WCAG 2.1, so as long as you are meeting WCAG 2.1 AA level, he said that “you are good to go”.

Another consideration that has emerged is something called the “VPAT [31]”. The “V” stands for “voluntary”, so this is yet another step. It is not required legally, but many people want and expect it, which means that he must conform. This is a document that he fills out that explains exactly how his client’s website meets whatever laws or guidelines with which they are trying to comply. His completion of the document allows his publisher clients to deliver the template to their customers and universities, etc., so that they can prove that they are meeting the guidelines and laws with which they need to comply. These are the challenges on which all technology providers must focus.

Unfortunately, standards evolve and laws and technology change. He cannot make everything accessible and then walk away. This is a continuous process, especially now that it is normal that, as for any website, that he and his staff do not build all the code. They use fantastic tools developed by other people, and those tools evolve as well. They can never assume that their site remains accessible just because it was on day one. Another challenge is that accessible versions of features can be more expensive. Financial decisions also need to be made so in reality decisions on features are made all the time with money as a factor. And the final one is that it is really hard to achieve full compliance.

Rickard then went on to describe the process that they go through to test a site. It sounds like a nightmare! They always do accessibility testing while building a website, not just at the end. Everything is tested against WCAG 2.1, level AA. They use a certain amount of automated testing, but even that has challenges. He provided an example that is one of the WCAG guidelines and it is about non-text content. The criterion is that all non-text content that is presented to the user must have text alternative that serves the equivalent purpose. So non-text content, such as an image, must have alt-text so that for someone who is using a screen-reader, the device will read out that alt text instead of showing the image. He showed a picture of

a nurse taking someone's pulse and the alt-text said just that - "nurse taking a person's pulse". He then went on to talk about the problems that this causes. They can readily fix problems when it is their own content with which they are dealing. But when they need to fix problems with content that they receive from other publishers it gets much more difficult (listen to the video and sympathize!).

In closing, Rickards said that despite automated tests you must do manual tests afterwards. And to summarize everything, if you are worrying about the law and the guidelines, the thing to focus on today is WCAG 2.1, level AA. As he already said, if you are meeting that, you are good to go. In addition, make sure accessibility is part of your regular testing processes. Whether you are making a site for yourself or someone else is making one for you, you need to make sure that that accessibility testing is ongoing all the time. And then finally, use manual testing on top of the automated testing as much as possible to ensure that the site is as accessible as possible.

7. Archiving and digital preservation

The morning of the second day of the conference opened with four parallel sessions, one of which was on archiving and digital preservation. This session looked at the challenges and opportunities of archiving and digital preservation, including some recent efforts to bring the community together to address them.

7.1. Applied digital preservation and risk assessment

The first speaker was Leslie Johnston, Director of Digital Preservation for the U.S. National Archives and Records Administration (NARA). She opened by saying that the discipline of digital preservation covers any digital object, whether born-digital or digitized. Digital preservation encompasses all format types: texts and images, databases and spreadsheets, vector or raster images, software, email and social media, games, movies, music and sound, and the web. With every new Information Technology innovation, digital preservation managers must respond by devising effective strategies for ensuring the durability and ongoing accessibility and usability of new digital materials, so digital preservation will remain an always-emerging challenge. Johnston noted that there are several digital preservation standards and models and, although listed on her slide along with the url's for access, she did not go into them (*note*: her slides are available in the NISO figshare repository) [32]. She asked a rhetorical question - Which one is "The Best?" and her response was none of them, all of Them, and the standard NARA response - it depends on your goals for assessing your digital preservation program and the risks related to your collections.

She then proceeded to talk about the applied work at NARA. The first step was to develop a guiding policy/strategy which they initially did in June of 2017 as a guide to its internal operations and it is publicly-available [33]. It outlines the specific strategies that NARA planned to use in its digital preservation efforts, and specifically addresses: infrastructure, format and media sustainability and standards, data integrity, and information security. The policy applies to born-digital agency electronic records, digitized records from agencies, and NARA digitization for access and preservation reformatting.

Johnston then went on to talk about the file format risks at NARA. She noted that an integral part of NARA's work is the issuance of guidance on all aspects of Federal electronic records management and transfer to NARA, including media types, file formats, and metadata [34]. By regulation, NARA cannot be 100% proscriptive in the formats it accepts. When records are transferred, they are validated to ensure that they are uncorrupted, and, if possible, meet NARA's format guidance. There are "Preferred" and "Acceptable" formats, but sometimes NARA must take in records in the format that the agencies have because those are the tools and formats that they use to do their jobs, and there must always be exceptions.

The work of the diverse federal agencies can be hugely different. NARA created a collections format profile so that they could see what types of material they have (email messages is the largest proportion of content at more than seven hundred and seventy-six million electronic mail messages!) and she highly recommends that every organization undertake such an effort. She said that the government memorandum known as M-19-21 [35] has stated that after the end of 2022, NARA will no longer accept physical transfers. They will only accept material that is either born-digital or has been digitized, as well as their own digitization for preservation reformatting and access of records that have come to them.

Johnston went on say that the creation of the format profile of their holdings was critical for undertaking the risk assessment. In 2018 NARA created an extensive risk assessment matrix, designed to apply a series of weighted factors related to the preservation sustainability of the file formats in the collection. Each question has a relative weighting that maps to the level of risk for each question and, to the extent that it can be defined, resource costs (staff time or budget). The matrix also includes high level factors that assess the preservation actions that could be taken vis-à-vis their current environment and capabilities. The Matrix calculates numeric scores, which are mapped to high, moderate, and low risk and the risk thresholds are subject to review and revision over time. She added that risk assessment is not enough - it must be translated into actionable plans to mitigate the risks. Their plans identify essential characteristics for electronic records held by NARA, document file format risk, and collate links to specifications and other digital preservation resources. The recommended preservation tools and actions for formats included in the plans are based on current NARA decisions and capabilities. The plans consist of two sets of documents: (1) record type plans which document the characteristics of different categories of records, e.g., audio, video, navigation charts, e-mail, etc. and (2) preservation action plans that is a single spreadsheet containing almost seven hundred file formats across all record types. NARA released these plans in 2020 to the public. Their digital preservation framework, which is their risk matrix, the record type, and the Preservation Action plans are all available on GitHub [36], and these are updated quarterly.

In closing, Johnston noted that in 2021 and 2019, NARA completed a self-assessment of their programs and systems using the PTAB (Primary Trustworthy Digital Repository Authorization Body) instrument based on ISO 16363:2012 [37]. She acknowledged that they had gaps in their documentation, but when they did their second self-assessment things had improved and she recommended that organization regularly undertake self-assessments. In fact, at NARA they intend to do such assessments every two years.

Johnston gave a similar talk at the NISO Plus 2020 conference and published a paper in *Information Services and Use* which you might find interesting [38].

7.2. A model preservation policy for digital publishers and preservers

The second speaker in this session was Alicia Wise, the Executive Director of CLOCKSS. This is a collaborative effort of the world's leading academic publishers and research libraries to provide a sustainable dark archive to ensure the long-term survival of digital scholarly content [39]. Wise opened by saying that digital preservation is not just a technology issue, even though we tend to talk about it that way. It is a commitment of resources over the long term - time, attention, and active management. It is actually a series of decisions. A preservation policy documents those commitments and decisions, helps you make deliberate, transparent choices about what aspects of a publication is being preserved and how and, just as importantly, what aspects of a publication is *not* being preserved. A preservation policy makes choices explicit, not implicit or, perhaps, unintended. A preservation policy is a long answer to two short

questions: (1) if I publish with you or cite this work, what will a future reader see? and (2) how close will that be to what I see now?

Wise went on to say that we probably cannot preserve everything - echoing the prior speaker. To do so would be expensive, and resources are finite. But a policy forces you to think about the following:

- What does a future user really need?
- Which version do we need to preserve? Want to preserve?
- How important is presentation?
- Does the content exist independently of the software through which it is delivered?
- How do we keep metadata associated with the content?
- What partners, skills, and resources do we need now and for the long term?
- How can libraries encourage authors and publishers?

She said that the NASIG [40] Digital Preservation Committee conducted a survey in 2018, the results [41] of which showed that many organizations in scholarly communications lack policies for preservation. Out of that survey came a recommendation to develop a model policy or template as a resource for the development of a preservation policy, so the committee set up a model policy working group in August of 2020 of which Wise ultimately became a member. She talked briefly about the process that the working group used in developing the policy which, at the time of the 2022 NISO Plus Conference, was still a work-in-progress. She said that they envisage that libraries and publishers will use the policy to get started in digital preservation to really document their organization's mandate and how their commitment scope and goals for digital preservation support that organization's mandate. The policy will not go into much detail. There will be additional need for platform level or repository level policies and procedures. She cautioned that it is not possible for the model policy to be "one size fits all" because context, resources, and content/collections vary from institution to institution and therefore policy approaches will differ as well.

In closing, she noted that the policy was to be officially launched at NASIG's 37th Annual hybrid Conference that was scheduled to take place in June of this year. I can add that it was announced and that the NASIG Model Digital Preservation Policy [42] is now publicly-available.

7.3. *COPIM WP7: Archiving and digital preservation*

Then final speaker in this session was Miranda Barnes, Research Assistant, Archiving and Preserving Open Access Books, Loughborough University, UK. She said that COPIM [43] (Community-led Open Publication Infrastructures for Monographs) is an international partnership of researchers, universities, librarians, open access book publishers, and infrastructure providers who collaborate to build community-owned, open systems and infrastructures to enable Open Access book publishing to flourish. They have seven work packages and COPIM WP7 [44] is focused on the archiving and preservation of Open Access monographs.

Barnes said that some of the questions being addressed are what are the boundaries of a book? how are complex digital monographs preserved? and what are the risks to archived content? In terms of the policy landscape, it must be said that while COPIM is an international project, Barnes was framing her comments within the UK context, because the UKRI (UK Research and Innovation) [45] has announced a new Open Access policy for all their funded projects. That policy now includes monographs, so any monographs published on or after January 1, 2024 must be made Open Access within twelve months of publication. This is very influential, and it is expected that the U.K. Research Excellence Framework

(REF) may follow with their own Open Access policy that will require a similar Open Access mandate for monographs. This will include digital monographs that are born-digital and cannot be easily replicated in print form. Such digital monographs may interact with different parts of the content on other parts of the internet, they may link to content on a video platform, they may have embedded content within them. They are beyond the traditional idea of what a book is, but they are an important part of digital humanities or creative practice works where these boundaries are being pushed. They may include embedded content, e.g., images, videos, audio files, geospatial data, 3D models, QR codes etc. She said that we need to identify the boundaries for preservation. Whose responsibility is it to preserve this content? how is it preserved? where is it preserved? This is particularly an issue for the smaller scholarly publishers who do not have a supporting organization to archive their content and we are looking at these issues from their perspective even though these questions impact all publishers. It is certainly not as complicated as what Leslie Johnston described earlier.

