

# An Overview of the NFAIS 2018 Annual Conference: *Information Transformation: Open, Global, Collaborative*

Bonnie Lawlor\*

Guest Editor, NFAIS Honorary Fellow, 276 Upper Gulph Road, Radnor, PA 19087, USA

**Abstract.** This paper offers an overview of the highlights of the 2018 NFAIS Annual Conference, *Information Transformation: Open, Global, Collaborative*, that was held in Alexandria, VA from February 28 - March 2, 2018. The goal of the conference was to take a close look at the initiatives that have emerged as a result of the increasing global acceptance of Open Science and Open Access ideologies and policies. These include the rise of private funding foundations that are mandating more open and collaborative research; innovative new technologies and tools that have opened dialogue in the research community; and increased interoperability and seamless access to not only the scholarly article, but also to all associated digital research objects. It became clear that the ultimate and common goal of all stakeholders in the scholarly and scientific communities is the rapid dissemination of scholarly communication with the parallel advancement of scientific research. It is also become clear that there are divergent views on when and how this goal will be reached. The NFAIS 2018 Annual Conference provided a very interesting overview of how the scholarly community is attempting to work together towards a more open, global, and collaborative future.

Keywords: Scholarly communication, open access, collaborative research, open data, open science, community building, knowledge creation, blockchain technology, peer review, Sci-Hub, preprints, Open Access 2020, science funding, future of the book, Unpaywall, Kopernio, artificial intelligence, University of Oklahoma, collaborative knowledge foundation

## 1. Introduction

Much of today's research and scholarly communication landscape was foreseen almost twenty years ago when the U.S. National Academies issued a report entitled *Issues for Science and Engineering Researchers in the Digital Age* [1]. The preface of the report opens as follows:

“The advance of information technology presents enormous opportunities in the conduct of research. In many ways, today's electronic tools of communication and computing make possible heightened productivity and creativity. At the same time, use of these tools *challenges many of the traditions of academic research...*”

The report went on to say

“...tomorrow they (the tools) will be universal, viewed as necessities. In the meantime, applications of information technology grow rapidly more sophisticated and automated, adding media and capabilities that open up new ways of learning, communicating, and creating knowledge. Scientists and engineers may be

---

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

concerned about the broader implications of these changes, including how they may affect *and possibly threaten existing norms and conventions.....*”

The report foreshadowed today’s reality for certainly the then-existing norms of scholar communication and publishing have since come under attack. Ironically, only a year after the report was published its implications were reinforced by demands for Open Access to scientific information in 2002 by the Declaration of Budapest [2] and again in 2003 by the Declaration of Berlin [3]. The die was cast, the conversations began, and scholarly communication began to move swiftly towards an environment of openness and sharing as exemplified by this very condensed timeline:

- 1998** BioMed Central founded
- 2000** PubMed Central launched
- 2001** Wikipedia launched
- 2002** Budapest Open Access Declaration  
Creative Commons Licenses first released
- 2003** Berlin Open Access Declaration  
Directory of Open Access Journals founded  
Public Library of Science (PLOS) founded
- 2008** Open Access Publishers Association founded  
Annual Open Access Week launched
- 2010** SpringerOpen launched
- 2013** PeerJ launched
- 2014** Nature Communications becomes fully Open Access  
Elsevier launches its 100th Open Access Journal
- 2017** Unpaywall launched

While the demand for changes in scholarly communication has grown, the impact that these changes, particularly Open Access, will ultimately have on traditional publishing practices and revenue models is not fully understood. The conversations initiated almost twenty years ago have not ended and continue to this day. One of them took place earlier this year at the 2018 Annual Conference organized by the National Federation of Advanced Information Services (NFAIS™). The meeting entitled, *Information Transformation: Open, Global, Collaborative*, attracted a large group of researchers, publishers, librarians, policy makers, and technologists who gathered together in an attempt to learn how all stakeholders in scholarly communication are attempting to handle the sociocultural shifts in how research is accomplished - global collaboration, use of sophisticated technologies in the laboratory, the need for and expectation of rapid dissemination of research results, the ability to share and reproduce research results, the demand for ease-of-access to and convenient use of information, etc.

This particular conversation went on for two-and-a-half days while stakeholders discussed funding policies, new technologies for free journal access, guidelines for building effective communities, open access models, etc. While at the close of this particular conversation no consensus was in sight, I believe everyone left understanding that most stakeholders really are working, albeit within the constraints of their businesses, missions, etc., to evolve and build a new information infrastructure that will accommodate the rapidly-changing requirements of the new digital information order.

## 2. Setting the stage

The keynote presentation by **Cameron Neylon**, Professor of Research Communications at Curtin University in Western Australia, was absolutely perfect for setting the stage of this year's "conversation." His focus was to identify what is driving change in scholarly communication, what is actually changing, and how the different perspectives of the stakeholders - scholars, publishers, funders, platform providers, and the myriad of information professionals - lead to a partial focus that can make us simultaneously fearful of the change we see and blind to the shifts that actually matter. He noted that he spoke on change at the 2010 NFAIS Annual Conference and at that time predicted that scholarly communication was undergoing major disruption and that some of the organizations represented in the audience might find themselves replaced. In preparing for the 2018 talk he looked back at his 2010 paper and had to admit that while some things have changed, the major changes he thought would occur did not. He had predicted that if an *open* network was created that used consistent standards for how people communicated and everyone adopted it, the traditional framework for scholarly communication would no longer be needed. He used FriendFeed as an example. This was a real-time feed aggregator that consolidated updates from social media and social networking websites, blogs, etc. with which it was possible to create customized feeds to share, as well as originate new posts-discussions, (and comment) with friends. It allowed people to not only track the social media activities of their own friends, but also to track such activities across a broad range of different social networks. The network was purchased by Facebook in 2009 and shut down in 2015. He was convinced when he spoke in 2010 that such a network would be a major disrupter, but admitted that his 2010 vision of an open network failed because people prefer to do things in the way that they are used to and prefer to choose with whom they communicate.

He then made two assertions. The first is that knowledge grows and it matters that it grows. He gave examples of work done by Derek de Solla Price on tracking the output of research through most of the last century [4], but added that it is not only research output that has grown, but also the number of people who seek knowledge. He used the number of students per capita who study at universities as the basis for this and noted that the number had grown from close to zero percent in the year 1900 to almost twenty percent of the world's population by the year 2000 [5]. He said that the growth of knowledge is essential to the continuation of civilization.

His second assertion is that knowledge is made by groups. An individual may come up with an idea, but until it is shared with others it cannot become knowledge. He said that Ludwig Fleck wrote about this concept in 1935, but it was not until his work was translated into English (from German) that his concept became more widely known [6]. Fleck said that knowledge is created by groups. It is the process of one group (experts) sharing their information with another external group (others) and then feedback being returned to the first group that actually creates knowledge. Others have since built models of this concept and they agreed that knowledge is created within groups (research groups, departments, etc.), but it is the process of transferring that knowledge externally - perhaps across disciplines or even through peer review - that allows it to be validated, shaped, and built upon. He said that the knowledge claims that are transferred between groups are not fully understood by either group. He gave the ideal gas law as an example. Robert Boyle had the concept, but he never would recognize the equation that ultimately expresses that concept. The equation is the result of discussion, further studies, and increased understanding of the concept. Knowledge is a product of translation and that general knowledge is produced at the boundaries of groups in contact and/or conflict.

But, Neylon asserted, if you agree that knowledge grows and that knowledge is created by groups, then you must accept that there is a problem because groups do not scale. Information management systems

for a laboratory do not work across an entire university. If you increase the size of a group it will fracture politically as people make their own relationships. If you create more groups they tend to each go off in their directions. He said that what supports social scaling is shared culture and the institutions involved in knowledge growth. But these, too, have scaling limits. In retrospect, what we have seen over time is knowledge growth until we hit a wall and a crisis occurs. This is followed by a process of innovation, solutions will be discovered and widely-adopted, and growth will be managed and continue. He gave the example of the growth in research after World War II when the scholarly communication process almost collapsed due to the increase in the number of articles being published and the number of new journals being launched. What had been groups of “cottage” publishing houses within individual countries could not scale to support this global distribution. He said that Robert Maxwell saved the day because he brought an industrial discipline to the process, consolidated publishing into the large companies we know today, and the growth became manageable. But there will always be a “next crisis.” By 1975 there was a concern about the quality of what was being published and it was decided that there needed to be a definition of what constitutes scholarly output. Peer review was the solution. In the early 1990s, the problem was information discoverability and integration in a world of knowledge silos. The web solved that problem, but since has ultimately caused problems of its own. Today, he believes the crisis that is brewing is one of trust. People do not know if they have discovered all of the information that they need on a given topic and are unsure if the information that they do have is reliable and useful. They question what they can and/or should use. He does not believe the problem is at crisis level, but he said that we need to prepare to handle the crisis when it comes.

He noted that science and scholarly communication is an old culture that has been able to support scaling and that “openness” has always been at the heart of that culture. He referred to a quote from Robert Boyle circa 1665 as follows:

“Of my being somewhat prolix [...] I thought it necessary to deliver things circumstantially, that the Person I addressed them to might, without mistake, and with as little trouble as is possible, be able to repeat such unusual Experiments.”

He reiterated that knowledge is created by groups and that the transfer of information between groups is key. For this to happen groups need to be open to others, but such openness is in direct conflict with the need for community identity which ultimately results in exclusion. It may not be a bad thing for a community to close down while they absorb and discuss feedback from the outside among themselves. He noted that after World War II research “opened up” only to again become closed with the notion of peer reviewed science, and then to re-open with the World Wide Web. Perhaps when the issue of trust must be resolved the community will close again. Because of this cycle it is important to cultivate a culture of openness that supports community identity. Why? Because interesting things happen at interdisciplinary boundaries, so building productive boundaries between communities is an important goal. Neylon closed by saying that it is the shared values underpinning scholarship and the various ways in which we identify with the process of building knowledge that drive us forward. If we are to take advantage of change, we need to understand what it is that must stay the same.

Dr. Neylon’s slides are available on the NFAIS website.

### **3. Evaluating information - an issue of trust**

Regina Joseph, Founder of Sibylink and Co-founder of pytho, was the second speaker and her focus was on the use of quantified forecasting for detecting trends and assisting in decision making.

She began by saying that information has never been more accessible or in more demand, but it is simultaneously under attack. There are challenges to information veracity, there is mistrust, and there are complexities in archiving. We are also at a point in time when the technologies for fraud can outsmart the technologies for documenting truth; e.g. photo shop, Lyrebird (which can take audio files of a person's voice and generate recordings of that voice reading a script of anyone's choosing) [7], and Face2Face (which can do similar things with video) [8]. In the last fifteen years technology has changed how we obtain our information, especially in a world of social media. She noted that twenty-six percent of news retrieval is via social media, and while in 1983 there were about fifty companies who controlled mass media in the United States, media convergence has resulted in the fact that today ninety percent of that market is owned by only six companies. Our minds can be acted upon and we can willingly be more easily controlled. Unfortunately, what is stated as fact often reflects the bias of the person/organization who is delivering the information. Many news outlets have become distributors of opinions rather than providers of news. How can we differentiate between the two?

She went on to say that we are faced with a global "expertise paradox." Higher education and job requirements have traditionally resulted in specialization – students focus on becoming financial experts, chemists, etc., rather than developing broad-based knowledge. Our digital world certainly still requires areas of expertise, but also requires that a person have broad general knowledge that can help in sorting through the information with which we are presented on a daily basis. She noted that the Pew Research Center puts up a weekly news quiz online (see: <http://www.pewresearch.org/quiz/the-news-iq-quiz/>) and tabulates the results across age, gender, education levels, etc. She has been following this for about two years and has noted that there is an increased gap (in the double digits) between men and women in their ability to correctly answer basic questions about what is in the news. She commented that 57% of women use social media as opposed to 47% of males [9], and that we really need to learn more about how diverse delivery channels are shaping the news gap between genders. She believes that today we need an evidence-based system for information verification.