In terms of the formats that publishers tend to use, the primary one is PDF, and it is the one most widely-adopted. However, it closely mimics the traditional book format and has serious negative implications for embedded files (It seems that most of the speakers on the first day have problems with PDFs!). In many cases, PDF is *not* recognized as a container by the software that scans files in archives. While an XML or a zip file will be recognized as a container and all the files within it will be scanned, the PDF does not have that benefit. Also, there are many different PDF formats (Johnston said that NARA has sixteen). XML is often preferred by key preservation players. It integrates data, metadata, and infrastructure. It is more ideal for long-term preservation than PDF, and relationships between content can be defined, but the problem is that many small publishers simply do not have the technology, resources, or staff resource to convert monographs into XML on a regular basis.

Barnes said that in terms of what they have learned so far key issues are related to file format. How they are packaged and how they are preserved will directly influence each other. And the fact that PDFs are widely-used presents complications to preservation that they plan on digging into much more thoroughly, particularly the differences in functionality between the different PDF formats.

In closing, she asked what are the longer-term solutions? What can the community build for all levels of monograph publishers because this is a part of the scholarly record that absolutely must be preserved for future scholars and future publishers.

8. Artificial intelligence and Machine Learning: Less theory and more practice

Artificial Intelligence and Machine Learning systems have been heavily discussed, but what is actually being done using them? The speakers in this session will present a variety of practical applications with which they are currently involved.

8.1. AI and Machine Learning: Less theory and more practice - a few examples

The first speaker in this session was Andromeda Yelton, an adjunct faculty member at the San Jose State University School of Information. Basically, she gave almost the same presentation that she gave at NISO Plus 2021, First she talked about Teenie Week of Play, so-called because it was a week of investigating various computational approaches to the Charles Teenie Harris Archive. He was a famous photographer for The Pittsburgh Courier. The projects used AI to automatically shorten titles, extract locations and personal names, and look for the same person in different photos across the collection [46].

Yelton again spoke about Transkribus [47], a comprehensive platform for the automated recognition, transcription, and searching of historical documents. It is a project of READ-COOP [48], which is a European Union (EU) cooperative society organized for social benefit rather than for profit. It is open to non-EU members as well and it has been used in the Amsterdam City archives and the Finnish archives. It supports Arabic, English, old German, Polish, Hebrew, Bangla, and Dutch, and it can be trained to handle more languages and scripts. There is a demo on the READ-COOP website along with an option to download the software and install it to have a full set of features. Yelton said that as with the Teenie Collection, this can help accelerate the human labor that is involved in archiving.

She again spoke about a project that she herself created entitled “Hamlet [49]” that she used to explore a database of graduate theses. The algorithm was trained on a corpus of about forty-three thousand Master’s and PhD theses, mostly from science, technology, and engineering-type subjects. There was some metadata associated with it, but not much. It did not have subject access, nor did it have full-text search capabilities. While one could look at all the documents that came from the same department and hope that they had something in common, the fact is that there can be thousands of theses from the same department that do not have much in common. Two documents can be in the same department and have very unrelated topics. For example, the Department of Electrical Engineering and Computer Science has some electrical engineering theses that could be basically mechanical engineering, and computer science theses that could be basically math. This problem was interesting to her because if you are a graduate student, you probably know about research that your lab group has done recently, but you do not necessarily know about research that was done a couple of decades ago or was done in other departments. The ability to find things that are related to your work, but are not co-located by the metadata that is available, is very useful. And, because, under the hood, Hamlet is putting things closer together or further apart based on similarity, it lends itself to a visual display of information. But Hamlet did, in fact, find interesting ways to group documents that are more or less similar to one another.

I suggest that you read my Overview of the 2021 Conference - there was not much new here, although Yelton is brilliant [50].

8.2. Why explainable AI is sometimes not explainable

Barry Bealer, the Chief Revenue Officer of Access Innovations, Inc. was the second speaker. He opened with a rhetorical question - what is Artificial Intelligence (AI)? He noted that there are a lot of words that describe it so it can be very confusing when talking to people. Is it artificial intelligence? Is it assisted intelligence? Is it advanced algorithms? You need to know the context of AI as seen by the person with whom you are having the conversation. Gartner has hype cycles for pretty much everything, and within the AI umbrella they track thirty-four different technologies [51]. He noted that AI is evolving, and as more and more people invest in it, it will keep on improving and there will be more clarity around how it can help your organization.

He then showed a slide (his slides are on the NISO figshare site) from an MIT class came that asks fourteen different questions. And as you go down through the tree structure list, it attempts to determine if you are actually using the technology. Bealer then said that the point of the slide is that it is important to ask questions when you start implementing or evaluating AI technology. Ask a lot of questions to figure out what is going to be explainable to you in your environment. Why? Bealer then showed a slide that listed eight definitions of AI, and said that this is leading to frustration on the industry. In many cases organizations are talking past each other. He noted that we do not even have a common definition of what

Artificial Intelligence and Machine Learning (ML) are. He then displayed the definition of AI that is in the Oxford Dictionary:

“Artificial Intelligence (AI) is the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages”.

He agreed that this is a straight-forward definition and went on to display the standard Venn diagram [52] of AI/ML/Deep Learning with which most of us are familiar and said that from a simplistic standpoint, these are the three main areas that are discussed when talking about AI and these seem to resonate with everyone. He referred to a White Paper that was published in 2022, *The Future Impact of Artificial Intelligence on the Publishing Industry* [53]. The first thing that jumped out at him when he was reading it was the following statement: “As discussions on AI increase so does the hype and therefore the confusion surrounding it”. And he went on to say, that is the problem today. We just make up new phrases every day about what AI truly is.

The fact of the matter is that AI has penetrated the entire publishing life cycle - from author, content creation, whether it is in word or some other form, to editorial, to production, to aggregation and distribution, to the end-user experience. The above-mentioned White Paper makes it clear that AI can be applied anywhere in publishing and even predicts the departments where the benefits will be most visible. But the important question to answer is where it can be most beneficially applied in your organization. (I scanned the White Paper. It looks like it is worth a read).

Bealer said that he did an informal survey of half a dozen executives in the publishing industry, three from very large global publishers and three from smaller publishers, just to get a sense of where they thought things are going with respect to the impact of AI. In 2020, for both large and small organizations, their comments were all about fiscal responsibility and making sure that they were surviving because of uncertainty about how long the pandemic was going to last. By the end of 2021 though, many of them were seeing an uptick in revenue because everything was moving online or more people were subscribing to databases. The larger publishers told him that they are starting to invest in AI-based technologies and already testing things out or implementing them in their production cycle. The smaller publishers, however, are still sitting on the sidelines and, for the most part, are not investing in AI-based technologies. Which, he said, is unfortunate because early adoption, while not necessarily for everyone, can help you prove out specific use cases.

In closing, Bealer said that he believes that AI is explainable, but that it is complicated, and that you need to know your audience. Also, if your organization is using AI, you need to verify and measure the results.

8.3. *A knowledge discovery system integrating knowledge organization & machine learning*

The final speaker in this session was Jiao Li from the Agricultural Information Institute of the Chinese Academy of Agricultural Sciences (CAAS). She opened by saying that Elsevier’s 2019 report, *Trust in Research* [54], stated that on average, researchers spend just over four hours a week searching for research articles, and more than five hours reading them. The rate is five to six articles per week, and half are considered useful. In addition, compared to year 2011, researchers read 10% less literature, but spent 11% more time finding the articles. She asked how we can help researchers with search and discovery and went on to describe the semantic search system that is under development by her organization. She reviewed the technological underpinnings that are being put in place with the objectives of this new system:

- Establishing connections between users and the right resources.
- Providing high-quality data and semantic knowledge models for Machine Learning.
- Using text mining technology to extract information/knowledge, identify concepts, entities, and relationships, with a special focus on key semantic relationships, and exploring computational causal reasoning.
- Offering intelligent Q&A-based on collaborative reasoning and knowledge computing.
- Providing human-computer interaction that is conducive to providing users with accurate literature/-knowledge retrieval and realizing contextual retrieval, cognitive search, etc.

If you are interested in the technical design, I suggest that you watch the video - no way that I can do it justice. I also recommend that you take a look at the Elsevier report - very interesting.

Note: The moderator for this session was Clifford B. Anderson, Chief Digital Strategist, Vanderbilt University Libraries, who did not give a presentation, but spoke to me about “Data Lakes” - a term which I had never hear before. He told me that a data lake is a platform for delivering big, heterogenous datasets to library patrons. A data lake gathers datasets across an organization into a single location, allowing latent connections between datasets to be discovered dynamically. When deployed with appropriate metadata and access controls, data lakes become effective ways for libraries to utilize the text and data mining rights they have licensed for proprietary datasets. I asked him to submit a paper on the topic and it appears elsewhere in this edition of *Information Services and Use*. If you are unfamiliar with the topic it is well worth a read.

9. 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 [55].

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). And this year the award was presented to Dr. Patricia Flatley Brennan, Director of the U.S. National Library of Medicine (NLM).

Dr. Brennan is the first nurse, industrial engineer, and woman to be the NLM director, and since joining NLM in 2016, she has positioned the organization as a global scientific research library with visible and accessible pathways to research and information that is universally actionable, meaningful, understandable, and useful. Prior to joining NLM, Dr. Brennan was the Lillian L. Moehlman Bascom Professor, School of Nursing and College of Engineering, at the University of Wisconsin–Madison. Brennan received a Master of Science in nursing from the University of Pennsylvania and a Ph.D. in industrial engineering from the University of Wisconsin–Madison. She served as chair of the University of Wisconsin–Madison College of Engineering’s Department of Industrial Engineering from 2007 to 2010. Dr. Brennan has received numerous accolades recognizing her contributions to her field. In 2020, she was inducted into the American Institute for Medical and Biological Engineering (AIMBE) College of Fellows. She is

also a member of the National Academy of Medicine and a fellow of the American Academy of Nursing, the American College of Medical Informatics, and the New York Academy of Medicine.

In her presentation, Dr. Brennan noted that NLM has been serving science and society since 1836. They are a research enterprise for biomedical informatics, and the world's largest biomedical library. From 1836 to 1968, NLM was a collection of books and journals, beginning as a shelf in a field surgeon's tent and now have moved on through congressional legislation, to the Public Health Service. NLM became part of the Department of Health and Human Services around 1966, and in 2000 NLM began the beginning of a 21st-century library. Dr. Brennan's lecture answered the following questions: How does a modern library meet its mission to acquire, organize, preserve, and disseminate the many outputs of contemporary science? What role do standards play? How does NLM help a diverse set of stakeholders make meaning from their resources? Overall, she described NLM's history, detailed current strategies, and discussed the future where partnerships form the key to advancing NLM's mission.

Dr. Brennan submitted an article based upon her presentation and it appears elsewhere in the issue of *Information Services and Use*.

10. Visions of the future

This session had six speakers giving very brief presentations on supposedly visions of the future. The first talk is a must-read, but more importantly a must-look at the video as I cannot replicate in words what I saw.

10.1. Visual-Meta

The main speaker in this session was Frode Hegland, Director, the Augmented Text Company supported by Vint Cerf, Vice President and Chief Internet Evangelist for Google (yes, *that* Vint Cerf!).

Hegland spoke about Visual-Meta, a new service from his company that 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. (To me it appears to be a very good answer to the problems with PDFs that several prior speakers highlighted).

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 moved it to the back of the document. He put on screen the proceedings of the Association of Computing Machinery (ACM) Hypertext Conference from last year. At the end of each document, there is Visual-Meta. The formatting is inspired by BibTeX [56]. He said that not everyone knows what BibTeX is, but it is an academic format, part of LaTeX [57]. 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.

The ACM Hypertext example uses the mandatory Visual-Meta. There is also an optional Visual-Meta that provides even more information. This includes document type and references. In a readable way, it provides everything that the document cites. Document headings provide the actual structure of the document. You can have a glossary and endnotes. He then visually collapsed the material and showed an appendix at the end of the document along with a “pre-appendix” which has content that verifies the importance of the document. It can also have three types of appendices afterwards. One is errata which he described as an old-fashioned piece of paper coming in saying there are errors in the document, which is a useful part of its history. Then you have another kind of history which is who has read the document and what have they done with it? That information could be part of an intelligence workflow of who has approved the document. And finally, there is information about augmented views, which he discussed later. The video [58] was fantastic. I cannot describe and do it justice - you need to see it.

The demo paused and Vint Cerf said that he wanted to talk about the implications of what Hegland has been able to do so far. The first observation he made was that by making Visual-Meta simply text at the end of the PDF, they have 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. 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.

Hegland said that one of the things that he is very excited about is how Visual-Meta can help Virtual Reality (VR), and this links back to the opening keynote about the potential importance of VR. When a researcher enters a workroom now, they can easily bring their laptops or their computer screen with them. But what they cannot do is take what is on their screen and lift it into the space to share it. It can be shared on a flat panel only. He went on to ask the audience to imagine a knowledge graph richly multidimensional being shared and manipulated. That can be done now - his group is experimenting with this using Visual-Meta because they know the headings and the structure of the document, and via the extra appendices there are the augmented views. All locations of every item are recorded, and all the attributes, etc. can be put into a new Visual-Meta at the back of the document so that when it is opened in a VR space, all of the information will be put back exactly where it was in the original document (mind-blowing!)

He said that his team uses the Oculus headset. They are not big fans of Meta (Facebook) because privacy is an issue, but he assumes that in about a year, Apple will release its own system. And, because Apple is good with consumer products, everyone within the next few years will have Apple VR headsets. But Apple is no angel either. All of these Big-Tech companies will try to own the whole space. Hegland believes that it is up to us as a community to make sure that we can own the data that goes in and out of VR (again, echoes of the opening keynote).

Vint Cerf added that there is another point to be made here apart from owning data, and that is regarding the interoperability of the various systems. Encouraging the producers of these products that interoperability is very important is not going to be easy to achieve. It was achieved almost by accident with electronic mail for the most part, but it never worked with instant messaging. That turned out to

be walled garden products, and the operators of those applications did not agree to interconnect in an actionable way. Cerf believes that as a community we have a lot of work to do to persuade the makers of the products and the servers of the products to conclude that interoperability among their various systems is more valuable than isolation.

In closing, Hegland added that his company's approach here is completely legacy safe, it does not add any data to a PDF other than text. And it can be printed, scanned, OCR'd, and nothing is lost. He believes that they have delivered on their goal of enabling rich interactions, flexible views, and easy citing for published documents in a robust way. And they are now looking to expand who uses the system and to build with the community in a completely open way so that the project can go where users want it to go. There is some interesting additional reading material if you want to learn more [59].

This was, for me, one of the most exciting presentations of the entire conference!

10.2. *The state of open data*

The second speaker was Ana Van Gulick, Government and Funder Lead at Figshare, who gave a summary of their state of open data survey from 2021. She said that they have been conducting the state of open data survey in collaboration with Springer Nature for six years. Over that period, they have had more than twenty-one thousand respondents from one hundred and ninety-two countries. And this has provided them with a great way to take a sustained look at the state of open data, data sharing, open science, and over that time period.

The 2021 survey was conducted in the summer of 2021, and they had nearly forty-five hundred responses that were analyzed as part of the results. The respondents came from all over the globe, with Europe, Asia, and North America represented most heavily, but also from the southern hemisphere. The fields of interest from the respondents included the biomedical sciences as well as many other scientific fields such as the applied sciences and physical and earth sciences as well as the humanities and social sciences and other fields. A majority of respondents were late-career researchers. However, nearly 30% were also early career researchers. She said that it is worth noting with a survey like this about open data that the survey population is probably likely to skew towards those already interested in open data, practicing data sharing, and being advocates of open science. However, this still gives them the best glimpse at a large population and what is happening with these practices.

Van Gulik said that an important thing to note in the 2001 survey results is the impact of the COVID-19 pandemic. It highlighted the importance of open data and open science and having scientific results shared quickly around the globe. Also, it showed that scientific results that have not been officially published and appear in preprints and in open data sets are reaching the mainstream media. About one-third of the respondents indicated that they had reused their own or someone else's openly-accessible data more during the pandemic than before.

She then went through some of the key take aways from the survey.

- There is more concern about sharing data than ever before, with four percent of the respondents saying that they have no desire to share their data.
- There is more familiarity and compliance with the FAIR [60] data principles than ever before.
- Repositories, publishers, and institutional libraries have a key role to play in helping make data openly available (35% rely on repositories, 34% rely on publishers, and 30% rely on institutional libraries).
- 30% of survey respondents said they would rely upon their institutional library for help making their research data openly available; almost half of respondents share their research data in an institutional

repository for a public audience; 58% of respondents would like greater guidance from their institution on how to comply with their data sharing policies.

- 47% of survey respondents said they would be motivated to share their data if there was a journal or publisher requirement to do so; 53% of survey respondents obtained research data collected by other research groups from within a published research article; 53% of respondents said it was extremely important that data are available from a publicly available repository.
- 52% of survey respondents said funders should make the sharing of research data part of their requirements for awarding grants; 48% of respondents said that funding should be withheld (or a penalty incurred) if researchers do not share their data when the funder has mandated that they do so at the grant application stage; 73% of survey respondents strongly or somewhat support the idea of a national mandate for making research data openly available; 33% would like more guidance on how to comply with government policies on making research data openly available.

In closing, Van Gulick noted that she has only presented a glimpse of the survey and she encouraged everyone to read the full report, look at the results and the expert essays that are included in the report, and really dig into it a bit more. The report and the raw data are freely-available on Figshare [61].

10.3. How the research nexus and POSI will help build a more connected scholarly community

The third speaker was Jennifer Kemp, Head of Partnerships at Crossref. She said that over the past several years at Crossref, they have adjusted their focus from talking about DOIs and persistent identifiers on their own to a broader view of the information that is associated with those identifiers, specifically better metadata. While DOIs are critical, so is the quality and completeness of the information that is included in each record. She added that more recently, they have evolved this further to concentrate on connections or relationships among their records, which now number more than one hundred and thirty-two 132 million. Of course, relationships of various kinds among different outputs have long been part of their shared work. Some of these relationships/connections are:

- Components (e.g., supplementary data)
- Translations
- Citations/Cited-by
- Preprints to Versions of Record (and vice versa)
- Peer reviews
- Links to datasets & DataCite
- Event Data
- Funding
- Other.

Kemp said these are only a sampling of the connection types, and they really need the information linked up in the metadata to know, for example, that there is a data set available that is associated with a journal article that may have had a preprint and has software and probably some funding associated with the research behind it. She went on to say that having information available in an openly-structured way as opposed to reading about it in a journal article is especially helpful because computers do a lot of the heavy lifting when it comes to using the metadata. But it also just makes a scholarly record much more powerful and accurate because it then mirrors all the outputs and all the contributors.

She said that at Crossref they have about a million and a half relationships in the metadata, and you can imagine what the full breadth of those connections would be for all of the one hundred and thirty million records (and growing) that they have today. With all this information she can picture a Research Nexus for which relationship metadata is the foundation.

She said that their vision is ambitious and largely aspirational as of today, but she gave two recent examples. First, in November of last year, Crossref made available the grant records that are registered by their funder members, which means, among other things that publishers can use this information in their own records for better linking between funding and published outputs. And second, last month, they announced that the persistent IDs from the Research Organization Registry or ROR IDs are now available in Crossref's open APIs. She noted that Crossref is aware that collecting and including this kind of information and deposits is a necessary and a very welcome first step. But the information really needs to be linked together and made openly available and she knows that it takes work to build these connections and make use of them.

In closing, Kemp mentioned the Principles of Open Scholarly Infrastructure (POSI) [62], which is a set of guidelines around governance and sustainability for scholarly infrastructure, organizations, and initiatives. POSI makes explicit how a lot of organizations have already been working, but because the availability and persistence of all this information requires a healthy network of organizations, she hopes that it serves as a useful resource for understanding how the research support community can sustainably operate.

Kemp has submitted an article based upon her presentation and it appears elsewhere in this issue of Information Services and Use. Take a look because it includes a visual representation of Crossref's vision of the Research Nexus.