She noted that while we are all in awe of how IBM's Watson can extract ideas and concepts from large amounts of structured data it is the humans that are providing the information for Watson's use and there are cases where humans can outperform computers. She spoke at length about the four-year Aggregative Contingent Estimation (ACE) program, the goal of which is to "dramatically enhance the accuracy, precision, and timeliness of intelligence forecasts for a broad range of event types, through the development of advanced techniques that elicit, weight, and combine the judgments of many intelligence analysts" (see: <https://www.iarpa.gov/index.php/research-programs/ace>). The program was funded by the Intelligence Advanced Research Projects Activity (IARPA) from 2011–2015. There was a lot of skepticism about the program, but in the first year the test "generalists" who were using only publically-available information gave far more accurate (75%) predictions than those who had access to the government's closed classified information systems. The results continued throughout the program and everyone was surprised that open source indicators could be so powerful. She also talked about human-computer hybrid systems in which the best of computer analytics and human thinking can be combined. Hybrid approaches hold promise by combining the strengths of these two approaches while mitigating their individual weaknesses; e.g., humans get tired; computers do not! An example of this is another IARPA program, the Hybrid Forecasting Competition (HFC), that seeks to develop and test hybrid geopolitical forecasting systems - see: <https://www.iarpa.gov/index.php/research-programs/hfc>. In closing, she commented that we need better training on how to approach and analyze information and search systems that not only provide page ranking, but also provide indicators of information veracity. Joseph's comments supported Cameron Neylon's prediction that "trust" may be the next crisis in scholarly communication.

Joseph's slides are not available on the NFAIS website, but a brief article based upon her presentation appears elsewhere in the issue of *Information Services and Use*.

#### 4. Unpaywall – a subscription model alternative

The next speaker was **Jason Priem**, Co-founder of Impactstory, who was one of several speakers throughout the conference that spoke on Open Access. He began by the telling story of a test flight of a plane built by Samuel Pierpont Langley in 1903 [10]. The plane crashed into the Potomac River and the news made much of the event, predicting that airflight would never happen. Nine days later the Wright Brothers proved the naysayers wrong and everyone was surprised. Priem noted that if they had been following the field of aviation they could have foreseen eventual success as much had already been accomplished; e.g. in 1853 Cayley manned glider flies; in 1890 Clement Ader made the first powered takeoff [11]; in the 1890s Lilienthal made serial controlled flights in gliders [12]; in 1894 Chanute published *Progress in Flying Machines*; in 1902 St Louis Aeronautical Exposition took place; and earlier in 1903 Karl Jatho made several short powered flights [13]. Keeping abreast of information is important. Priem then turned to scholarly communication and said that he would focus on three things: (1) that in 2018 Open is the new default model for information access; (2) Value in the industry is now a level up for while Open Access destroyed some value it also has created value; and (3) Unpaywall, a free database of 18,062,575 scholarly articles helps create value now (see: <https://unpaywall.org/>). He then went on to talk about the current state of Open Access and the following numbers quoted are from a recent publication that he co-authored [14]. The study was based on all ninety million scholarly articles that have a DOI.

He said that 75% of all articles published before the mid-1990s are behind a firewall and that nearly half of the articles published in 2015 are Open Access. Based on this growth an aggressive forecast is that by the 2040 all articles will be Open Access. He then said that a more realistic forecast is that Open Access is steady-state from the mid-1990s through the year 2020, followed by an acceleration in the percentage of Open Access material due to a 2020 mandate, with the percent growth rate reaching a steady-state at 90% in the year 2030. Most important to keep in mind, however, is that the most used articles tend to be Open Access. He showed a graph based on the decline of toll-based articles and how it matches the decline of mules and horses *per capita* in the U.S.A. from 1900 to 1960 (the graph and the data behind it can be accessed at <http://bit.ly/horses-per-capita>). He reiterated that the most used articles tend to be Open Access – perhaps because people are reading more recent articles or because authors tend to publish their best results via open Access. Another measure of value is looking at what people cite, and he mentioned an article that had just been released [15] that looked at articles cited by Swiss researchers over the past two years. In 2015, 38% of the articles cited were Open Access and in 2016 that number jumped to 41%. Fee-based articles are not going away, they are simply declining as a percent of total articles published.

Priem then said that there is money to be made via Open Access and that is by adding value on top of the corpus of Open Access articles. He called this “moving up the abstraction stack;” i.e. moving from articles to groups of articles, and compared this to Eugene Garfield seeing the value of looking at a body of work (citations) and seeing the relationship across papers. He then talked about the database, Unpaywall. It contains structured data for every Crossref DOI (95 million articles); it has been built for copyright compliance from the ground up; it is accurate, with 98% precision and 75% recall compared to Google Scholar (this has been independently assessed); and it is updated weekly. It is used for Open Access Assessment efforts by the U.S. National Institutes of Health; it is used by browser-based access tools such as Kopernio; it is used by link resolvers at MIT, Harvard, by 1,500+ libraries (including the

British Library), and by value-add aggregation tools such as the Web of Science. He invited everyone to go try it out at: <https://unpaywall.org/>.

Priem's slides are available on the NFAIS website at: <https://nfais.memberclicks.net/assets/docs/ANCO2018/Jason%20Priem.pdf>.

## 5. Piracy - a form of Open Access

The next speaker was **Sari Frances**, Manager of Digital Licenses compliance at IEEE, who spoke about the impact of digital piracy on libraries and publishers. She used Sci-Hub [16] as a case study. Sci-Hub was founded by a Kazakhstani graduate student, Alexandra Elbakyan, in 2011, as a reaction to the high cost of research papers that reside behind paywalls. It is a website with more than sixty seven million academic papers and articles available for direct download. It bypasses publisher paywalls by allowing access through educational institution proxies which Frances said are accessed through compromised user credentials (in a 2016 Science article, Elbakyan denied that the credentials are stolen) [17]. Frances said that Sci-Hub is used because it is free, it fits in with the culture of openness and sharing, and it is very easy to use. It is quite popular and more than four hundred and fifty articles have been written about it since 2015. Sci-Hub has been sued by publishers such as Elsevier and the American Chemical Society and the courts have judged in their favor. However, it is unlikely that the publishers will see any payments as Sci-Hub has no assets in the United States.

Frances said that this type of hacking is going on every hour of every day. Publishers are losing money and such theft is undermining business models. She noted that IEEE is very diligent in monitoring Sci-Hub activity and regularly alerts universities if IEEE becomes aware that a university's security has been breached. The institutions have been very responsive when alerted and some universities have been aggressive in prohibiting access to Sci-Hub via their systems.

Frances then went on to discuss how publishers are responding to piracy and talked briefly about RA21, the Resource Access in the 21st Century initiative that was established in 2016 as a joint effort between the STM Association and the National Information Standards Organization (NISO). It aims to "optimize protocols across key stakeholder groups, with the goal of facilitating a seamless user experience for consumers of scientific communication [18]." The assumption is that if legal access to information is easier for students and researchers alike, piracy will decline. In closing she said that publishers are working together to do what they can to at least minimize piracy, if not eradicate it.

Frances's slides are not available on the NFAIS website; however, a brief article based upon her presentation appears elsewhere in the issue of *Information Services and Use*.

## 6. Kopernio - an antidote to information piracy

The next speaker, **Jan Reichelt**, Co-founder of Kopernio.com, provided support to Frances' comment that if legal access to information is easier for students and researchers alike, piracy will decline. Kopernio can help to do just that. He noted that 75% of the articles downloaded from Sci-Hub by students at the University of Utrecht were theirs to have legally, but the students could not easily get to them. He noted that the right to access an article does not mean that it is readily available. Libraries and publishers enter into legal agreements that permit access to and use of the publishers' journals by those who visit the library (physically or virtually). But there is no technology that gives "life" to the contracts. Users still

hit redirects, popups, and even firewalls. He asserted that it is not in the DNA of publishers to build such technology because they are not dealing directly with end users. They are dealing with institutions and the relationship is business-to-business. What is needed is a technology platform that is business-to-customer.

Reichelt views this as an opportunity. He said that there are between nine and ten million researchers that are the core group of users for STM publishers and each researcher requests about 250 PDF's per year. This means that there are 2.5 billion download requests annually and 2.5 billion opportunities to make an end user happy. But we do not - Sci-Hub does! We need to think about how we provide *convenience* for users, not just access.

Kopernio attempts to provide a guaranteed direct line between the user and the best possible version of the document that the user seeks. This is done by integrating a browser plug-in into the user's workflow – Web of Science, PubMed, Science Direct, etc. He said that there is a very popular social media site (he did not name it) that is widely-used by students and Kopernio is able to deliver 80% of their PDF requests by redirecting the requests back to their institution's holdings. Kopernio facilitates *convenient* access to journals to which universities subscribe. With one click users can tap into their university library holdings to retrieve articles and can also access free material that is held elsewhere (e.g. PubMed). There is no need to go to multiple platforms. Kopernio eliminates user frustration and ensures that publishers' journals are *legally* available from multiple platforms.

Reichelt closed by saying that Kopernio was not deliberately created to thwart Sci-Hub - it was created to provide users with convenient access to documents that they are permitted to use. (*Note: on 10 April 2018, shortly after the NFAIS Annual Conference, Clarivate Analytics announced that it had acquired Kopernio "to create the definitive publisher-neutral platform for research workflow and analysis for scientific researchers, publishers and institutions worldwide. Jan Reichelt has become Managing Director of Clarivate's Web of Science) [19].*

Reichelt's slides are not available on the NFAIS website.

## 7. The value of preprint servers

The final speaker of the day was **Shirley Decker-Lucke**, Publishing Director, SSRN, Elsevier, who talked about the value of preprint servers.

She began by saying that the scholarly world is under intense pressure to produce research that is open, accessible, collaborative, measurable, useful, and quickly shared. All of these demands are in addition to the work involved in the traditional research process: enabling research (strategy development, obtaining funding, establishing partnerships, etc.); doing the research (search, read, experiment, analyze, etc.); and sharing the research (publish, promote, etc.). So she posed a question: how can sharing early stage research/preprints be part of the solution to juggling all of the demands? Scholarly communication has undergone a lot of change in recent decades and preprints have existed throughout this period of change. She broadly defined a preprint as a document that exists prior to submission to a publisher and admitted that in the past people were skeptical of preprints - they want the version of record that has gone through peer review, editing, etc. But that perspective has changed. Preprints are now acceptable to most, but not all, journals; have citable DOIs; are creditable and viewed as valid; and they are versionable, archivable, and discoverable. She compared the version of record in a journal to fine dining – total perfection and expensive, while a preprint is convenience food – fast, easy, and cheap.

She then presented a chart on the growth of global preprint services. The first, arXiv, came out of Cornell University in 1991. As of her presentation it had 1,356,224 prints loaded in the fields of Physics,



Mathematics, and Computer Science. SSRN was launched in 1994. It is part of Elsevier and has 777,588 prints in Social Science and Economics/Multidisciplinary. There were a total of five preprint servers before the year 2000. Two more were added between 2000 and 2009 and two in 2013. But then growth accelerated with five preprint servers being launched in 2016, thirteen launched in 2017, and four already launched or planned for launch in 2018. Part of the growth is a result of changes in funding policies. Funders have a positive view of preprints and like to see them included in grant applications or at the end of a grant report. But other drivers of growth is that preprints are increasingly seen as proof of progress and central to scholarly sharing practices and that they are in alignment with recent sociocultural shifts in research: (1) the expectation of and a general cultural comfort with speed and ease over perfection; (2) a scrutiny of the peer review process and reproducibility concerns; and (3) a growing demand for free access to content.