10.4. Accessible discovery via mobile for an economically diverse world

This presentation had two speakers, Michael Napoleone, Vice President of Product Management at EBSCO Information Services and Dr. Monita Shastri, Chief Librarian at Nirma University (India). Napoleone opened by saying that EBSCO has been striving to improve research and discovery workflows for libraries and their users through mobile solutions. As part of that journey, they released a mobile app to address some specific problems that they had found in the market. Some of the problems/concerns were: (1) academic research is a nonlinear process with intermittent steps along the way; (2) it is a cross-device process with needs and expectations for seamless synchronization across the devices being used; (3) there is a need for ubiquitous access to research and making it easy to conduct some of the steps in an anytime, anywhere fashion; (4) the mobile downloading of e-books downloads is difficult; and (5) making research accessible to anyone with a smartphone.

Napoleone went on to say that EBSCO's App serves as kind of a Swiss Army Knife in that it performs some key functions as a subset of desktop capabilities, but in a portable easy-to-use package. They consider this a start in a way for them to continue learning and iterating further as they continue their journey for improving research in the markets that they serve through mobile solutions. To learn more, they have made a pivot in their research and design approach. He said that in the past they had modeled their approach around role-based personas, e.g., undergraduate students, masters, faculty, librarians, professional, and researchers all with key variations across markets. This approach worked to an extent for identifying distinct solutions suited to each role. But when they look at today's generation of academic and professional researchers across roles, those professionals share common expectations for an efficient and effective research experience, and from this perspective there is less variation across roles. But at the

same time, it is a more complex world with huge variation within each role, based on factors such as their use and skills of technology, the differences in the environment in which they conduct their research, and their goals and intentions in conducting research. As a result, EBSCO has shifted to more of a needs-based persona model in order to draw distinctions and therefore design solutions between users with different types of needs. This transcends across roles and even markets.

He gave an example of a student with a high level of digital competency, who is very savvy with modern software apps and their functions, and who can adeptly conduct advanced searches in a discovery platform, but chooses not to, simply because they were not inspired by a particular class assignment. It was not a limit of technology, but rather just not something that suited their needs in that particular use case. The question here is how is EBSCO able to tap into the student's savviness in a familiar way that might pique their curiosity a bit more and gain further engagement? Or how does EBSCO get them to use research apps, rather than Google, for researching topics in which they are interested in and therefore are more likely to use some of the advanced features for more robust searching? Perhaps that involves going to where they are, which could be other apps and hooking into those apps via interoperability.

He also gave an example of an enthusiastic researcher who might be so locked into their desktop habits for research that they dismiss or overlook opportunities for mobile to play a role in their research. Or maybe they are simply not aware. Again, it is not a limitation of mobile technology, but a matter of raising awareness and helping researchers understand where or how mobile can fit into their workflow in a seamless manner.

Napoleone then turned the session over to Dr. Shastri who opened by talking about the philosophy of librarianship in India. She said that Dr. S. R. Ranganathan [63] was the founder of library science in India and in 1931, he presented five laws of library science [64] which are applicable even today. First, books are for use. Second, every reader has his or her book. Third, every book has its reader. Fourth, save the time of the reader. And fifth, the library is a growing organism. She said that the term "book" can be information in any format, and the readers or the users are the techno-savvy generation that is comfortable with e-devices because they are portable and the devices' search capabilities save them time (the 4th law of Librarianship).

She went on to say that during her information literacy sessions, she finds that she has to make students understand that while open sources such as Google are easy to access and use, the students will have to authenticate the information that is retrieved to ensure that it is correct. Since they are students, they are not subject experts, and they struggle to do this. However, her library has subscribed to academic products for them, and they should be able to easily access these products. In the past the students had to access these products through laptops, desktops, etc. and they found it clumsy. In 2021, Shastri introduced the EBSCO Mobile App and students found this a welcome change because wherever they go they are carrying their mobile devices. Now, with the mobile app, they can even easily search the academic resources at school. As a library professional, she believes that this kind of App will encourage students to use academic resources. She suggested that EBSCO also make the App usable on smartwatches or other wearable devices as ultimately the students will rely on them and not the hand-held devices.

In closing, Napoleone thanked Dr. Shastri, and went on to say that it is pretty clear to say that mobile devices are now the most used and most important computing platforms. And as EBSCO looks at the world of academic research, they still believe that mobile platforms are under-utilized and can be much more prolific in the research process going forward. It is an important area for them as they work to improve the research process.

10.5. Uber-like scholarly communication system

The final speaker in this session was Mohammad M. Alhamad, the E-resource Strategist at Missouri State University who presented on what he called an uber-like scholarly communication system. He noted that the current method of scholarly communication has its roots in the 17th century when the growth of experimental science led to the need to share the results of research with fellow scholars. Over time, publishing in scientific journals came to serve additional functions, including serving as a method for evaluating individual scientists' performance and validating the quality of research via the Impact Factor and other metrics.

He said that the purpose of scholarly communication remains the same today. Publishing provides scholars the platform through which they can share their theories and discoveries. However, the increased pace and specialization of research in the mid-20th century led to a rapid increase in the number of journals published, which attracted the attention of commercial publishers. This introduced a major defect in the current scholarly communication model. Publishers can make large profits by selling research to the very universities, taxpayers, and grant vendors that pay to produce them. Furthermore, this model prevents research findings from being disseminated as equitable as possible.

In the current model, publishers set prices with little reference to economic realities. Subscription prices have no impact on demand. When specific journals are determined to be core as a discipline, libraries have little options but to pay for them. As a result, access is not equitable rather it is based on the ability to pay. The consequence of the current scholarly communication model is that the results of research are not disseminated as widely as they could or should be. While Open Access is a good solution to get the scholarly communication system back on track, there are challenges with the Open Access system, including high article processing charges, double-dipping, funding, predatory journals, copyright issues such as Sci-Hub [65], and the demise of small publishing societies (Sound familiar? Echo from the session on ethics).

He then asked a question. With all the technology currently available, why not rethink the scholarly communication system? And before launching into a possible solution, we should ask how many researchers still choose a specific journal when researching a topic search and why, since it is now common for users to go directly to the library discovery service and search or go to Google Scholar.

He then described his idea for an open scholarly communication system - a platform for communicating research outputs to the community and for sharing across disciplines. This system will include new tools to evaluate individual scientists' performance and validate the quality of research aside from their impact factor and other existing metrics. Here, the focus would be on the articles rather than the repetition of individual journal titles. The quality of scholarly work, of course, would still be maintained through the peer-review process. In fact, reviewers will have profiles and gain scores for their review to maintain credibility and accountability. The new platform will not be a repository system for hosting preprints of research results. Instead, it will be a platform on which to publish original articles with a unique digital object identifier (DOI).

In closing, he said that the new system will not necessarily replace the current scholarly communication system. It will complement it. This may parallel the current scenario in which social media and streaming services enrich, complement, and in some cases, have overtaken older forms of media, e.g., radio, newspapers, and cable TV. – a scenario in which the news media technology has provided new methods for interacting with each other, managing our well-being, studying, and working. Most importantly, it provided us the freedom of choosing when, what, and how we want to read, listen, or watch. Also, it opened doors for entrepreneurs to innovate and thrive. There is also a great potential for big tech to step in

and collaborate with libraries to make that happen, considering the success of Highwire, Google Books, and Google Scholar.

Note: Organizations around the world are already using blockchain technology for just such a system [66]. I am in the process of writing a White paper for the International Union of Pure and Applied Chemistry (IUPAC) on the use of blockchain technology along the scientific research workflow and publishing is a hot area!

11. The role of the information community in ensuring that information is authoritative

This session focused on answers to the following question - How much responsibility does the information community have to ensure that the content that we provide is authoritative? Preprints are a great way to make early research results available, but it is not always clear that those results are not yet thoroughly vetted. Peer review - a key element of scholarly publication - can help, but it is far from foolproof. Retractions are another important tool, but most retracted research is still all too readily-available. A panel of experts in this area attempted to address what can and should be done to safeguard the integrity of the content that is being created, disseminated, and used.

11.1. Design as a tool for producing authoritative information in online crowdsourcing

The first speaker in this session was Samantha Blickhan, the Humanities Lead of Zooniverse.org [67] and Co-director of the Zooniverse team at Adler Planetarium in Chicago, IL. She opened by saying that she is going to share her perspective on where questions of authority rise in the world of online crowdsourcing and the responsibilities of those who are creating spaces for this work to take place have to help ensure that the data being produced is at a quality level that meets the expectations of the people who are using it. She said that she is not a member of the information standards community, but she does believe that there are some valuable lessons from the field of crowdsourcing and public data creation that are extremely relevant to this question of responsibility and how to ensure that the information community is producing authoritative content.

Blickhan then gave some background information on her organization. Zooniverse is the world's largest platform for online crowdsourced research. She said that they refer to themselves as a platform for people-powered research, and they do this by providing a space for researchers to build and run projects, which invite volunteers to help teams process data to aid in their research efforts. They work with hundreds of research teams around the world. Since the platform was launched in 2009, more than 2.4 million registered volunteers have collectively produced almost six hundred and fifty million classifications across more than three hundred projects. The main tool that supports this work is called the Project Builder [68], which is a browser-based tool that allows anyone to create and run a Zooniverse project for free. More than two hundred and fifty peer-reviewed publications have been produced using data from Zooniverse projects [69].

The main part of her presentation was about the ways in which authority can be a barrier to crowdsourcing and what some of her responsibilities are as a practitioner and platform maintainer to break those barriers down. She put up a quote from the *Collective Wisdom handbook* [70], which was written last year by a group of leaders in the field of cultural heritage crowdsourcing, herself included.

“Crowdsourced data is variable much in the same way that all human-generated data is variable. There will be a margin of human error, but that is not unique to crowdsourced data. A frequent misconception of crowdsourced data is that it is less trustworthy or of lesser quality than data produced by experts. In fact, crowdsourced

projects across disciplines have shown that participants can produce very high-quality results (even at the level of experts) when given adequate instruction and support”.

She said that the quote really sums up the most common barrier that she sees, which is authority as a barrier to trusting the data that is produced through crowdsourcing, generally because it is not coming from a typically-authoritative source. She believes that the question of whether crowdsourced data is trustworthy is not that straightforward for many projects and context is extremely important. She works with a broad range of data, and each tends to have its own method for determining quality. If it is not clear how a team is planning to use or evaluate their data from the onset of the project, that is when they end up with results that they are not able to use or that might be considered to be low quality. In the *Collective Wisdom handbook*, the authors defined the three main dimensions of data quality for cultural heritage crowdsourcing projects as fidelity, completeness, and accuracy. She went on to provide more information on what is meant by “fidelity” and gave the following example. A text transcription project might ask volunteers to type out the letters in a historical document as written, including preserving spelling mistakes. Therefore, a corrected transcription would have increased accuracy, but reduced fidelity to the original. So, if a desire for fidelity is part of a team’s goals, that data is going to require an adjustment of the methods that they would typically use to determine accuracy of the results.

She then went on to say that alongside the three main dimensions, the authors also talk about auditability, which is a particular consideration with crowdsourced data as well. Authority can be given to crowdsourced data by ensuring that information about how the data was produced is provided. Such information allows users or reviewers of the data with the opportunity to audit the origin of the information in relation to the project through which the data was created.

Blickhan said that there are many publications which show the quality of data produced through crowdsourcing and she gave an example from the Gravity Spy project [71]. The team on this project works with data from LIGO, the Laser Interferometer Gravitational-wave Observatory, which detects gravitational waves in space. It is incredibly sensitive and therefore susceptible to a lot of noise known as glitches. The team is trying to identify what causes these glitches so that they can be removed from the data. The size of their data sets makes it extremely difficult for experts to examine the entirety of the data, so the project is aiming to train Machine-Learning algorithms to help with this work. However, they know that Machine-Learning algorithms do not always do a great job with unexpected features, and volunteers on a crowdsourcing project might not recognize new features without guidance, or if they do, they may be uncertain as to how to interpret them. The team is using both approaches, and an article [72] compares instances of volunteer-discovered and computer-discovered glitches and ultimately shows that Zooniverse volunteers can identify new glitches at a level similar to experts.

She said that this is an example of how the success of a project is contingent upon intentional and well-thought-out project design. On the Zooniverse platform, a multi-track approach to data collection is used as a quality control method. Basically, that just means multiple people will engage in the exact same task to produce many versions of the same data. And project teams must then aggregate the data together to determine a consensus response for a given task. But rather than wait until data is produced in order to implement quality control or validation steps on the results, they actually design [73] projects based on their own experiences and the experiences of others. And she went on to describe several examples.

In closing, her final point was about testing. Zooniverse, teams are required to go through a process of internal review by her team, as well as beta review by the Zooniverse community before their project can launch publicly. And teams are required to take this feedback very seriously, including demonstrating to the reviewers how they edited their project based on the beta feedback that they received. This requires

that the teams put in more effort upfront to ensure the usability and reliability of the data being created through their project.

11.2. *Open science and signals of trust*

The second speaker in this session was Nici Pfeiffer, Chief Product Officer at the Center for Open Science, who talked about open science in the research lifecycle today. Transparency and openness is happening across all the lifecycle stages and this is a culture change for research and science that is happening across all disciplines. She said that her organization talks a lot about how the practice of open science can be pushed forward - what infrastructure is necessary to support it. She said that the key steps are making open science: (1) possible (need an infrastructure); (2) easy (offer a good user experience); (3) normative (via communities such as Zooniverse); (4) rewarding (offer incentives); and (5) a requirement (via policies).

Pfeiffer said that today incentives of individual success are based upon getting research published and not necessarily on getting it right. Research reproducibility is a big issue. If the end goal for a researcher is to publish results as a preprint or as an article and to share what they found in their study, they need to make sure that they provide all the information that is necessary for another researcher to validate and use the results. She then went on to question how open science and today's scholarly publishing system fit together and said that we need new models to advance research rigor, transparency, and the rapid dissemination of results - certainly the idea of preprints where data can be uploaded and reviewed by others can only help the researcher to ensure reproducibility before submission to a journal for publication and she mentioned Peer Community [74], a non-profit organization of researchers offering peer review, recommendation, and publication of scientific articles in open access for free.

In closing, she briefly talked about the Center for Open Science's pre-print service [75] that was lunched to meet the following goals:

- Facilitate and accelerate new models of scholarly communication across multiple research disciplines
- Improve accessibility of scholarship
- Facilitate timely feedback on scholarship
- Address delays of research publishing
- Improve scholarly communication
- Provide a single search interface to access a growing number of preprint sources.

11.3. *Providing information and building trust*

The third speaker was Dr. Bahar Mehami who holds the position of Reviewer Experience Lead, Global STM journals at Elsevier who opened by saying that the role of scholarly publishing and information hosting platforms is to provide trustworthy scientific information to the research community who then can build on top of the existing knowledge for the advancement of science. That is why journal publishers and platforms owners all try to ensure that the content that they are providing is trustworthy and this is done primarily through communication channels, workflow tools, best practices, information on the platform, collecting and analyzing data, and by providing more infrastructure to the users and the community.

But, she asked, does the community trust the information that it receives? Mehami said that according to a survey run jointly by Sense About Science and Elsevier in 2019, sixty-two percent of two thousand seven hundred and fifteen surveyed academics across the globe trust the majority of research output that they encounter, but a third doubt the quality of at least half of it.

She went on to say that Publishers and Societies are working together to improve the trustworthiness of their published content through being more transparent about their editorial and peer review processes and one example is the STM/NISO peer review terminology [76] that has been adopted in a pilot phase by several STM members. She also described the many steps that are now being taken to ensure that there is no intentional or unintentional bias introduced during the peer review process to ensure that marginalized groups receive a fair and transparent review of their manuscripts.

In closing Mehami briefly mentioned that Elsevier is working on a method to ensure that retracted articles are not cited in their journals and that such citations are identified at the time of manuscript submission so that they can alert editors and reviewers and reach out to the authors to determine if the citation is necessary.

11.4. Safeguarding the integrity of the literature - the role of retraction

Mehami's closing remark was the perfect segue to the final speaker in the session, Jodi Schneider, Assistant Professor at the School of Information Sciences, University of Illinois, Urbana-Champaign where she also runs the Information Quality Lab. Schneider gave several examples of retracted articles that continue to be cited long after the article was retracted. She said that she has been working for about two years now with a group of stakeholders to try to understand how the inadvertent spread of retracted science can be reduced and to develop recommendations for action. She added that she presented on this topic last February at the NISO 2021 conference and the draft recommendation that was presented then has recently been officially released in final form [77]. The goal is to develop a systematic cross-industry approach to ensure that the research community is provided with consistent, standardized, interoperable, and timely information.

In closing, Schneider said that because of those conversations last February, there is a new NISO Working Group that has been established - CREC [78] - to look at the communication of retractions, removals, and expressions of concern, which is meant to focus on the dissemination, the metadata, and the display of information. The Group will not focus on what a retraction is. It will focus on how the information is most effectively disseminated in the community once something is retracted. And she encouraged everyone to help them figure it out.

Schneider and other speakers in this session submitted an excellent article based upon this session and it appears elsewhere in this issue of *Information Services and Use*.

12. Open science: Catch phrase or a better way of doing research?

12.1. What is open science?

The first speaker was Shelly Stall, the Senior Director for Data Leadership at the American Geophysical Union (AGU) and a familiar face at NISO Plus conferences. She opened by saying that the single-author paper is essentially disappearing. Today researchers need collaborators and a research network and a recent study by *Nature* showed that the majority of papers in a journal are authored by people from multiple countries. Such papers outnumber those authored by a group of authors from a single county, and this is really important to understand, especially for early career researchers.

She went on to say that for the future of research, scientists will work in global teams. Researchers will require good tools to find the research that is being done worldwide and look beyond their lab partner. They are going to need good documentation to understand other researchers' data and software to determine if

there is something relevant there for their own work. And they, too, will need to share data and software in a way that is easy for other researchers to consume it. Everything will need to be interoperable and accessible. No matter which research team is creating it, it needs to be accessible, understandable, and usable by others. Software will need to be something that is commonly-used in a given discipline. And then researchers need to have licensing that supports reuse of data and software in a way that is supportive of the scientific ecosystem.

Stall went on to say that there is a lot of activity on Open Science today and that the UNESCO Recommendation on Open Science [79] was adopted by the General Conference of UNESCO at its 41st session in November 2021. She quoted one of the value statements from the document -“*Increased openness leads to increased transparency and trust in scientific information*”. She said that this is where the scientific community is headed. We really need to work collaboratively across our scientific community and know that we can all trust one another other because we are all working towards the same goal. This requires a cultural change, and we need to encourage all scientific cultures to move in this direction.

She then described the key pillars of Open Science as the following:

- Open Science Knowledge (scientific publications, open research data, open-source software and source code, open hardware)
- Open Science Infrastructures
- Scientific Communication
- Open engagement of Societal actors
- Open dialogue with other Knowledge Systems.

She went on to say that from the very beginning of the research process, the researcher both contributes to Open Science and takes advantage of the Open Science practices of other members of the research community. Researchers need to make sure that their research is discoverable and visible and to do that they need to think about persistent identifiers and make sure that the metadata is out there so that they can get credit for their work.

Stall said that at AGU they sometimes have a really hard time working openly because the cultures across global countries are not the same. She works with countries that are very willing to work openly, and with those who are not. The teams need to determine the minimal criteria for working together and that can be very hard. Working across disciplines can be even harder and require even more communication. Each researcher and each discipline and each country is adopting Open Science at their own pace, and we are definitely not all in the same place... yet.

In closing, Stall invited those interested to visit the AGU website as they have a lot of resources available on how to navigate Open Science [80].

12.2. Enabling open science with Dryad

The second speaker in this session was Jennifer Gibson, the Executive Director of Dryad, who opened with an overview of Dryad and characterized it as a not-for-profit data publishing platform and a community that is committed to the availability of all research data and its routine re-use. Dryad is for all research domains and exclusively handles fully-curated research. She said that as of today, Dryad has more than forty-three thousand data publications contributed by more than one hundred and seventy-five thousand researchers who are associated with more than thirty-two thousand international institutions, and twelve hundred academic journals. She stressed the importance of collaboration and working in concert

with other organizations and provided several examples. She emphasized that with collaboration she is looking to reach around the world and get more people to the table than in the past. She is also looking to take whatever action that she can, as an individual, to make Dryad's environment more inclusive and more equitable. She plans to showcase data through Dryad no matter where it is from. Her own personal ambition is to help make data from communities worldwide a first-class citizen in research and in research assessment.

Gibson has submitted an article based on her presentation and it appears elsewhere in this issue of *Information Services Use*.

12.3. Data standards and FAIR principles of open semantic dynamic publishing

The third speaker was Dr. Yongjuan Zhang, Senior Strategist at the Shanghai Institutes for Biological Sciences. She opened with a definition of "Smart Data", quoting James Kobiellus [81] as follows:

Smart Data is a strategy that sets the groundwork for smart decisions, rather than a particular type of data. "It is a particular set of practices for how you leverage your data for greater insights in a wider range of scenarios", allowing smart decisions to flow organically from data that meets a range of criteria.