She went on to say that there are a lot of benefits to authors when they share their early stage research via preprints. Their research is quickly disseminated globally and this can lead to feedback from and collaboration with other researchers. It demonstrates their productivity and independence while showcasing their scholarly output and research accomplishments. It allows them to claim priority over their discoveries and provides a vehicle for the sharing of research results not suitable for traditional journal publishing. She admitted that there are potential concerns around preprints including the dissemination of poor quality and irreproducible data, but that this can be mitigated by basic quality control.

Decker-Lucke said that SSRN takes a very broad approach to preprint content – from very short concept (idea) papers, working papers, conference proceedings and traditional preprints to papers under consideration by publishers and, with publisher permission, peer reviewed accepted manuscripts as they appear before other enhancements are made by the publisher. SSRN has 2.2 million users, 360 thousand authors, and have had 120 million downloads and have moved into chemistry, biology, and engineering. And they are always asking the following question: how early can they go in the research process? She said that they became part of Elsevier just about two years ago and this gave them access to more technology than ever, so they have been experimenting to answer that question.

One idea that they worked on was to determine if they could capture what users are currently researching/working on so that this information could be showcased much earlier in the researcher lifecycle. The premise was that by asking authors what they are currently working on, SSRN can harvest valuable information that will provide value both to SSRN and to their users. They measured the quantity and quality of the current ideas that they obtained through email and targeted authors in the fields of Biology, Economics/Finance, and Law through an email campaign to capture their current ideas. In August 2017 ten thousand emails were mailed using three different templates and this resulted in a 2.5% response rate (they were hoping for 5%). But while the quantity metric was poor, the responses that they received were of very high quality so they continued with the project. In October 2017 they added a feature by which the author can submit his/her idea directly to the server and this resulted in 1,654 ideas being submitted. One author put forth his idea and asked for input, so they will be adding additional features to support collaborative efforts.

Decker-Lucke closed by re-iterating that the scholarly world is under intense pressure to produce research that is open, accessible, collaborative, measurable, useful, and quickly shared. She said that she firmly believes that preprint servers in general and SSRN in particular help to address many aspects of this pressure. She noted that SSRN will continue to run experiments and explore ways to bring tomorrow's research (tomorrow's published journal article) to today (early stage research).

Decker-Lucke's slides are available on the NFAIS website.

## **8. Open Access NOW**

The morning session of the following day was opened by Dr. Ralf Schimmer, Deputy General Manager and Head of Information, Max Planck Digital Library, who, like others before him spoke on Open Access to scientific research. He opened by saying that there is velocity and turbulence in the information industry. This has been the environment since 2010 and is only accelerating. But there is inertia in the eye of the storm and that inertia is caused by a stagnant paywall system which, after fifteen years of Open Access movement, remains largely unaffected! Indeed, after more than a decade of global effort, paywall access and the subscription system are as prosperous as ever; only fifteen percent of content is immediately Open Access. He asserted that the paywall is the primary roadblock to openness, innovation, and sustainability in scientific communication.

Schimmer said that there are smart and innovative work-arounds to the current access and copyright limitations. As examples, He referred to Unpaywall and Kopernio, both of which have been mentioned earlier in this article. He noted that while these initiatives ease the symptoms and make our lives better, they cannot provide a cure for the disease. They are simply patching a broken system while expending enormous effort and growing in complexity. He said that real innovation will only come when energies can be focused on forward-looking solutions in an open environment.

He then referred to the “evil twins,” Sci-Hub and RA21, both of which were also mentioned earlier. He said that while these two seem to be diametrically opposed, they are actually twins and together are the epitome of what is wrong with the current system. Everyone uses Sci-Hub. It is an expression of end-user frustration, but is essentially tied to the paywall system. He asserted that RA21 was not requested and that it is unneeded and unwanted. It had already been defined in the Declarations of Budapest (2002) [20] and Berlin (2003) [21]. He firmly believes that Open Access is the only legitimate resource access in the 21st century and that an open system will provide opportunities for publishers who would continue to provide core services based on a transparent cost structure. He said that the paywall is equivalent to \$10 Billion U.S. dollars and that this money needs to be shifted to new business models. He believes that if we are to drive innovation and exploit technological opportunities, the subscription-based paywall system needs to be overcome as radically and quickly as possible. Open Access on a large scale can only be accomplished if we change the underlying business model of today’s scholarly journals and leave the subscription system behind.

Schimmer firmly believes that the way forward is via a new initiative entitled Open Access 2020 (OA2020 - see: <https://oa2020.org/>). This is a global alliance that is committed to accelerating the transition of Open Access. This is based upon the assumption that there is more than enough money in the system for Open Access to be sustainable. As of his presentation, one hundred and three institutions had signed on (individuals are not permitted to do so). Their goal is to transform a majority of today’s scholarly journals from subscription to Open Access publishing in accordance with community-specific publication preferences, and to pursue this transformation process by converting resources currently spent on journal subscriptions into funds to support sustainable Open Access business models.

Schimmer called subscriptions a “read access” model that is one-dimensional and is no longer good enough. Publishing and reading are two sides of the same coin. They are interrelated and need to be combined in library service level agreements with the publishers, for instance, through offsetting or publish and read models. He noted that all German Research organizations have joined OA2020.

He said that based upon the 2015 STM Annual Report, the annual revenue generated from English language STM journal publishing is estimated at about \$10 billion in 2013 [22] and that this translates into a spending level well in excess of \$5,000 per research paper through subscriptions. He closed by

saying that the Max Plank Library is committed to divesting of subscriptions and stated that “We have the leverage to bring down the \$5,000 per article we are putting on the table in the subscription system. By virtue of our own spending decisions we can drive Open Access into the system. We do not need further mandates for researchers, we need a mandate for our money.”

For more details see Schimmer’s slides on the NFAIS website.

## 9. Flipping from subscriptions to OA: easier said than done

The next speaker, **Michael Levine-Clark**, Dean of University Libraries, University of Denver, gave a very thought-provoking presentation on Open Access and what he thinks it will take to transition the publishing industry from a subscription-based model to Open Access. He admitted that his is a very U.S.-centric view, and opened with a question: How open is scholarly literature? On his first slide he showed the results of a study of three hundred randomly-selected articles and found that of them two hundred and sixty (87%) were available through Sci-Hub and only one hundred and sixty-six (55%) were available through some legal form of Open Access [23]. He then talked about the usual subscription agreement - at least for consortia and large schools, and that is the “Big Deal” [24]. He noted that these agreements are generally based on print spending and the journal “bundle” is a mix of journals to which an institution usually subscribes along with additional “free” titles. They are most often negotiated at the consortium level and each deal is constructed somewhat differently. These deals are difficult to disentangle and make it difficult to understand costs at the journal level and almost impossible at the article level.

Levine-Clark noted that many libraries assess the value of a journal subscription on usage and librarians tend to look at cost versus use. This assumes that all use is good use, but he said that the use of many articles could be a sign of inefficiency. Perhaps increased choice means more use, but less *critical* use. He noted that as long as we assume that this is the way to measure value, it’s very hard to move beyond subscriptions and does not believe that it will be easy to flip from a subscription-based world to an Open Access world. Why? The change will not be the result of flipping a switch, it will be a gradual transition. Because of the uneven distribution of subscription levels across universities and the publishing activities of researchers working at those institutions, library budgets will be impacted differently. Subscriptions cost money, but so does Open Access publishing when Article Processing Charges (APCs) must be paid by university researchers, and he provided some examples of this using the University of Denver and the California Institute of Technology as examples. Based upon his calculations, during a transition period from a subscription-based world to one of Open Access, Cal Tech would pay about \$3.1M in subscriptions and \$7.5M on APCs while the University of Denver would pay \$4.1M in subscriptions and \$1.3M in APCs. He said that the problems that must be overcome are the wide variety in pricing for subscriptions; the wide variety in research output; the fact that there is no central funding in US; and that fact that most academic libraries work through consortia for their subscriptions.

A potential solution is to renew Big Deals, but with some portion of the current subscription costs being applied to open up all articles by authors at participating institutions, including older articles. Across multiple consortia this would make a difference and after a predetermined time, there would have to be a plan to transition to a different model.

Levine-Clark closed by asking some questions: If everything is open - what about discovery? Will libraries pay more for discovery? Will publisher business models be built around enhancing access to open content with such features as profiling, metadata, discovery tools, enhanced user features? He sincerely

hopes that Open Access becomes a broader reality, but has some concerns about how soon that can really happen and what the ultimate result will be.

Levine-Clarke's slides are available on the NFAIS website and an article based upon his presentation appears elsewhere in this issue of *Information Services and Use*.

## 10. Funding open science: a bigger issue

The next speaker was **Katja Brose**, Science Program officer at the Chan Zuckerberg Science Initiative, who provided an overview of her organization and its efforts around scholarly communication and Open Science. She herself had worked at Elsevier for seventeen years and spent time as a researcher in neuroscience before entering the publishing world, so she approaches information from many perspectives – researcher, publisher, and now funder.

Her organization was founded about two years ago by Mark Zuckerberg and his wife, Priscilla Chan, upon the birth of their first child when they made a decision to use the vast majority of their wealth for philanthropic causes. It is not a charity. It is an LLC that can do advocacy work and venture investing, and it is built around advancing human potential and promoting equal opportunity through initiatives related to education, science, and advocacy.

The science division where she works was established just eighteen months ago and its mission is to cure, prevent or manage all diseases by the end of this century through investments in science technology and information. She knows that this is an aggressive long-term ambition and said that their focus is on basic science, largely in the field of biomedicine. They hope to fulfill their mission by (1) fostering collaboration between scientists, engineers, and clinicians; (2) enabling open tools (lab tools, computational tools, etc.) and technologies; and (3) building support for science – changing the culture of science to make Open Science the norm, and by improving the public perception of science as well as how scientists perceive themselves and their insular world.

Brose provided some examples of their initiatives. The first, Biohub, was really started by Mark and Priscilla before they established CZI and its goal is to support collaborative medical research in the San Francisco Bay area (see: <https://www.czbiohub.org/>). A second example is the Human Cell Atlas (see: <https://www.humancellatlas.org/>) which they did not establish, but with whom they have a partnership. Her group funds the development of tools for the project (all tools will be made openly-available); they are building the data platform; and their computational scientists work closely on the project. The objective is to “create comprehensive reference maps of all human cells - the fundamental units of life – as a basis for both understanding human health and diagnosing, monitoring and treating disease.”

She said that they are also trying to accelerate Open Science in a space that they are calling “Knowledge Environments.” They recently acquired an organization called Meta. This is a group that enables literature discovery by using artificial intelligence. They are actively supporting the preprint movement by funding and collaborating with bioRxiv, a preprint server for biology. They just started to work with Protocols.io, an open access repository for science methods, primarily in the life sciences. (The founder of this company, Lenny Teytelman, spoke at the NFAIS 2015 Annual Conference at the time that his company launched Protocols.io. An article on the service appeared in *Information Services and Use* [25].)

In closing, Brose said that we all need to work towards building tools and an environment for Open Science. We cannot only focus on the published article which is simply the end product. We need to think about open data and building platforms and repositories that are interoperable. We need to have dissemination plans in place for the open tools and resources that we build as well as for methods and

protocols. And we need a sustainable infrastructure, especially for data platforms that are built with grant funds. And while technology can be an issue, even more importantly we need cultural changes - we need to find a way of rewarding scientists who share.