She said that smart data is a data format that is computable, self-explanatory, and actionable in a network environment and went on to explain the technical underpinnings of the system that they have developed as the foundation for providing researchers with information via the dynamic semantic publishing model [82]. If I heard correctly, the overall process is quite complicated and has six steps (the recording is not the best for a variety of reasons, but the slides that she used helped. Unfortunately, she did not deposit her slides with NISO). The process extensively uses Machine Learning, structured encoding, data linking, etc. on Big Data, ultimately creating Knowledge Graphs and resulting in what she called "smart data" that leads to informed decision-making and facilitates dynamic publishing.

I had never heard of dynamic semantic publishing until I listened to her presentation, and I find it fascinating. Apparently, it was developed by the British Broadcasting Company (BBC) in 2010 when the Lead Technical Architect for the News and Knowledge Core Engineering department decided to move away from traditional publishing solutions and implement semantic web technologies for BBC's FIFA World Cup coverage and it has been adopted by others since then. Zhang's talk is worth watching because it gives you a good sense of what they are doing, and you will get even more out of it if you are tech savvy (which I am not!).

12.4. The F1000 publishing model

The fourth speaker in this session was Dr. Rebecca Grant, Head of Data and Software at F1000,

Grant presented a high-level overview of the F1000 publishing model which is fairly straight forward. Authors submit their article after some technical checks by the F1000 editorial team. The article is made available nearly immediately alongside any research data that underpins the conclusions of that article. Then all the peer review takes places completely openly on the F1000 platform. Peer reviewers provide their reports which are made public. Readers can comment. And then all the revisions are made public as well. Every revision of the paper, which has its own persistent identifier, and remains on the platform. She then went into finer details and their work with stakeholders in the publishing process.

Grant has submitted an article based on her presentation and it appears elsewhere in this issue of *Information Services Use*. An again, I recommend that you read Ana Heredia's paper that appears

elsewhere in this issue of *Information Services and Use* in which she gives an overview of Latin America's current scientific information infrastructure, highlighting its key role in the adoption of Open Access and Open Science in the region. It is a perfect complement to this session.

13. The state of discovery

This session focused on the process of "discovering" information and part of the discussion was "trekking into the semantic frontier". There was a panel discussion among a small group of librarians about what the "semantic frontier" means. For one, Ashleigh Faith, the panel moderator and a librarian/data scientist at EBSCO, the definition is finding whatever you want with whatever word you want without barriers. Another, Lisa Hopkins, the Head of Technical Services at Texas A&M University-Central Texas, built on that statement and gave examples of how she helps students do their research to find papers that they need on a given topic. When a student hasn't a clue how to start, she just asks them about their research and starts jotting down words as they describe what they are interested in. When she has about five to ten words, she shows them how to start plugging those into the keyword search. And the students start to realize that they do not need a polished idea with which to start. She said that she becomes a translator/guide from where they started to where they can go, and the process unlocks and opens up the student's idea of what they wanted to do. Together with the student she starts the search and travels down the paths to where the searches lead them. She said again that the process makes her feel like she is a translator. Rather than just writing down synonyms, she attempts to identify concepts from what the student is saying. She makes it look easy and it is for them when they can learn from the search box how they can expand their own searches.

A third panelist, Daniel Eller, Electronic Resource Librarian at Oral Roberts University, who also teaches information literacy to the behavioral sciences and some others. He said that one of one of his central concerns as he is trekking into this new semantic frontier is knowing where we are going and being able to explain the journey to others. He believes that for librarians, transparency is paramount if they are going to be part of mapping out this journey. A large part of what librarians do is teach information literacy. Understanding how query expansion and enhanced subject precision can help searchers requires that librarians are able to teach others how it works and how their search results and resources are generated.

He went on to say that the king of natural languages searching is obviously Google, and they do not offer much transparency when students use it. They get a mixed bag of results, and sometimes they get exactly what they want, and they are happy, but they do not know why or how Google generated the answer. In a lot of ways, Google is a black box. And librarians need to take care when teaching that they themselves do not create black boxes, that they themselves understand the process that brought them to the answer. Librarians want the system to begin to learn from them and make these connections for them. But during that process, the most important issue is that librarians never lose the transparency, so that when they plug in a subject heading or a natural language term, and it is translated, they can explain to the students every time why they are getting the results that they see. Eller said that he believes that the key to this is for librarians, teachers, and vendors to be in a constant dialogue, and work as a team. He believes that the future is a team dialogue that will allow us to create a better dialogue with the system and with the students.

Ashleigh Faith came back to close the session and said that librarians are all focused on helping their users find information more easily and they need some help from NISO - they need more standards. She went on to say that mapping is not very standardized. There is not much out there to help a librarian to implement a knowledge graph in search and she asked them to investigate that and to move librarians -

and libraries - into that great beyond, into Star Trek, where they will not have to worry about language. They will all speak the same language, and they will all find whatever they want.

Note that Daniel Eller submitted a paper, “Transparency and the Future of Semantic Searching in Academic Libraries”, that appears elsewhere in this issue of *Information Services and Use*.

14. Research infrastructure for the pluriverse

In the closing keynote (which I thoroughly enjoyed), Dr. Katharina Ruckstuhl, Associate Dean and Senior Research Fellow at the University of Otago’s Business School in New Zealand, discussed what is meant by a “pluriversal” approach to a research infrastructure, and how Indigenous scholars and allies have been thinking about and implementing research infrastructure processes, with implications for standards, both for today and into the future. Ruckstuhl is an expert on the topic as she is the Māori [83] lead of a major New Zealand “grand challenge”, Science for Technological Innovation [84]. She has leadership roles with her tribe of Ngāi Tahu [85], is an ORCID Board member, and a member of the IEEE working party on standards for Indigenous people’s data. She has published on Māori language, Māori economy, and Māori science and technology and she is currently co-editing a book on Indigenous Development.

Ruckstuhl opened with an incantation in her native Indigenous language which told the journey of knowledge - moving from that which is obscure, through various states of understanding, then to enlightenment. She talked about her tribe, the Ngāi Tahu, which had a three-thousand-year journey to where it is today in the South Island of Aotearoa, New Zealand. The journey started in Southeast Asia, the tribe island-hopped its way through Micronesia [86], then to East Polynesia, finally settling in New Zealand in about 1,200 A.D. in a series of ocean migrations. The story that she was about to tell carries traces and memories of those voyages.

In this story, Aoraki and his brothers descend from the heavens and travel in a canoe to the South Island to explore and fish. After being there some time and wishing to return to the heavens, Aoraki begins his incantation, similar to the one with which Ruckstuhl opened her presentation. However, because Aoraki’s brothers were very quarrelsome, he lost his concentration, and consequently he did not correctly perform the incantation that would take them home. The canoe broke into several parts with the prow at one part of the island and the stern at the other. And in between, the canoe and its crew were turned into stone and formed the mountains that today are known as the Southern Alps, of which Aoraki is the peak. After that, Aoraki’s grandson traveled through the whole land and across the coastlines, naming each place and these are the first names. But these first names were changed by those who came next. However, the first names have not been forgotten. She then asked a question - what can we understand from this narrative?

She went on to say that at its simplest it could be considered a legend to explain some geographical features - a mountain range that looks like an upturned canoe. But if one is attentive to the details, Aoraki is a name that can be found in Tahiti and in the Cook Islands, mirroring the tribe’s three-thousand-year journey - a cultural continuance. The story could also be considered moral instruction - how leaders are supposed to behave. For despite being a demigod, Aoraki failed in his key role, which was to protect his canoe and its crew from misfortune through correct adherence to protocol and incantation. As a consequence, they were turned to stone and became a mountain range. Aoraki failed to observe the correct standard that had been laid down for him, and the consequence was fatal. But those of the tribe who came after were very grateful.

Ruckstuhl said that she told the story because she believes that it will help to elucidate the following quotation.

“A system’s effectiveness in organizing information is part a function of an ideology that states the ambitions of its creators and what they hope to achieve [87]”.

She then applied the quotation to the Aoraki story and focused on the term “ideology” and defining it as follows: an ideology, according to Merriam-Webster, is a manner of thinking characteristic of an individual group or culture, the integrated assertions, theories and aims that constitute a sociopolitical program, a systematic body of concepts, especially about human life or culture, and visionary theorizing.

She went on to talk about the ideology of the ancient Māori, saying that the main ideology was continuance through observation of correct behavior. And much of that correct behavior was through their relationships, not only with tribal members, but also with the gods and the natural environment. Naming places created those relationships. Relationships are a key ideology of the Māori, and indeed of many Indigenous people. She then asked several questions: in terms of organizing information as a function of an ideology, how did the Māori do that? did they have a systematic body of concepts? how were the concepts integrated? what were their characteristics?

Continuing, she said that the Māori organizing principle was “whakapapa”. At its simplest, whakapapa means to place in layers, to stack flat, and it also means to recite genealogy and she then displayed a slide of her genealogy. She said that she can name the ancestors who came from the Pacific, how these relate to other tribal groups in Aotearoa, and how they relate to more recent kin. She is also able to name these back to certain gods, or atua, as was shown on her slide. She added that whakapapa does not only organize human and god relationship; it also organizes relationships between humans, gods, and all animate and inanimate beings in the universe. The organizing principle is around continuance through right relationships with the natural world and with the universe and she said that this this raises the following questions.

Given that the Māori’s ideology (and many other Indigenous peoples have similar ideologies) is about maintaining right relationships with the world, of which people are only one aspect, can modern research infrastructures be effective? And not just efficient in organizing information to reflect an ideology of right relationships. And what is a right relationship with the world of Indigenous people? What is the wrong relationship? She said she will try to lay out some of the rights and wrongs, and ultimately close with some reflections on potential future applications of these ideas for a Pluriverse research infrastructure.

The first problem that she discussed was *terra nullius* [88] thinking. *Terra nullius* itself means a land that is no one’s. And she gave the example of the instructions given to Captain James Cook in 1769 as he was about to head off to the Pacific on his first voyage - he undertook three - to document the transit of Venus. The voyage was funded by the Royal Society of London and Cook was instructed by the admiralty to take possession of uninhabited countries. Now Cook did not discover any lands that were uninhabited, but he was justified to take lands from those who did not inhabit land in the same way as Europeans. Therefore, those who used the land, but whose residences might have been cyclical or seasonal were deemed to *not* inhabit the land and that land was open to being taken or alienated. Through history, that has been the pattern that can be seen in New Zealand, Canada, the U.S., and Australia. Even though there have been many treaties with the Indigenous peoples of those countries, in effect there was a type of *terra nullius* thinking going on.