The slides for this presentation are not available on the NFAIS website.

The next speaker, **Margaret Tait**, Research Associate, Robert Wood Johnson Foundation (RWJF), also spoke on the funding offered by her organization and their vision that “we, as a nation, will strive together to build a Culture of Health enabling all in our diverse society to lead healthier lives, now and for generations to come.” From 1972 to the present they have focused on improving health and the health care of all Americans. In 2014 they shifted to a Culture of Health vision and a broad focus on all that impacts health, and in 2015 they released the Culture of Health Action Framework. The framework, developed in collaboration with the RAND Corporation, sets a national agenda to improve health, equity and well-being. Informed by rigorous research on the multiple factors which affect health, it recognizes there are many ways to build a Culture of Health, and provides numerous entry points for all types of organizations to get involved.

They have a \$10 Billion dollar endowment and give out about \$66 Million in grants each year. She said that they are motivated by engagement with other funders and developed an interest in Open Access when the Gates Foundation said that efforts that they fund must be open. So, in October, 2015 RWJF convened a meeting in cooperation with the Scholarly Publishing and Academic Resources Coalition (SPARC). This meeting offered a unique opportunity for participants to share experiences, concerns, strategies, and questions regarding Open Access and Open Data. It included representatives from more than fifty organizations and resulted in the creation of the Open Research Funders Group, a partnership of funding organizations committed to the open sharing of research outputs (see: <http://www.orfg.org/about>). On 7 September 2016 the Group sent out a call for proposals for initiatives that would make research more transparent and accessible. One of six proposals that was funded was to convert the *Annual Review of Public Health*, a leading public health journal, to Open Access and to develop a sustainable model for other publications.

In closing Tait said that their plans for moving forward are to continue to explore a Foundation-wide policy; to put greater emphasis in entering into a dialogue with the publishing community; and to continue engagement with other funders and grantees. She said their role is not to lead, but rather it is to support their key stakeholders.

For more information, refer to Tait’s slides on the NFAIS website and take a look at the Robert Wood Johnson Foundation site as well (<https://www.rwjf.org/en/how-we-work/building-a-culture-of-health.html>).

## 11. Shark Tank Shoot-Out

The final session of the morning was a “Shark Tank Shoot Out,” in which three start-ups each had ten minutes to convince a panel of judges that their idea was worthy of potential funding (the “award” was actually a time slot for a future NFAIS Webinar). The session Moderator was Eric Swenson, Director, Product Management, Scopus, Elsevier, and the Judges were Jason Rollins, Senior Director of Innovation, Clarivate Analytics; Neil Kleinberg, Founder and CEO, DiliVer; and Andrea Michalek, Managing Director of Plum Analytics and Vice-President of Research Metrics Product Management, Elsevier.

The first speaker was **David Celano**, Business Product Manager, North America, SciencePod.

Celano said that research technology often gets buried in complicated words. SciencePod's objective is to be a "story teller" that makes science more accessible to a broader audience using clear, concise summaries that make specialized scientific and technological ideas understandable.

SciencePOD typically deliver bundles of content, including plain language summaries which help authors raise the profile of their work and collaborate with a broader community. They also produce magazine-style articles, infographics, and podcasts to showcase the most exciting research scholarly publishers produce in order to support the publisher's content marketing activities. They have a cloud-based solution that utilizes artificial intelligence and natural language processing to automate the process of creating content suitable for delivering digital and print publications on a very large scale. And they use a stable of science-educated writers who are paid on a piece-by-piece basis. They have a dual business model: one where they do the content creation and another soon-to-be-released software-as-a-service (SaaS) option where their clients can do their own content creation. The former model has a fee for the content bundles that are created; the latter is a standard SaaS model plus a per item percentage fee. The company is four years old and they have been profitable since day one. Customers include publishers and pharmaceutical companies. In 2017 they doubled their revenue over 2016. He noted that the science-related marketing industry will be 5.3 Billion Euros by 2019.

He admitted that they do have competitors and these include Raconteur (<https://www.raconteur.net/>), Content Central (<http://www.contentcentral.se/>), and Contently (<https://contently.com/>). The latter is headquartered in New York, while the others (including SciencePod) are European-based. SciencePod was initially funded by the Irish government. But Celano believes that SciencePod offers more capabilities, especially their smart magazine tool which none of their competitors offer. He closed by saying that SciencePOD provides their people, process, and technology platform to give their clients their own dedicated, agile, content creation team – all available at the click of a button! For more information go to: <https://sciencepod.net/splash>.

The second speaker was **Mads Holmen**, Founder of Biblio, a recommendation service founded in 2014 that helps publishers make the most of every visitor to their site by displaying relevant and engaging recommendations using Artificial Intelligence (AI). Their goal is to solve the discovery problem and give the right content to the right person at the right time. Holmen said that the ingredients of a good recommendation are the content, the user, and behavior. They use AI to quickly analyze those three ingredients in order to deliver the best recommendations possible. Biblio can be either an end-to-end solution or a tool kit to complement the software developed by their clients. They view the total market as seventy five million publishers/websites that run a publishing content management system and see their piece of the market as about ten million paying customers who have about ten thousand monthly visitors to their websites. At present they have seventy customers (publishers, media, libraries) and have a target of five hundred clients for 2018. They received \$1.4 M in funding at the end of 2016 to help them continue their work. Their revenue in 2017 was \$600 K and they hope that in 2018 they can increase that by 120–130%. They plan on doing a SeriesA funding at the end of this year. For more information see: <http://www.biblio.org/>.

The third speaker was **Craig Tashman**, CEO of LiquidText. Tashman opened by saying that research is the heart of knowledge work - but it is time-consuming. The process can consume 40% of a researcher's time gathering, reading, and distilling information in order to prepare their manuscripts and reports. He asked how computers help us and answered his own question by saying that they do not. People must navigate tons of information without tools that let them annotate, outline, and connect ideas while they do their research. He said that 80% of Knowledge Workers still prefer paper. He then gave a fascinating demo of what LiquidText can do. Basically, through intuitive interactions, it allows the user to compare

sections of a document by squeezing the document, pulling out key passages, organizing ideas, finding context, etc. You really need to see the video to appreciate its power (<https://liquidtext.net/product/>).

The first piece of the platform is a document reader app for iPad. Tashman said that it has been downloaded over a million times, Apple named it “Editors’ Choice” and the “Most Innovative iPad app” of the year when it was launched, and it has received glowing reviews from *MacWorld* to *Mashable* to *CIO Magazine*. The product was first launched in 2016 and that year monthly sales averaged \$5,100 with total sales in 2016 at \$61,300. They started advertising in 2017 and monthly sales averaged \$61,300 with total 2017 annual sales at \$610,000. They have two sales approaches, one for end users and one for businesses. The former is a freemium/subscription model that provides access to the core product on the web and permits collaboration, basic sharing, and retrieval. The price is \$50/year. The latter is a premium model that, in addition to the above, allows for enterprise management and internal sharing. The cost is \$120/sear/year.

He said that there are excellent tools that are niche tools for reading, annotating, etc., but there is really no other software tool that brings all these functions together seamlessly. He sees their competition as the workflow itself and the paper, pencils, browsers, windows, etc. that comprise the Knowledge Worker’s daily life – that is what LiquidText aims to replace.

For more information see: <https://liquidtext.net/>.

Later in the afternoon the judges announced that LiquidText was the winner of the Shoot Out. They will receive a plaque and the opportunity to present their business in a future NFAIS webinar. All of the slides used in the Shoot Out are available on the NFAIS website.

## 12. Members - only lunch: Open science and other NIH Initiatives

The next session was the Members Only lunch and the featured speaker was **Neil Thakur**, Special Assistant to the NIH Deputy Director for Extramural Research, who spoke on using Open Science to speed dissemination of research, reduce burden on researchers, and measure impact. He noted that there is a growing recognition that “interim research products” could speed the dissemination of science and enhance rigor. “Interim research products” are broadly-defined as complete, public research products that are not final. Preprints fall into this category. They are complete and public drafts of scientific documents. They speed research dissemination, establish priority, generate feedback, and may reduce publication bias. He said that many disciplines have been using preprints for years and there have been suggestions that expanding preprints could increase the impact of NIH research and ensure better science. He noted, however, that such change is occurring at different rates across scientific disciplines, and that NIH rules are narrow except for the reference section of applications. In addition to preprints, the pre-registration of protocols (i.e. publicly declaring key elements of a research project in advance) also falls into the interim research product category. As of March 2017 NIH guidelines state that interim research products now can be cited anywhere research products are cited, although DOI’s are required so that there is a sense of permanence. The full new grant guidelines on this issue can be accessed at: <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17-050.html>. These include Best Practices for repositories. He said that NIH received very strong support from the scientific community in favor of this change, although there were negative comments mainly related to the fact that preprints are not peer reviewed. So in the guidelines to reviewers NIH notes that interim research products are not peer-reviewed. He noted that NIH is neutral on whether reviewers should read references or not. During this process he found that 90% of the reviewers do look at what is cited in a grant.

Thakur then shifted topics and talked about the burden of many tasks that fall on researchers – one of which is filling out grant applications. He said that the process has:

- *Duplicative requirements*: Researchers have to curate and combine data that is scattered across public and private sources- ORCID, SCOPUS, PubMed, RPPRs, Vivo, Trellis, etc., and must do this multiple times in multiple systems.
- *Poor tracking and measurement tools*: Funders cannot track their impact on researcher careers, especially across different funders.
- *Inefficient research networks*: Researchers and associated groups do not use modern technology for networking and hiring (e.g., finding mentors, collaborators, employees, reviewers, etc.).
- *Bad incentives*: The current measures of research productivity do not adequately incentivize openness, rigor, and impact. Current fragmentation in research and career data and reporting makes it difficult to implement new measures.

NIH has established a series of goals to improve the overall Grant impact infrastructure and these are to:

- Track funder impact
- Encourage development of better productivity measures and incentives
- Support efficient collaboration and networking services
- Maintain researcher control and privacy
- Reduce researcher burden - facilitate more science, less paperwork

He said that one of the problems of the duplicative work is that there are a lot of information silos and there is no way to get them to talk to one another. Funding databases do not interact with university databases and there is no way to seamlessly pass information back and forth, so researchers are entering the same data over and over. He said that it would be great to have a CV Hub where all the information is linked. He noted that NIH has about 300 K scientists in their system whereas ORCID has 4.5 million people registered. ORCID also interacts seamlessly with the publishing community. He raised the question: Can we create a comprehensive research impact infrastructure using unique identifiers that will facilitate the seamless sharing of information across researchers, funders, publishers, and institutions using RRID, DOI, ORCID for products, DOI's for funding, ORCID for people, and Institutional identifiers as well?

As part of this vision NIH is looking at utilizing the publications tracking infrastructure (DOIs) to track grants in order to:

- Better track people across their careers and funding agencies
- More accurately identify research products
- Obtain more robust data to identify potential reviewers and assess conflicts of interest
- Validate grant /product associations

As an overlay, a universal funding number system for all funding agencies would be used to:

- Provide a 'common denominator' funding identifier format to harmonize NIH's grants system and contract system, and harmonize with other funders
- Create an inexpensive way for funding agencies to develop unique identifiers for their funding. This will require a permanent location for funding information

ORCID is enhancing their data model and third- party service integrations to:



- Broaden connections to research and career data usually reported on CVs
- Link researchers to funding and professional activities with verified and structured data
- Serve as an open hub for other systems
- Explore institutional identifiers

Phase one of ORCID integration into NIH systems has already been completed and they are now in phase two which will allow ORCID IDs to be incorporated into the profile section of NIH's electronic research administration system (eRA) in order to facilitate data exchange and funding/ORCID linkages. A third phase is planned for the future.

Thakur closed by saying that he firmly believes that if we can build this comprehensive research impact infrastructure using unique identifiers that will facilitate the seamless sharing of information across the diverse stakeholders in the scholarly community we will all be able to be more innovative in our work. NIH is working with ORCID, CrossRef, Publishers, and other Finding organizations to make this happen.

Thakur's slides are available on the NFAIS website.

### 13. Miles Conrad Lecture

The first afternoon session was the Miles Conrad Lecture. This presentation is given by the person selected by the NFAIS Board of Directors to receive the Miles Conrad Award - the organization's highest honor. This year's awardee was **Dr. C. Lee Giles**, David Reese Professor at the College of Information Sciences and Technology at Pennsylvania State University. He spoke on Artificial Intelligence (AI), defining it as machines that think, understand, reason rationally (although he noted that we can make machines think irrationally as well), make plans and decisions, and follow-through. The machine then re-evaluates the process and starts all over. He said that AI assists scientists (AI for the people) and AI can replace scientists (AI as the people), but most often it is a combination of the two. With regard to information services, AI assists in understanding and communicating knowledge using automated methods and operates at scale. It also automatically creates new knowledge in formal data structure. He briefly discussed machine reading and writing.

Giles also discussed scholarly Big Data and broadly defined it as all academic and scientific research documents - journals, books, theses, conference papers, technical reports, etc. He included presentations, experimental data, facts, formulae, code, and equations as associated data. Most of this information resides in large, sophisticated networks and those interested in the data include businesses, governments, social scientists, funders, policy makers, educators, economists, and scholars in general. He, together with some colleagues did a study to determine how much scholarly Big Data is available. They estimated that as of 2014 there were at least 114 million scholarly articles in English on the web, 24% of which were publicly- available. Google Scholar has at least 100 million articles. The study will be extended to distinguish the types of articles and include languages other than English.

AI and machine learning are used with scholarly Big Data to extract and link metadata, build knowledge structures, and process natural language queries. Giles et al. have developed the CiteSeerX system (<http://citeseerx.ist.psu.edu>) to perform some of these operations on the literature of computer science, such as author searching and name disambiguation; identification of tables in documents and extraction of the data from them; citation indexing; and full text indexing. He talked at length about the system and in closing said that AI is not a disruptor of information services - it simply makes the services easier to use.

Giles' slides are not available on the NFAIS web site.

#### **14. Libraries as Technology Innovators**

The final session of the day was a joint presentation by **Carl Grant**, Associate Dean, Knowledge Services & Chief Technology Officer, **Dave King**, Founder & CEO, Exaptive, Inc., and **Ken Parker**, CEO/Co-Founder, NextThought. The focus of the talk was on how the University of Oklahoma is applying new technologies in order to transform research and teaching in higher education. Grant joined the University of Oklahoma in February 2013 to fill a totally new position. Almost exactly three years later he spoke at the 2016 NFAIS Annual Conference to discuss what he was doing to at the University to improve information access and discovery in a world of information silos [26]. His current presentation demonstrated how much Grant has accomplished in his five years on the job. The University has been innovating with immersive visualization, 3D printing, microcontrollers, software, etc., and have put in place a suite of new tools for the transformation of scholarly communication. Working with innovative collaborative technology firms such as Exaptive, Inc. and NextThought, the University has extended that tool set and has set in motion the adoption of those tools across research libraries everywhere; e.g., classes and research are now regularly run not only locally, but also in virtual reality across wide geographic areas.

Grant opened his section of the presentation by saying that he wanted attendees to walk away understanding three things: (1) why information containers need to be further opened to unleash information's additional value; (2) The fact that major, new value creation is happening now on top of open information and products; and (3) that collaborative environments, physical and virtual, will fuel the creation of that value. He set the stage by talking about library budgets and how funding has changed over the years; e.g. in 2010 expenditures were 19% on databases, 24% on books, and 57% on journals. By 2017, those numbers were 23%, 10%, and 67% respectively. While budgets get tighter, the amount of information continues to explode. He said that by 2020, about 1.7 Mb of new information will be created every second for every human being on the planet. And our accumulated universe of data will grow to 44 trillion Gb [27]. He believes that Information is a commodity. Additional value is being locked away because information is "contained" in: database silos; document containers (PDF, .Docx, .AZW, etc.); behind paywalls; legal contracts/restrictions; systems with difficult access. This means that the additional value of information is being held back. It needs to be unleashed for librarians, for vendors, for society. He then said that we are all in the same boat and to create new value, we need to move together from viewing information as the source of value, to unleashing its full potential via virtual tools and physical spaces.

Grant then went on to talk about why libraries serve as great "laboratories" for innovation as a result of all the changes that librarians have had to face over the last fifteen years or more. He gave many fascinating examples of what the University of Oklahoma is doing now to provide new collaborative works spaces for faculty and students and to offer them new technologies that will enhance the overall educational experience.

Following Grant's presentation his two collaborators each spoke briefly about how they have been working with the University. Ken Parker spoke about the importance of "connections" in education; e.g. tutors, internships, apprentices, etc. and reviewed some of the learning tools and education platforms that his organization, NextThought offers (see: <https://nextthought.com/>), while Dave King discussed how his organization, Exaptive (see: <https://www.exaptive.com/>), has helped the University build complementary, cross-disciplinary teams in order to maximize the likelihood of innovative output. Grant closed the session with a summary of the three sections and thanked his collaborators.

All of the slides from this session are on the NFAIS website and a more detailed article by Grant based upon the three presentations appears elsewhere in this issue of *Information Services and Use*.

## 15. Practical applications of artificial intelligence and machine learning

The first session of the final day of the conference had three speakers discussing the practical applications of Artificial Intelligence. The opening speaker was supposed to be **Rajan Odayar**, Vice President, Head of Global Enterprise Management Solutions, Proquest, who was unable to attend due to illness. His replacement, **Mathew Devapiriyam**, Director of Technology, gave a very brief talk about chatbots and Proquest's internal use of system called Aristotle Analytics, currently in Beta version. The system is for the sales staff, but may eventually be a service. There was no update on the service that Rajan spoke about at the 2017 NFAIS Annual Conference and if you would like to know more about chatbots you can refer to the 2017 NFAIS Conference Overview [28].

The second speaker was **Ruth Pickering**, Co-Founder and Chief Business Development and Strategy Officer, Yewno, a company that offers a knowledge discovery search platform, Yewno Discover, that uses machine-learning and computational linguistics to analyze and extract concepts, and discern patterns and relationships in order to make large volumes of information more effectively understood through visual display (see: <https://about.yewno.com/>). (Note that Pickering participated in the Shark Tank Shootout in 2017, but Yewno was not the winner).

She opened with a brief description of Artificial Intelligence (AI) and where it stands in a long line of innovation, from the first written language in 3500BC, to the law Code of Hammurabi in 1790BC [29], all the way through to Mark Records in the 1960s [30], and the World Wide Web in the 1980's. She said that she would look at AI from two perspectives: (1) how it can help in finding information and (2) how it can expose content that has been hidden, perhaps due to poor search algorithms.

With regard to the first she made the point that searchers often cannot find what they are looking for because of the language required to do the search. You need to know the key concepts and terms and if it is an area with which one is unfamiliar that can be a problem. She added that even if you are an expert in a field, when you do a search you are faced with perhaps reams of search results that you must click-on, read, and determine if it is of value or not, then move on to the next result and start all over again. She raised the question: What if an AI platform could read text and present content along with a knowledge graph rather than present a long list of results? She also raised the point that publishers have an enormous amount of data and asked the question: what if an AI platform could delve into information, be it a book or a database, and identify what topics are covered and to what degree (e.g. chemistry comprises 10%). She then went on to demonstrate how Yewno Discover is an AI platform that does both. A description on the Yewno website states that "at the core of its technology is the framework that extracts, processes, links and represents atomic units of knowledge - concepts - from heterogeneous data sources. A Deep Learning Network continuously "reads" high-quality sources, projecting concepts into a multidimensional Conceptual Space where similarity measures along different dimensions are used to group together related concepts. In accord with prominent cognitive theories of conceptual spaces, our space allows for both geometrical, statistical and topological operations, and it permits to aggregate basic concepts into more complex representations."

Pickering's slides from both the 2017 and 2018 NFAIS Annual Conferences are available on the NFAIS website.

The final offering in this session was a joint presentation by **Jonathan Griffin**, Head of Product Development, IFIS Publishing, and **Jignesh Bhate**, Founder and CEO, Molecular Connections. They spoke on how IFIS Publishing was able to add value to its content and create new market segments using Big Data technologies developed by Molecular Connections.

Griffin began by saying that IFIS (originally known as the International Food Information Service, see: <https://www.ifis.org/>), is fifty years old this year and for the first forty seven years it focused on compiling food information for both the industrial and academic communities. By then, they came to realize that in order to grow they needed to do something. Database usage had gone flat as younger researchers prefer to use Google Scholar and IFIS does not have the resources to create and manage their own technology center. They had millions of complex records, but did not know how to maximize their usage so they reached out to Molecular Connections (see: <http://www.molecularconnections.com/>) with whom they had already worked to improve their indexing and who fully understood their content. After doing some market research across the academic, industrial, and government sectors it became very clear to both organizations that the food industry was having difficulty complying with international food regulations. There was no one-stop shopping in order to find regulations across all countries. Molecular Connections was already familiar with the IFIS content (one million abstracts, more than ten million metadata records, and almost fifty years of highly-curated, granular legacy data) so they easily were able to participate in the conceptualization of a new product that would fill the market need. The challenge was how to deal with unstructured text that was in different languages, different formats (PDF, HTML, etc.), and that was updated on irregular schedules around the world. Ultimately, Molecular Connections mined the legacy IFIS data and combined it with new data from the web, applied their Artificial Intelligence and Machine Learning techniques (including a human review by those with domain expertise to ensure quality), and added a linked data store to create a new database, Escalex (to see a video go to: [https://www.youtube.com/watch?v=r5LVL\\_Gg-IM](https://www.youtube.com/watch?v=r5LVL_Gg-IM)). Bhate noted that the database is user-friendly, has links to full text, and is customer centric. The combined IFIS - Molecular Connections vision is that they will release at least four new products within the next five years as a result of creating the Escalex database.

The slides for this presentation are not available on the NFAIS website. However, an article describing the Molecular Connection Technologies that were used to create the Escalex database appears elsewhere in this issue of *Information Services and Use*.

## 16. Building effective working communities

The next speaker was **Katherine Skinner**, Executive Director, Educopia Institute. Skinner noted that her organization, a non-profit founded in 2006, serves as a catalyst for collaboration among cultural, scientific, and scholarly institutions, Educopia's motto is "With others, you can accomplish what you cannot accomplish alone." Their mission is to build networks and collaborative communities to help cultural, scientific, and scholarly institutions achieve greater impact. Based on more than ten years of work with scholars, librarians, archivists, curators, and publishers in various fields, she shared her observations about the impetus, process, and impact of building and sustaining targeted cross-sector collaborative networks.

Skinner noted that a community of communities provides: (1) cohorts based on common vectors; (2) a network of common experiences (3) strategies and models for different stages of growth; (4) scaling of support services; and (5) healthy community development. Cross-sector networks are essential because

system-level change can only be effectively orchestrated through the deliberate work of *all* stakeholders across an entire system. She said that there are three perspectives that come into play on any issue or problem. The perspective of the individual; the perspective of the organization for which that individual works; and the perspective of the system in which that organization lives. All of these must be brought together, respected, managed, and facilitated so that trust can be developed and the “community of communities” can move forward together as they build an interdependence.