She went on to say that not only was land alienated, but other things, both tangible and intangible were alienated as well. And these included alienation from one’s physical environment - mountains, rivers, seas, animals, rocks, stones, all the things that she mentioned when talking about the whakapapa. There was alienation from traditional food sources, from traditional labor and economy, from cultural pursuits, including language, narrative, songs, and ritual - even from physical belongings such as everyday objects and even revered objects. Indigenous people were alienated from themselves as a tribal people. Yet despite

this alienation, communal knowledge has persisted in pockets orally, through elders, through wise people, through protest, and through legal challenges. And of course, it has persisted in documentation in records, archives, museums, photos, sound recordings, etc. The impact of *terra nullius* was to transfer not only land, but also knowledge to other people and into infrastructures where it could be stored, categorized, studied, and then reassembled for the purposes and ideologies of the owner of the infrastructure.

The ideologies of the owners of the “new” infrastructure could be categorized as ownership and trade, accounting and record keeping, documentation, preservation of public education, and knowledge dissemination, whether in books or articles, or legal texts, or public records published by Societies such as the Royal Society of London, and in New Zealand, the Royal Society of New Zealand. Today, the ideologies are similar, but are driven by today’s concerns - accessibility, Open Source, etc. This is the idea of universality, the idea that knowledge can somehow freely move and serve the common good. However, Indigenous people often ask, whose good is served by the common good? Ruckstuhl then put up the following quotation:

“Research” is probably one of the dirtiest words in the Indigenous world’s vocabulary because research has taken specific Indigenous collective goods, tangible and intangible, and used them to benefit other people [89].

She went on to say that Indigenous people have not forgotten their collective goods, their collective knowledge, or their collectively-held places and environments. These collectively-held goods are emblematic of distinct worlds based on particular types of relationships with particular types of environments. However, what we have had until quite recently with regards to research infrastructure is what Colombian academic Arturo Escobar has described as the one world *world* approach as follows:

“What doesn’t exist is actively produced as non-existent or as a non-credible alternative to what exists[90]”.

Ruckstuhl said that this is the concept that worlds, such as the one that she described when she began her presentation, are actively produced as non-existent and a non-credible alternative to the current dominant, a colonized world. By this she means that by viewing the story of Aoraki and his brothers as a myth or a primitive religious belief, one denies other possibilities for such coded knowledge. Answering questions of today is the mission of current research infrastructures, and some people are going to disagree that so-called myth and science can be viewed in the same breath. She said that many countries, including her own, have maintained a very strict division between what is science and what it is not. And she went on to give some examples. She displayed a Google screengrab from 2016 based on keying in the word “scientist. The scientist is depicted as laboratory-based, with chemicals, white coats, and a sterile environment. But if you type in the term “Indigenous Scientist” you see something quite different. There is no lab, people are in a field with a number of elders present and it is browner.

Fast-forward to February 2002 and there is quite a shift. When “Indigenous scientist” is displayed it is very feminine and much more into the *one world world* of the laboratory. Something has changed in the algorithm to include Indigenous people as scientists, but in a one world sense. But it is leaving open the possibility that science and traditional knowledge can sit side-by-side. She said that this leads to yet another question - how do we allow for a pluriverse research infrastructure where Indigenous worlds and the traditional, modern research world can co-exist without subsuming each other? Shen then displayed another quote which reminded me of what Darcy Cullen discussed on day one of the conference in the session on Indigenous Knowledge that I summarized earlier in this paper:

“Today there is a paradox of scarcity and abundance for Indigenous data. There is A scarcity of data that aligns with Indigenous rights and interests and which Indigenous People can control and access. There is an abundance of data that are buried in large collections, hard-to-find, mislabeled, and controlled (legally and literally) by others [91]”.

She raised the issue of while infrastructures are talking to infrastructures where do Indigenous rights and responsibilities fit? She said that the answer is the key CARE [92] principles which are around Collective benefit, Authority to control, Responsibility and Ethics - these are very people-oriented principles, very relational and the sit alongside the FAIR (Findable, Accessible, Interoperable and Reusable) [93] data principles. The goal is to practice CARE in collection; engage CARE in stewardship; implement CARE in the data community; and use FAIR and CARE principles together in data applications. She said that there is a Working Group that is tackling this goal, but they are only beginning. From her perspective it will boil down to what she termed the four P's- Power-equity, People, Process, and Policy. At this point she specifically recommended that those listening view Darcy Cullen's presentation from earlier this week, specifically with regard to the labels being developed for Traditional Knowledge (TK labels) and Bio-Cultural labels (BC labels). These labels are appended to a digital resource, and they work by creating a space in the metadata to allow for indigenous relationships to the data, whether that is retrospectively to already-existing collections or data, or potentially prospectively to a potential re-use, perhaps even a commercial re-use.

Ruckstuhl said that she wanted to close with some speculations. The fact that NISO has a three-year plan around diversity, inclusion, and equity suggests that NISO may want to tackle something serious, not only ethically, but also perhaps in terms of the infrastructure. She also reflected on the conference's opening keynote on the Metaverse and went on to say that based on the following quotation, the Metaverse seems remarkably similar to the "one world world" concept:

"Internet functionality requires land-based infrastructure. The relationship between the digital and land is inextricable, and it is erroneous to think of cyberspace as landless [94]"

In closing, she said that she prefers "Pluriverse [95]" over "Metaverse", but if the latter is the next big thing, she proposed that the CARE principles are very applicable to that new world - collective benefit, authority to control, responsibility, and ethics. She commended NISO for allowing diverse voices to be brought forward, because these are not just ethical ideas, and they need to work on both sides - that is what a "Pluriverse" infrastructure is all about.

15. 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-six sponsors who made it possible and allowed the global audience to participate from twenty-eight countries at an affordable rate (*note*: this was two more countries than in 2021!). 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 2022 conference would bring together six hundred and thirty 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 that 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. NISO also hopes to have an in-person meeting later in 2022 pending the status of the pandemic (it took place in September), but NISO Plus 2023 will again be virtual and is scheduled for February 14th – 16th *So mark your calendars!!*

16. 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 year's conference.

- Open Science and sharing, citing, and reusing datasets require a cultural and behavioral shift among researchers. The global research community is not there yet.
- The process of making information accessible to and usable by those who are disabled is difficult, time-consuming, and complicated. The limitations related to PDFs in particular are problematic.
- Creating rich metadata is essential to facilitate information discovery and preservation (also a theme in 2021).
- Preservation Policies are living documents that are essential for libraries, publisher, and archives.
- Making ethics a core value in the creation, evaluation, and dissemination of authoritative scholarly information is an ongoing effort in the Information Community.
- Using standards is essential to the global sharing of data and scholarly information.

The majority of the presentations that I “attended” were excellent. I thoroughly enjoyed the opening keynote, perhaps because I am fascinated by Virtual Reality and the Metaverse, but I found the speaker compelling. I was especially impressed with the presentations in the session on working towards a more ethical information community as I was unaware of the Sustainable Libraries Initiative, the SDG Publishers Compact, and the Climate Change Knowledge Co-operative (CCKC) which were mentioned in the talks.

I always like it when (1) I am blown away by a technology of which I was unaware such as Virtual-meta (which could solve some of the PDF problems that were discussed!); (2) when I learn a new word/concept such as “coopetition [96]”, (even though the concept has been around since 1913! *Where have I been?*); or (3) when I am made aware of an issue to which I had never given much thought such as the effort that goes into making information accessible to and usable by disabled information seekers and how challenging the process can be - both with regards to policies and with regards to documents such as PDFs. Those are the take-aways that, for me, make attending a conference worthwhile and those are the things that made attending the 2022 NISO Plus conference worthwhile for me.

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. But perhaps I am being unfair. 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 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? Perhaps those questions get answered to a certain extent in the “Plus” part of the conference and will always be a smaller portion of the conference symposia. I do wonder if a new conference will emerge to focus on such issues.... I hope so!

The only suggestion for improvement that I have for the next global conversation is that speakers be given guidance on their presentations - such as, do not repeat what you said last year. I was disappointed

that some speakers did do just that thing. And, please, offer up your slides for the repository. And, *double please*, speak more clearly and louder when recording. The amount of “inaudible” was frustrating. Perhaps NISO should audit all recordings for just those issues (but in reality, NISO probably does not have the time and staff to do 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!!

Additional information

The NISO 2023 Conference will take place virtually from February 14–16, 2023, and registration [97] is now open.

If permission was given to post them, the speaker slides that were used during the 2022 NISO Plus Conference are freely-accessible in the NISO repository on figshare [98]. If permission was given to record and post a session, they are freely-available for viewing on the NISO website [99]. The complete program is there as well. I do not know how long they will be available, but it appears that the recordings from 2021 and 2020 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 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*.