Educopia uses the Center for Creative Leadership (CCL)’s “Boundary-Spanning Leadership” model [31] to guide their facilitation efforts. And, as they facilitate multi-stakeholder groups, they also rely on many of the principles of “Collective Impact,” a methodology that swiftly rose to prominence in the social sector after the publication of a 2011 article by John Kania and Mark Kramer in the *Stanford Social Innovation Review* [32]. She presented several case studies, one involving the Software Preservation Network where they worked with gamers, lawyers, archivists, engineers, artists, etc., and another involving the Library Publishing Coalition where they worked with publishers, editors, librarians, students, administrators, etc. While each of these two organizations had very diverse stakeholders, in each case Educopia was able to bring them together to work effectively on a common cause. In closing, Skinner said that through their work they have found that groups of institutions acting in concert across fields and disciplinary boundaries accomplish more than any of the individual players could hope to do alone.

Skinner’s slides are available on the NFAIS website and a more detailed article based upon her presentation appears elsewhere in this issue of *Information Services and Use*.

## 17. A digital-first workflow

The next speaker was **Kristen Ratan**, Founder and Executive Director, Collaborative Knowledge Foundation (Coko, see: <https://coko.foundation/>) who described their work in building open source tools for a digital-first workflow in which all aspects of the editorial, peer review, and production processes are done in a collaborative workspace. Ratan said that the problem with research communication is that it is slow, expensive, incomplete, static, and closed. The first step towards change is to change academic publishing by utilizing digital (not print) workflows, by increasing automation, and by publishing all of the outputs from the research process - data, code, and protocols - in order to broaden access to that information. She said that there are three ingredients necessary to engineer change. The first is collaboration - we need to move from closed and linear workflows to collaborative workspaces. The second is cooperation - we need to move from proprietary platform silos to an Open Source ecosystem. And the third is community - we need to move from the garage to the town square model of product development.

She noted that the typical workflow today is that after research is completed a manuscript is first submitted into a manuscript control system such as Scholar One. After peer review and acceptance, it then goes into production (proofing, XML coding, etc.), and then is put into the publisher’s web delivery system (PDF, static HTML) and offsite repositories. This process can take from months to years for completion depending upon the scholarly discipline involved.

But in a collaborative workspace the manuscript (HTML format) is at the center, processes can be automated, and the tasks can be quickly completed. Ratan went on to say that no single platform can solve all the problems. We need an ecosystem of tools and software and we should build modular and interoperable tools. Reinforcing Skinner’s comments as noted in the prior presentation, it is *essential* that the community create and own solutions to today’s scholarly communication problems.

She added that such an ecosystem is emerging and that Coko's PubSweet publishing platform (a free, open source toolkit for building state-of-the-art publishing workflows, see: <https://pubsweet.org/>), enables innovation and cooperation. The platform has three use cases to date in three areas: book publishing, journal publishing, and in micropublishing. The book platform is Editoria which facilitates the efficient production of format-flexible, standards-compliant books (see: <https://editoria.pub/>) and current users are the University of California Press and the California Digital library. The journal platform is xPub and is being developed in co-partnership with eLIFE (to create a journal submission solution) and with Hindawi (to create a platform for its open access journals). Ratan closed by saying that Coko aims to enable publishers to move from closed and linear workflows to collaborative webspaces, and from proprietary platform silos to an open source ecosystem.

Ratan's slides are on the NFAIS website and a more detailed article based upon her presentation appears elsewhere in this issue of *Information Services and Use*.

## 18. The future of the book

The final speaker in this session was **Bob Stein**, Founder and Co-Director of the Institute for the Future of the Book Founder, The Voyager Company and a computer pioneer [33]. His talk was fascinating and I wish NFAIS could post it in addition to the slides so that all could hear him speak. Stein noted that in 1981 he wrote an article entitled *Encyclopedia Britannica & the Intellectual Tools of the Future*. This article lead him to Atari in 1982 where he created a series of drawings to demonstrate the technologies that he described in his paper. One of them is of a mother with her children sitting by a tidal pool and she is holding a wireless terminal with an antenna. The drawing was an attempt to demonstrate how the encyclopedia would operate in the future - it would be intelligent, you could ask it questions (sound familiar), etc. In 1984 he was given a laser disk and he ultimately co-founded the Criterion Collection, Inc., an American home video distribution company that focused on licensing important classic and contemporary films such as Citizen Kane and King Kong. In 1992 Stein was given a prototype of a CD player that could be connected to a computer and he immediately created the complete CD-Companion to Beethoven's 9th Symphony. It was the first viable commercial CD ROM that allowed you to learn everything you want to know about that piece of music. In 1991 he decided to do something fun with Shakespeare's *Macbeth*. The text was combined with a performance of the play by the Royal Shakespeare Company (in sync no less!) and a karaoke element was included so listeners could themselves perform a role along with the actors.

Stein went on through the years with fascinating products. He left Voyager in 1996 because he and some others believed that they needed a new technology to be more innovative with the book. In 1998 they launched TK3 Author, a set of powerful, flexible tools that allowed users to assemble text, images, sounds, and video into sophisticated interactive documents. It also let users annotate and personalize TK3 books in many ways - highlight passages, write notes on "stickies" that stay on the page, and copy text or other materials into a personal notebook. In 2004 the McArthur foundation asked Stein to come back into publishing which he did and he established the Institute for the Future of the book. In 2005 he published an online book entitled *Gamer*. He considers this to be the first networked book because when it was released readers around the globe began sharing comments in the margins. In 2007 he wrote *A Unified Field Theory of Publishing in the Networked Era* in which he proposed that a book is a place where things happen. Once published, the book does not end there. Publishing is only the initiation of a conversation when authors and readers begin to communicate with one another via commentaries in the margins. Stein

noted that his definition of the book has shifted throughout the years and he does believe that “social” reading in a digital world will grow. In closing he invited anyone who is interested in what he has done to reach out to him and he would gladly provide input and advice.

Stein’s slides are on the NFAIS website and my advice to any publisher looking into “social reading” is to contact him (email: [futureofthebook@gmail.com](mailto:futureofthebook@gmail.com)). His knowledge and depth of experience is amazing!

## 19. Lightning talks

The final session of the morning was a series of six lightening talks, each six minutes in length, on a topic of the speakers choosing. There was no specific theme.

### 19.1. *Libraries as abstracting and indexing services*

The first presenter was **Marjorie Hlava**, President, Access innovations, Inc., who spoke about libraries becoming abstracting and indexing services (A&I). Her company is in a project with the Smathers Libraries at the University of Florida where they are creating the Portal of Florida History. The project requires digitization of huge amounts of data which she termed the “new microfilm.” They have processed more than fourteen million pages to date. Since users must be able to get to the materials, they are expanding and enhancing metadata in order to increase discovery access to the digital collection. The library will need improved and consistent metadata practices moving forward. Hlava noted that catalogers are becoming metadata librarians. The take-away was that Libraries need to be vital. They need to concentrate on search and retrieval rather than solely on storage. They are a new generation of A&I services and as such need to invert their process to metadata-driven and discovery-enabled.

Hlava’s slides are available on the NFAIS website.

### 19.2. *Effective strategic planning and implementation*

The second speaker was **Michael Cairns**, Managing Director, Digital Prism Advisors, who talked about how to use customer and market insight to deliver new digital products and services that drive customer engagement and revenue. His organization helps clients identify, plan, and execute digital business strategies that open up new markets, more deeply engage customers, and inspire compelling new services (see: <https://www.dprism.com/>). He talked about strategy and the common pitfalls in execution such as lack of clarity; department and project silos; lack of transparency on project execution; shifting priorities; and accountability. He gave an example of a recent project with the American Institute of Architects, who prior to 2014 was a 155-year-old company with problems that included an old website with 7,000 pages that was plagued with poor search and navigation, a legacy infrastructure, and departmental silos. After completing their strategic planning and execution in 2017 they had become a 158-year-old company with a new responsive web site containing 700 pages and a recommendation engine. As a result, they had improved customer visibility and moved forward with a content governance process that is user-centric and personas-driven. The process was successful, but not without problems. Cairns noted that the overall success of these kinds of programs is linked to a variety of factors, including organization, management commitment and successful strategic planning. The lessons learned were:

- Wade, don’t jump - start small and incrementally
- Obtain senior leadership buy-in from the start

- Take an agile approach (3-week sprints) that forces prioritization and ongoing value delivery
- Train staff on how to use and customize visualization tools
- Develop standards for your website dashboards (layouts and color palette)
- Have self-service as a goal: make it easy for end-users to find and use the data they need
- Be prepared for change management and expect challenges when bringing people on board

Cairn's slides are available on the NFAIS website.

### 19.3. Crossref

The third speaker was **Jennifer Kemp**, head of Business Development, Crossref (see: <https://www.crossref.org/>), the organization that makes research outputs easy to find, cite, link, and assess through the use of a persistent identifier, the DOI (digital object identifier). She noted that DOIs are used to link clinical trials, event data, grant IDs, organizational IDs, peer review reports and other new content types. Kemp emphasized the importance of metadata, and noted that Crossref has 632 million metadata queries per month which represents a 28% increase over 2017. Some of the notable statistics she mentioned are that Crossref has: 9,686 member organizations; 94,981,451 registered content records; 65,272,832 records with links to full text; 2,568,142 records with funding information; 1,878,477 records with a Funder registry ID; 1,245,543 records with ORCID IDs; 30,047,715 records with licenses; and 6,444,793 Crossmark counts (Crossmark provides readers with access to the current status of a piece of content. With one click, you can see if content has been updated, corrected or retracted and access valuable additional metadata provided by the publisher). One could easily see why NIH is working with them as noted in the Neil Thakur's presentation at the NFAIS members-only lunch that was discussed earlier.

Kemp's slides are on the NFAIS website and a more detailed article based upon her presentation appears elsewhere in this issue of *Information Services and Use*.

### 19.4. CHORUS

The fourth speaker was **Susan Pastore**, Director of business Development, CHORUS, an organization that is committed to ensuring that the output from funded research is easily and permanently discoverable, accessible, and verifiable by anyone in the world (see: <https://www.chorusaccess.org/>). She discussed the new CHORUS Institution Dashboard Service that offers cost-effective access to article metadata and public access and archive information.

She said that what CHORUS offers maximizes system interoperability by employing widely-used standards and infrastructure; that the organization is policy-agnostic and supports a wide-spectrum of funder policies, Open Access business models, and diverse publishing platforms; and that CHORUS' goal is to broaden engagement among participants in the research ecosystem. They work with Crossref, ORCID, and other sources of persistent identifiers as well as trusted archives and more than fifty major publishers in order to provide cost-effective public access to funded research information. CHORUS is now working with academia in order to help faculty be compliant with funder requirements by utilizing existing author workflows, minimizing researchers' compliance efforts, streamlining technology, and offering a scalable solution. In addition, CHORUS lowers overhead by offering a low cost for academic institutions and libraries and by providing transparency through dashboard monitoring and reporting.

Pastore discussed several of the pilot projects that CHORUS launched in March 2017 with La Trobe University and the Australian government, in September 2017 with JST and Chiba University, and in



November 2017 with the Universities of Florida and Denver. She talked about the lessons learned to date, such as: accurate article metadata can be hard to come by; linking authors to a university is complex; faculty research is being deposited, it's just not necessarily compliant; preservation in perpetuity has value; researchers need help to comply with funding agency requirements; and researchers are confused by both their usage rights and funder obligations.

Pastore's slides are available on the NFAIS website.