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/70fb682a-c3b7-4245-a097-4d45ef0fbd4c>, accessed September 19, 2022.
- [2] See: <https://www.theverge.com/2021/10/28/22745234/facebook-new-name-meta-metaverse-zuckerberg-rebrand>, accessed September 24, 2022.
- [3] See: <https://en.wikipedia.org/wiki/metaverse>, accessed September 23, 2024.
- [4] See: <https://www.molecule.to/blog/molecules-first-ip-nft-in-the-united-states>, accessed August 16, 2022.
- [5] Haptics definition: The use of electronically or mechanically-generated movement that a user experiences through the sense of touch as part of an interface (as on a gaming console or smartphone), <https://www.merriam-webster.com/dictionary/haptics>, accessed September 25, 2022.
- [6] See: <https://content.iospress.com/journals/information-services-and-use/41/1-2?start=0>, accessed September 25, 2022.
- [7] See: <https://sdgs.un.org/goals>, accessed September 28, 2022.
- [8] See: <https://sustainablelibrariesinitiative.org>, accessed September 28, 2022.
- [9] See: <https://sustainablelibrariesinitiative.org/resources/professional-development/roadmap>, accessed September 28, 2022.
- [10] See: <https://info.growkudos.com/climate-change-knowledge-cooperative>, accessed September 28, 2022.
- [11] See: <https://www.plantwise.org/KnowledgeBank/> accessed September 28, 2022.
- [12] See: <https://www.cabi.org/projects/prise-a-pesst-risk-information-service/>, accessed September 29, 2022.
- [13] See: <https://www.cabi.org/projects/action-on-invasives>, accessed September 29, 2022.
- [14] See: <https://www.cabi.org/ISC/>, accessed September 29, 2022.
- [15] See: <https://sdg.internationalpublishers.org>, accessed September 29, 2022.
- [16] See: <https://sdgs.un.org/HESI/sdg-publishers-compact>, accessed September 29, 2022.
- [17] See: <https://www.research4life.org>, accessed September 29, 2020.
- [18] See: <https://info.africarxiv.org>, accessed October 2, 2022.
- [19] See: <https://www.tcc-africa.org>, accessed October 2, 2022.
- [20] See: <https://www.masakhane.io>, accessed October 2, 2022.
- [21] See: <https://nisoplus.figshare.com>, accessed October 2, 2022.
- [22] See: <https://www.ubcpres.ca/ravenspace>, accessed October 2, 2022.
- [23] See: <https://np22.niso.plus/Title/b285c303-257d-4a7b-8706-eb3f9a94bbfc>, accessed October 2, 2022.
- [24] M. Sraku-Lartey, Connecting the world through local indigenous knowledge, *Information Services and Use* 41(1–2) (2021), 43–51. <https://content.iospress.com/journals/information-services-and-use/41/1-2?start=0>.
- [25] See: https://www.wikidata.org/wiki/wikidata:Main_Page, accessed October 2, 2022.
- [26] See: <https://www.wikidata.org/wiki/wikidata:Introduction>, accessed October 2, 2022.
- [27] See: <https://disability-studies.leeds.ac.uk>, accessed October 3, 2022.
- [28] See: <https://www.plutojournals.com/international-journal-of-disability-and-social-justice/>.
- [29] Formally, “Marrakesh Treaty to Facilitate Access to Published Works for Persons Who Are Blind, Visually Impaired or Otherwise Print Disabled”; see <https://www.wipo.int/treaties/en/ip/marrakesh/>, accessed October 3, 2022.
- [30] See: <https://www.w3.org/WAI/standards-guidelines/wcag/>, accessed October 3, 2022.
- [31] See: <https://www.section508.gov/sell/vpat/https://www.itic.org/policy/accessibility/vpat>, accessed October 3, 2022.
- [32] See: <https://nisoplus.figshare.com>, accessed October 4, 2022.
- [33] See: <https://www.archives.gov/preservation/electronic-records.html>, accessed October 4, 2022.
- [34] See: <https://www.archives.gov/records-mgmt/policy/transfer-guidance.html>, accessed October 4, 2022.

- [35] See: <https://www.archives.gov/files/records-mgmt/policy/m-19-21-transition-to-federal-records.pdf>, accessed October 4, 2021.
- [36] See: <https://github.com/usnationalarchives/digital-preservation>, accessed October 5, 2022.
- [37] See: <http://www.iso16363.org/iso-certification/preparation/https://public.ccsds.org/Pubs/652x0m1.pdf>, accessed October 5, 2022.
- [38] L. Johnston, Challenges in preservation and archiving digital materials, *Information Services and Use* 40(3) (2020), 193–199. <https://content.iospress.com/journals/information-services-and-use/40/3>, accessed October 5, 2022.
- [39] See: <https://clockss.org>, accessed October 5, 2022.
- [40] NASIG is an independent, non-profit organization working to advance and transform the management of information resources; see: <https://www.nasig.org>, accessed October 5, 2022.
- [41] See: https://nasig.org/resources/Documents/Publications/NASIG-Guides/NASIG_DPTF_Survey_Report_2019-05-01.pdf, accessed October 5, 2022.
- [42] See: <https://nasig.org/NASIG-model-digital-preservation-policy>, accessed October 5, 2022.
- [43] See: <https://www.copim.ac.uk>, accessed October 5, 2022.
- [44] See: <https://www.copim.ac.uk/workpackage/wp7/>, accessed October 5, 2022.
- [45] See: <https://www.ukri.org>, accessed October 5, 2022.
- [46] See: <https://cmoa.org/art/Teenie-Harris-archive/>, accessed October 5, 2022.
- [47] See: <https://transkribus.eu/lite/>, accessed October 5, 2022; and <https://readcoop.eu/transkribus/>, accessed October 5, 2022.
- [48] See: <https://readcoop.eu/transkribus/>, accessed October 16, 2022.
- [49] See: <https://hamlet.andromedayelton.com>, accessed October 10, 2022.
- [50] B. Lawlor, An overview of the 2021 NISO Plus conference: global connections and global conversations, *Information Services and Use* 41(1–2) 1–37. <https://content.iospress.com/journals/information-services-and-use/41/1-2>, accessed October 5, 2022.
- [51] <https://www.gartner.com/en/articles/the-4-trends-that-prevail-on-the-gartner-hype-cycle-for-ai-2021>, accessed October 5, 2022.
- [52] See: <https://sebastianraschka.com/faq/docs/ai-and-ml.html>, accessed October 6, 2022.
- [53] See: https://www.buchmesse.de/files/media/pdf/White_Paper_AI_Publishing_Gould_Finch_2019_EN.pdf, accessed October 6, 2022.
- [54] See: <https://www.elsevier.com/connect/trust-in-research>, accessed October 6, 2022.
- [55] See: <https://www.niso.org/node/25942>, accessed October 6, 2022.
- [56] See: <https://www.bibtex.com/g/bibtex-format>, accessed October 8, 2022.
- [57] See: <https://www.latex-project.org>, accessed October 8, 2022.
- [58] See: <https://np22.niso.plus/Title/d37428d2-249c-42ea-89a8-fae5c5b3d22>, accessed October 8, 2022.
- [59] See: The Future of Text: <https://futuretextpublishing.com/>, Augmented Text Software: <https://www.augmentedtext.info/>; Infrastructure for Rich PDF documents: <https://visual-meta.info/>, accessed October 16, 2022.
- [60] See: <https://www.go-fair.org/fair-principles/>, accessed October 16, 2022.
- [61] https://digitalscience.figshare.com/articles/report/The_State_of_Open_Data_2021/17061347, accessed October 7, 2022.
- [62] See: <https://openscholarlyinfrastructure.org>, accessed October 8, 2022.
- [63] See: https://en.wikipedia.org/wiki/S._R._Ranganathan, accessed October 7, 2022.
- [64] See: https://en.wikipedia.org/wiki/Five_laws_of_library_science, accessed October 7, 2022.
- [65] See: <https://www.sci-hub.sthttps://sci-hub.se>, accessed October 7, 2022.
- [66] See: <https://www.disruptordaily.com/blockchain-use-cases-publishing>, accessed October 8, 2022 and D.W. Gunter, *Transforming Scholarly Publishing with Blockchain Technologies and AI*, IGI Global, 2021, ISBN 9781799855903 Paperback.
- [67] See: <https://www.zooniverse.org>, accessed October 9, 2022.
- [68] See: <https://www.zooniverse.org/lab>, accessed October 9, 2022.
- [69] See: <https://zooniverse.org/publications>, accessed October 9, 2022.
- [70] See: <https://britishlibrary.pubpub.org>, accessed October 9, 2022.
- [71] See: <https://www.citizenscience.gov/catalog/368/#>, accessed October 16, 2022.
- [72] See: <https://arxiv.org/abs/2103.12104>, accessed October 9, 2022.
- [73] See: <https://hal.archives-ouvertes.fr/hal-02280013v2https://doi.org/10.1093/Ilc/fqx064>, accessed October 9, 2022.
- [74] See: <https://peercommunityin.org>, accessed October 9, 2022.
- [75] See: <https://www.cos.io/products/osf-preprints>, accessed October 9, 2022.
- [76] See: STM, the International Association of Scientific, Technical and Medical Publishers. A standard taxonomy for peer review [Internet]. OSF; 2020 Jul [cited 2022 May 12]. Report No.: Version 2.1. Available from: <https://osf.io/68rnz/>, and NISO. Peer review terminology standardization [Internet]. [cited 2022 May 12]. Available from: <https://www.niso.org/standards-committees/peer-review-terminology>, accessed October 9, 2022.

- [77] See: <https://osf.io/preprints/metaarxiv/ms579/>, accessed October 9, 2022.
- [78] See: <https://www.niso.org/standards-committees-crec>, accessed October 9, 2022.
- [79] See: <https://en.unesco.org/science-sustainable-future/open-science/recommendation>, accessed October 10, 2022.
- [80] See: <https://www.agu.org/>, accessed October 10, 2022.
- [81] See: <https://www.dataversity.net/big-data-smart-data-big-drivers-smart-decision-making/>, accessed October 11, 2022.
- [82] See: <https://www.ontotext.com/blog/future-now-dynamic-semantic-publishing/>, accessed October 12, 2022.
- [83] See: https://en.wikipedia.org/wiki/Māori_people, accessed October 15, 2022.
- [84] See: <https://www.sftichallenge.govt.nz>, accessed October 15, 2022.
- [85] See: https://en.wikipedia.org/wiki/Ngā_Tahu, accessed October 1, 2022.
- [86] See: <https://en.wikipedia.org/wiki/Micronesia>, accessed October 16, 2022.
- [87] See: R. Svenonius, *The Intellectual Foundation of Information Organization*, MIT Press, 2020.
- [88] See: https://en.wikipedia.org/wiki/Terra_nullius, accessed October 16, 2022.
- [89] See: L. Smith, *Decolonising Methodology*, 1999, p. 1.
- [90] See: A. Escobar, *Designs for the Pluriverse*, 2017, p. 68.
- [91] See: S. Russo Carol, *Operationalizing the CARE and FAIR Principles for Indigenous data Futures*, 2021.
- [92] See: <https://www.gida-global.org/care>, accessed October 15, 2022.
- [93] See: <https://www.go-fair.org/fair-principles/>, accessed October 15, 2022.
- [94] See: A. Morford and J. Ansloos, Indigenous sovereignty in digital territory: A qualitative study on land-based relations with #nativetwitter, *AlternativeNative: An International Journal of Indigenous People*17: (2021), 293–305. see: https://www.researchgate.net/publication/352960587_Indigenous_sovereignty_in_digital_territory_a_qualitative_study_on_land-based_relations_with_NativeTwitter, accessed October 15, 2022.
- [95] See: <https://en.wiktionary.org/wiki/pluriverse>, accessed October 17, 2022.
- [96] See: <https://en.wikipedia.org/wiki/Coopetition>, accessed October 10, 2022.
- [97] See: <https://niso.plus>, accessed October 10, 2022.
- [98] See: <https://nisoplus.figshare.com>, accessed October 9, 2022.
- [99] See: <https://np22.niso.plus>, accessed October 10, 2022.