#### 19.5. *Dimensions: the World's largest linked knowledge system*

The fifth speaker in this session was Ashlea Higgs, Founder, Über Research Part of Digital Science), whose mission is to build decision support solutions for science funding organizations (see: <https://www.uberresearch.com/>). Higgs discussed how a global collaborative effort within the scholarly community created the world's largest linked research knowledge system entitled Dimensions (see: <https://www.dimensions.ai/>). This system has gathered together in one place one hundred and twenty eight million grants, publications, citations, clinical trials, and patents along with four billion connections [34].

He said that Digital Science (see: <https://www.digital-science.com/>) and more than one hundred global research institutions have spent the better part of the last two years collaborating to solve three distinct challenges in the current research landscape:

##### (1) Research evaluation focuses almost exclusively on publications and citations data

Research evaluation tools are siloed in proprietary applications that rarely speak to one another. The gaps amongst proprietary data sources make generating a complete picture of funding impact extremely difficult (and expensive).

The goal of this collaboration amongst publishers, funders, research administrators, libraries, and Digital Science is to transform the research landscape by attempting to solve the problems resulting from expensive, siloed data research evaluation data.

Higg's slides are on the NFAIS website and a more detailed article based upon his presentation appears elsewhere in this issue of *Information Services and Use*.

#### 19.6. *Going Open Access: The Experience of the Routledge, Taylor & Francis Group*

The sixth and final speaker in this session was **Joseph Lerro**, Open Access Sales Executive, Routledge, Taylor & Francis Group. He noted that the global Open Access movement has undertaken a proposal to 'flip' from the traditional subscription model to an Open Access model (as we have already heard from many of the conference speakers). But common sense must prevail. With so many stakeholders involved, such a transition *must* appeal to the interests of researchers, librarians, funders, and publishers. Likewise, with an array of Open Access models, it is important to determine which one will be the *most* effective solution in the long term. Lerro said that Taylor & Francis is taking a flexible, evidence-based approach to this transition, piloting a variety of models, and he openly discussed their experiences in flipping journals from hybrid to full Open Access. In addition to converting twenty-eight subscription journals to full Open Access since January 2017, Taylor & Francis has established Open Access agreements with organizations such as the Max Planck Digital Library (from whom we heard in the first session on the second day of the conference with Dr. Ralf Schimmer's talk) and the VSNU Dutch Library Consortium. Lerro went on

to provide specific examples of the effects and implications of transitioning to Open Access for a global publisher.

Lerro's slides are on the NFAIS website and a more detailed article based upon his detailed presentation appears elsewhere in this issue of *Information Services and Use*.

## 20. Final keynote: academic publishing, blockchain, and shifting roles in a rapidly changing world

The final keynote speaker was **Dr. Joris van Rossum**, Director Special Projects, Digital Science, who discussed the opportunities and challenges that Blockchain technology offers within the broader context of the evolving roles of academic publishers in a world characterized by revolutionary technological changes. The NFAIS audience had been introduced to Blockchain technology during the 2017 conference when Christopher E. Wilmer, Managing Editor of *Ledger*, talked about cryptocurrency (*Ledger* is a peer-reviewed journal for publishing original research on cryptocurrency-related subjects) [35]. The final keynote went beyond Blockchain's role in supporting Bitcoin and discussed the fascinating role this technology might take in the publishing arena.

Van Rossum opened his presentation with a 1956 quote from Aldous Huxley, "Our technology produces a state of chronic revolution," followed by a brief overview of the milestones that scholarly communication has experienced - from the printing press through the web. He noted how long it took for various technologies to be embraced by one hundred million users: the telephone, seventy years; radio, forty years; television, thirteen years; the internet, four years; and Facebook, three-and-a-half years. Things are going incredibly fast!

He noted that the journal was started in the 17th century and that the processes have not changed all that much although technology has advanced since the printing press. He asserted that the role of publishers is to support researchers in what they can't do, or in what they don't feel like doing themselves. This is our guiding principle. The functions performed by publishers are: (1) registration (establishing the author's precedence and ownership of an idea); (2) certification (ensuring quality control by peer review); (3) dissemination (communicating the findings to the relevant audience); and (4) preservation (preserving a fixed version for a future reference and citation). Publishers are not technology companies and he noted that sometimes we are successful despite our technology. He asserted that publishers have partnered with authors for centuries and that we provide knowledge and services.

But that is not to say that things will stay the same. In fact, he said that the publisher's role is getting smaller as alternatives for the fulfillment of the publisher functions have emerged. Preprint services such as ArXiv.org, bioRxiv, and ChemRxiv are assuming the role of registration, as is figshare. Dspace and CLOCKSS are taking on the preservation role. Google has been a major disruptor in the dissemination of information as has Sci-Hub as we heard earlier in the conference. He also talked about ResearchGate as an information disseminator - not via subscriptions, but via social networking. He said that he is not taking a stance on whether or not Sci-Hub and ResearchGate are information "pirates," but rather wants to point out that if indeed our role is to support researchers in what they can't do, or in what they don't feel like doing themselves, we need to stop and ask - is there an new way of disseminating content and are we really needed? Is our traditional role better performed by others? It is not for us to impede researchers if there is a better dissemination model.

He then moved on to the publisher's role of certification because it is this role that has serious challenges in three areas: (1) reproducibility of search results; (2) peer review; and (3) metrics. With regard to the reproducibility of research results he noted that a recent study showed that only about one third of research

results could be duplicated and that more than fifty-percent of researchers today say that we have a crisis on our hands [36]. We have an issue of trust that is reinforced by the publishing of fake news – whether it be politics or science. With regard to peer review there is a lack of transparency (the reviews are not published) and those who do the reviewing are not recognized for their efforts. Other problems with the system have been noted as well, not the least of which is the variations in the thoroughness of the reviews themselves [37].

The final problem is the metrics. He noted that you can only reward people for what you know about them - how many papers they have published, how often they have been cited, etc. All of this happens *after* the paper is disseminated. What went into the research (study design, experiments, analysis, peer review, etc.) is all unknown. Our current metrics are limited and outdated and are tied to the print world. His conclusion regarding the current state of affairs is that the publisher's role outside of certification is becoming smaller, and in parallel while certification is becoming increasingly relevant in today's world, it faces serious challenges. It is in the area of certification where Blockchain technology can bring significant improvement.

Van Rossum said that he will discuss Blockchain technology on three levels. First, it is the underlying encryption technology used for cryptocurrencies such as Bitcoin and therefore could serve as a currency for science. When a scientist publishes a paper, performs peer review, etc. they could be given a “token” as a reward that in turn could be used to “buy” other services; e.g. journals, statistical analysis, etc. A closed economy could be built around functions that are performed in the science community and he has seen a number of initiatives emerge since Digital Science released their Blockchain report last November [38].

On the second level, use of Blockchain technology would move us from an Internet of information to an Internet of value. He explained that the Internet is great for information dissemination, but when he sends a copy of a something to someone over the Internet it is just that - a copy in code that can be rendered upon receipt. But he still has his copy. Unless it is a fraud, when you give someone a twenty dollar bill you truly give up ownership of that amount of funds. Blockchain ensures that any currency is truly transferred. It establishes ownership, prevents double spending, and allows for the exchange of value without the use of an intermediary such as a bank. He said that the technology is perfect for Digital Rights Management and that micropayments open way for a new business model in publishing. He noted that this is already happening via Katalysis, an organization that plans on democratizing the value of online content using Blockchain technology (see: <https://www.katalysis.io/>). He added that perhaps this technology could eventually replace the journal subscription model.

He said the third level is where it gets even more interesting because Blockchain technology can serve as a new form of database. It is a very special kind of data storage: decentralized; shared and immutable; and transparent, but pseudonymous. It could very well support a single repository for scientific research that would eliminate the certification challenges mentioned earlier, allowing for advanced metrics, transparency, validation, and reproducibility.

Van Rossum then went on to describe a new Digital Science initiative that he is leading and that was officially announced shortly after the NFAIS conference. It is pilot project for the development of a protocol where information about peer review activities (submitted by publishers) will be stored on a blockchain. This will allow the review process to be independently validated, and data to be fed to relevant vehicles to ensure recognition and validation for reviewers. By sharing peer review information, while adhering to laws on privacy, data protection and confidentiality, the project is hoped to foster innovation and increase interoperability.

He said that the advantages for reviewers is improved recognition and more targeted invitations to review. The advantages to editors are that there will be reviewer finding tools to use against a database of complete reviewer profiles so that they will be able to more easily identify qualified reviewers and hopefully get higher acceptance rates. And the advantages for publishers are that the obstacles in the review process will be eliminated; there will be better demonstration and justification of the publisher's role; and there will be more transparency in the process, hopefully resulting in increased trust. He noted that the project will result in a data store of review information for a select group of journals. Information will be sent to ORCID and the entire process will be tracked and audited.

In closing, van Rossum invited everyone who was interested to participate in the project. As of his presentation, the pilot included Digital Science, ORCID, Katalysis, and Springer Nature. The Taylor and Francis Group and Cambridge University Press have since joined. You can learn more about the pilot by visiting their website: <https://www.blockchainpeerreview.org/>.

Van Rossum's slides are on the NFAIS website and a more detailed article based upon his presentation appears elsewhere in this issue of *Information Services and Use*.

## 21. Conclusion

From Cameron Neylon's insightful opening keynote on the drivers of change and the fact that there will always be another "crisis" in scholarly communication to the closing keynote on the diminishing roles of publishers and the potential positive impact that blockchain technology offers for the future of scholarly communication, the conference put a lot on the table to think about. How easy will it be to "flip" from a fee-based journal subscription to Open Access? Michael Levine-Clark presented a well-thought out rationale for why this change will be a gradual transition rather than a short-term one as discussed by Dr. Ralf Schimmer. While Jason Priem supported Levine-Clark's perspective through his prediction that based on the current trajectory, it will not be until 2030 that 90% of scholarly journals will be Open Access, Joseph Lerro demonstrated that the switch can be flipped if a loss of revenue can be absorbed. I suspect that the debate will continue and while it does the current state of affairs encourages the creation and use of alternative services such as preprints.

Indeed, what about the role of preprint servers as we move forward? I found Decker-Lucke's presentation fascinating. The growth of preprint servers has accelerated - from a total of five preprint servers before the year 2000 to approximately thirty-one in 2018, with twenty-two of those being launched between 2016 and 2018! She contests that part of the growth is a result of changes in funding policies as funders now have a positive view of preprints and encourage that they be included in grant applications or at the end of a grant report. Her perspective was completely supported by Neil Thakur in his presentation on NIH funding guidelines. But it is one of the questions that Decker-Lucke posed that really caught my attention: How early in the research process can we go to capture information? Her organization is actively gathering information at the concept stage of research - what ideas are being worked on - and my guess is that others will follow.

Also, from many of the speakers' perspectives it seems that the scholarly community really must start addressing the issue of trust. Neylon said that trust might be the next "crisis" and Regina Joseph presented a very good case in support of Neylon's comment with her discussion of media bias and the need to better educate students on how to approach and understand information, and the need to build search and retrieval systems that not only rank pages, but also provide veracity indicators. Both Shirley Decker-Lucke and Joris van Rossum also raised the issue of trust in relation to the peer review process.

The Open Access - Open Science discussions and the related presentations on funding policies and technology solutions were excellent, but I believe that Katja Brose made a point of which we should not lose sight. She asserted that Open Science needs to be given more attention and that we are too focused on the end product, the article. To a certain extent that point was also made by Kristen Ratan in her call for interoperable systems. The practical examples such as those given by Jonathan Griffin and Jignesh Bhate on how to breathe new life into legacy content and by Carl Grant on how to do the same for a university library were equally compelling as was Katherine Skinner's discussion on building effective communities across diverse stakeholders (perhaps those of us in scholarly communication should seek Educopia's guidance).

What has made the NFAIS conferences so interesting and valuable over the years is that NFAIS provides a *neutral* venue in which controversial issues can be discussed productively and with respect for differing opinions, and this year was no different. In listening to the back-and-forth conversation, what was interesting was that none of the speakers were complaining. They were stating facts from their perspective and it appears that all are actually involved in *doing* things to make science better.

I leave you with the following two quotes with which Carl Grant opened his presentation. One is from President Theodore Roosevelt: "*Complaining about a problem without posing a solution is called whining.*" A second is from President Barak Obama: "*Change will not come if we wait for some other person, some other time. We are the ones we have been waiting for. We are the change we seek.*"

My takeaway from the conference is that the scholarly community *has* stopped whining about Open Access/Open Science and that everyone is working, albeit within the constraints of their businesses, missions, etc., to evolve, to build collaborative partnerships and to move forward with the goal of building an open and collaborative scholarly community. Congratulations, NFAIS on your 60th Anniversary Conference - it was one of your best!

Plan on attending the 2019 NFAIS Annual Conference that will take place in Alexandria, VA, from February 13–15, 2019. Watch for details on the NFAIS website at: <http://www.nfaais.org/>.

*Note:* If permission was given to post them, speaker slides used during the NFAIS 2018 Conference are embedded within the conference program at: <http://www.nfaais.org/2018-conference-program>. The term "slides," if they are available, is highlighted in blue.

## 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. Prior to NFAIS, Bonnie was Senior Vice President and General Manager of ProQuest's Library Division where she was responsible for the development and worldwide sales and marketing of their products to academic, public, and government libraries. Before ProQuest, Bonnie was Executive Vice President, Database Publishing at the Institute for Scientific Information (ISI - now Clarivate Analytics) where she was responsible for product development, production, publisher relations, editorial content, and worldwide sales and marketing of all of ISI's products and services. She is a Fellow and active member of the American Chemical Society and a member of the Bureau of the International Union of Pure and Applied Chemistry for which she chairs their Publications and Cheminformatics Data Standards Committee. She is also on the Board of the Philosopher's Information Center, the producer of the *Philosopher's Index*, and she serves as a member of the Editorial Advisory Board for *Information Services*

and Use. She has served as a Board and Executive Committee Member of the former Information Industry Association (IIA), as a Board Member of the American Society for Information Science & Technology (ASIS&T), and as a Board member of LYRASIS, one of the major library consortia in the United States.

Ms. Lawlor earned a B.S. in Chemistry from Chestnut Hill College (Philadelphia), an M.S. in chemistry from St. Joseph's University (Philadelphia), and an MBA from the Wharton School, (University of Pennsylvania). Contact: [chescot@aol.com](mailto:chescot@aol.com).

## About NFAIS

Founded in 1958, the National Federation of Advanced Information Services (NFAIS™) is a global, non-profit, volunteer-powered membership organization that serves the information community; i.e., all those who create, aggregate, organize, and otherwise provide ease-of-access to and effective navigation and use of authoritative, credible information.

Member organizations represent a cross-section of content and technology providers, including database creators, publishers, libraries, host systems, information technology developers, content management providers, and other related groups. They embody a true partnership of commercial, nonprofit, and government organizations that embraces a common mission - to build the world's knowledgebase through enabling research and managing the flow of scholarly communication.

NFAIS exists to promote the success of its members and for sixty years has provided a forum in which to address common interests through education and advocacy.

## References

- [1] *Issues for Science and Engineering Researchers in the Digital Age*, National Academy Press, Washington, DC, 2001, <https://www.ncbi.nlm.nih.gov/pubmed/24967477> (accessed 3 July 2018).
- [2] <http://www.budapestopenaccessinitiative.org/read> (last accessed 3 July 2018).
- [3] <https://openaccess.mpg.de/Berlin-Declaration> (last accessed 3 July 2018).
- [4] D.J. de Solla Price, *Little Science, Big Science*. Columbia University Press, New York, 1963.
- [5] E. Schofer and J.W. Meyer, The world expansion of higher education in the twentieth century, *American Sociological Review* **70**(6) (2005), 896–920 <http://faculty.sites.uci.edu/schofer/files/2011/03/Schofer-Meyer-Higher-Education-ASR.pdf> (last accessed 3 July 2018).
- [6] L. Fleck (author), T.J. Trenn (editor/translator), T.K. Merton (editor), F. Bradley (translator), *Genesis and Development of a Scientific Fact*, University of Chicago Press, 1979.
- [7] Produced by researchers at the University of Montreal, Canada, Lyrebird demos have featured fabricated “conversations” between Bill Clinton, Barack Obama and Donald Trump. <https://lyrebird.ai/> (last accessed 3 July 2018).
- [8] Produced by researchers at Stanford University, The Max Planck institute and the University of Erlangen-Nuremberg, Face2Face is described as “real-time face capture and reenactment of RGB videos.” <https://www.youtube.com/watch?v=ohmajJTcpNk> (last accessed 3 July 2018).
- [9] E. Shearer and J. Gottfried, *News Use Across Social Media Platforms*, Pew Research Center, 7 September 2017, <http://wqad.com/2018/02/13/study-women-more-reliant-on-social-media-for-getting-news/> (last accessed 3 July 2018).
- [10] Wikipedia, [https://en.wikipedia.org/wiki/Samuel\\_Pierpont\\_Langley](https://en.wikipedia.org/wiki/Samuel_Pierpont_Langley) (last accessed 3 July 2018).
- [11] Wikipedia, [https://en.wikipedia.org/wiki/Cl%C3%A9ment\\_Ader](https://en.wikipedia.org/wiki/Cl%C3%A9ment_Ader) (last accessed 3 July 2018).
- [12] Wikipedia, [https://en.wikipedia.org/wiki/Otto\\_Lilienthal](https://en.wikipedia.org/wiki/Otto_Lilienthal) (last accessed 3 July 2018).
- [13] Wikipedia, [https://en.wikipedia.org/wiki/Karl\\_Jatho](https://en.wikipedia.org/wiki/Karl_Jatho) (last accessed 3 July 2018).
- [14] H. Piwowar, J. Priem, V. Larivière, J.P. Alperin, L. Matthias, B. Norlander, A. Farley, J. West and S. Haustein, The state of OA: A large-scale analysis of the prevalence and impact of OA articles, *PeerJ* (2018), PubMed 29456894.

- [15] F.S. Muller and P. Iriarte, , Measuring the Impact of Piracy and Open Access on the Academic Library Services, 15th Interlending and Document Supply Conference (ILDS), Paris, France, October 2 – 6, 2017, <https://archive-ouverte.unige.ch/unige:102345> (last accessed 3 July 2018).
- [16] Wikipedia, <https://en.wikipedia.org/wiki/Sci-Hub> (last accessed 3 July 2018).
- [17] J. Bohannon, Who's downloading pirated papers? Everyone, *Science* (2016) <http://www.sciencemag.org/news/2016/04/whos-downloading-pirated-papers-everyone> (last accessed 3 July 2018).
- [18] See: <http://www.stm-assoc.org/standards-technology/ra21-resource-access-21st-century/> (last accessed 3 July 2018).
- [19] H. Else, Web of science owner nays tool that offers one-click access to journal articles, *Nature.com.*, April 10, 2018, <https://www.nature.com/articles/d41586-018-04414-8> (last accessed on 3 July 2018).
- [20] <http://www.budapestopenaccessinitiative.org/read> (last accessed 3 July 2018).
- [21] <https://openaccess.mpg.de/Berlin-Declaration> (last accessed 3 July 2018).
- [22] M. Ware and N. Mabe, *The STM Report: An Overview of Scientific and Scholarly Journal Publishing*, 4th ed. International Association of Scientific, Technical and Medical Publishers, 2015.
- [23] M. Levine-Clark, J. McDonald and J. Price, Availability of freely available articles from gold, green, rogue, and pirated sources: How do library knowledge bases stack up?, *Electronic Resources & Libraries* (2017).
- [24] R. Poynder, The big deal: Not price, but cost, *Information Today* (September 2011) <http://www.infotoday.com/it/sep11/The-Big-Deal-Not-Price-But-Cost.shtml>, (last accessed 3 July 2018).
- [25] L. Teytelman and A. Stoliartchouk, Protocols.io: Reducing the knowledge that perishes because we do not publish it, *Information Services and Use* **35**(1-2) (2015), 109–115 <https://content.iospress.com/journals/information-services-and-use/35/1-2?start=10> (last accessed 3 July 2018).
- [26] G. Grant, Supporting a passion for new ideas through open APIs, *Information Services and Use* **36**(1-2) (2016), 65–72 <https://content.iospress.com/download/information-services-and-use/isu798?id=information-services-and-use%2Fisu798> (last accessed 3 July 2018).
- [27] B. Marr, 20 mind-boggling facts that everyone must read, *Forbes* (30 September 2015) <https://www.forbes.com/sites/bernardmarr/2015/09/30/big-data-20-mind-boggling-facts-everyone-must-read/1a53bef817b1> (last accessed 3 July 2018).
- [28] B. Lawlor, An overview of the NFAIS 2017 Annual Conference: The big pivot: re-engineering scholarly communication, *Information Services and Use* **37**(3) (2017), 299 <https://content.iospress.com/journals/information-services-and-use/37/3?start=10> (last accessed 3 July 2018).
- [29] Wikipedia, <https://en.wikipedia.org/wiki/Hammurabi> (last accessed 3 July 2018).
- [30] Wikipedia, [https://en.wikipedia.org/wiki/MARC\\_standards](https://en.wikipedia.org/wiki/MARC_standards) (last accessed 3 July 2018).
- [31] J. Yip, C. Ernst and N. Campbell, *Boundary Spanning Leadership*, Center for Creative Leadership, 2016, <https://www.ccl.org/wp-content/uploads/2015/04/BoundarySpanningLeadership.pdf> (last accessed 3 July 2018).
- [32] J. Kania and M. Kramer, Collective impact, *Stanford Social Innovation Review* (Winter 2011) [https://ssir.org/articles/entry/collective\\_impact](https://ssir.org/articles/entry/collective_impact) (last accessed 3 July 2018).
- [33] Wikipedia, [https://en.wikipedia.org/wiki/Robert\\_Stein\\_\(computer\\_pioneer\)](https://en.wikipedia.org/wiki/Robert_Stein_(computer_pioneer)) (last accessed 3 July 2018).
- [34] G. Bode, C. Herzog, D. Hook and R. McGrath, A guide to the dimensions data approach, *Digital Science* (January 2018) <https://www.digital-science.com/resources/portfolio-reports/a-guide-to-the-dimensions-data-approach/> (last accessed 3 July 2018).
- [35] B. Lawlor, An overview of the NFAIS 2017 Annual Conference: The big pivot: re-engineering scholarly communication, *Information Services and Use* **37**(3) (2017), 300 <https://content.iospress.com/journals/information-services-and-use/37/3?start=10> (last accessed 3 July 2018).
- [36] M. Baker, 1,500 scientists lift the lid on reproducibility, *Nature* (2016) <https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970> (last accessed 3 July 2018).
- [37] G. Kabat, The crisis of peer review, *Forbes* (23 November 2015) <https://www.forbes.com/sites/geoffreykabat/2015/11/23/the-crisis-of-peer-review/2/409ef3696981> (last accessed 3 July 2018).
- [38] J. Van Rossum, *Blockchain for Research: Perspectives on a new Paradigm for Scholarly Communication*, Digital Science Report, November 2017, [https://figshare.com/articles/\\_/5607778](https://figshare.com/articles/_/5607778) (last accessed 3 July 2018).