

An overview of the NFAIS 2019 Annual Conference: Creating strategic solutions in a technology-driven marketplace

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 2019 NFAIS Annual Conference, *Creating Strategic Solutions in a Technology-Driven Marketplace* that was held in Alexandria, VA from February 13 - February 15, 2019. The goal of the conference was to focus on how technological innovations, especially Artificial Intelligence and machine learning, along with changing market demands, are creating new opportunities within the information community. Speakers were invited to demonstrate that such innovations have the potential to provide researchers with new tools with which to advance their quest for scientific discovery and also have the potential to provide the much-needed insights to assist business leaders in making their strategic decisions with confidence. The diverse speakers made their point - technology is driving us forward. But the real message of the conference was all about the value of content and how that value can be increased by leveraging appropriate technology. Like changing a rough stone into an incomparable diamond, technology can transform traditional content into a faceted gem!

Keywords: semantic tagging, semantic search, information discovery, open access, open science, university presses, pre-print servers, Plan S, artificial intelligence, blockchain technology, born-digital scholarship, internet-in-a box, scholarly communication, scholarly publishing, academic research libraries, library consortia

1. Introduction

Technology has been transforming scholarly communication for centuries - slowly at first and much more rapidly over the past century. From proto-writing [1] in the early days of man, through to the invention of paper [2], the printing press [3], computers [4], the Internet [5], and now blockchain technology [6], innovative technologies have been disrupting how information is created, disseminated, managed, stored, and even perceived. With each change has come opportunity for some and ultimate irrelevance for those too focused on their traditional world to understand the potential impact of change.

Today's information technology has created a landscape that is now ripe for a major change in how the scholarly communication process will proceed into the future. With the launch of digital publications in the mid-1990s, information has become much more widely and easily disseminated. These digital publications ultimately met the needs of the born-digital generation that developed from the launch of personal computers in the 1980s - a generation that by the late 1990s began to demand free access to

*E-mail: chescot@aol.com.

digital information. That demand has since been heard and supported by research funders around the globe who now require that the outcomes of the research that they fund be made publically available. And the researchers, who are also authors and users of information, expect that such free sharing of their scientific data be allowed. The stars are now well aligned for change!

Indeed, gone is the traditional world of scholarly communication and publishing where publishers held the reigns. At the 2018 NFAIS Annual Conference Dr. Joris van Rossum, Director Special Projects, Digital Science, noted that the publisher's role is getting smaller as alternatives for the fulfillment of the publisher functions have emerged [7]. Indeed, his comment applies to all content providers - from primary publishers through to the abstracting and indexing services - for today advances in technology have provided information alternatives that are often "good enough" for the majority of those that seek it.

Now that the stars are aligned how will our world change as change it will - process has already begun. This question was raised several times during the 2019 NFAIS Annual Conference that took place earlier this year when a group of researchers, publishers, librarians, policy makers, and technologists gathered together in an attempt to learn how all stakeholders in scholarly communication are grappling with the changes that advances in technology have wrought upon them. How do they remain relevant to their users and sustainable as an organization well into the future? This dialogue went on for two-and-a-half days while attendees discussed the new initiatives in open access, the growth of pre-print servers, user expectations, the democratization of information, and the innovative technologies that may allow them to retain relevance in the eyes of information seeker - a discussion sprinkled with philosophical debates over who does versus who should own data. It was no surprise that at the close of this conversation no one could predict the future configuration of the information landscape. What was a surprise was the call for less passionate and non-constructive rhetoric from both sides of the Open Access issue, and more rationale and collaborative conversations about the future of the information community and its stakeholders. Perhaps we are making our way to the table to work things out?

2. Opening keynote

The opening keynote presentation was given by Dr. Samuel Zidovetzki, Global Health Director, University of California Riverside Emergency Medicine Program, who spoke on the need for global access to up-to-date medical information. He noted that information is a commodity, one that is tightly controlled and regulated by many different entities, both public and private. And medical information is no different. For developing countries, gaining access to valuable and time-sensitive medical information can mean the difference between life and death. Equally important is access to quality medical education versus what exists now in many countries - educational curricula using decades old medical texts. There are many barriers to the dissemination of accurate and timely medical information and there are organizations such as Wikipedia, Wiki Project Medicine, and other non-governmental and governmental organizations working to get medical information to those who need it most.

To prove his point, he gave several examples based upon his work in refugee camps and noted that the medical conditions are quite rudimentary. In the Dominican Republic newly-minted doctors go right from medical school to work in the field and, unlike in the USA, they do not have the benefit of supervision by experienced physicians. They identify a patient's symptoms and then rely upon outdated textbooks, old equipment, and even old medical posters hanging on the walls in order to make a diagnosis. While they may possess a smartphone, Internet access is unreliable and costly – the best that they can do is download pdf's to the phone when they have the opportunity and refer to them as needed. One very interesting and

novel solution is the “Internet-in-a-Box” [8] that originated in Cuba. Users come to a park and download what they need from a hot spot (see video at <http://internet-in-a-box.org/>). Content can be customized based upon the setting and the box is available commercially. The Box can also be used as a teaching tool and a side use has been the creation of a mobile classroom. He said that a study will be done in Guatemala in May of this year to see what information is most used and most helpful.

Zidovetzki said that he uses the Internet-in-a-Box in his own work and noted that physicians actually do rely upon online information such as Wikipedia (even in the United States) [9], and offered the following statistics on Wikipedia usage in the medical field: 50%–100% of physicians, 35%–70% of pharmacists, and 94% of medical students - even policy makers access the information and it was the number one resource during the 2014 Ebola crisis [10]. The use of online materials by physicians was confirmed by Violaine Iglesias, a speaker later in the conference.

Zidovetzki and others are working diligently to gather relevant, quality medical information that can be stored on Internet-in-a-Box and utilized offline in rural areas. It is important to have the information in native languages and it is a struggle to get the information translated so that the information is current. But he noted that progress is being made and the ability to have offline access to quality medical information will make a big difference in underdeveloped areas. He believes very strongly that ensuring the flow of high quality, timely information to medical professionals either at home or abroad can make all the difference in improving patient care.

Dr. Zidovetzki’s slides are not available on the NFAIS website, but more information can be found in an interview with him on his work with the WikiMedicine Project [11].

3. Advancing knowledge and research in the 21st century

Mary Lee Kennedy, Executive Director of the Association of Research Libraries (ARL), was the first speaker in this session and she discussed the work of research libraries in advancing knowledge in today’s research and learning ecosystem - an ecosystem that is focused on open science, the adoption of artificial intelligence and other fourth industrial revolution technologies [12], and the debate about what constitutes research in the 21st century.

She noted that ARL’s mission is to advance research, learning, and scholarly communications by fostering the open exchange of ideas and expertise, promoting equity and diversity, pursuing advocacy and public policy efforts, forging partnerships, and by catalyzing collective efforts.

She said that they have been giving a lot of thought as to where ARL can add the greatest value and they believe it is at the intersection of three communities: (1) Research libraries and their parent organizations; (2) Public Policy makers; and (3) Research and learning Communities. She noted that each community has its own topics under discussion. For Research Institutions the topics are: accountability/value/public good, trust, affordability, diversity, equity and inclusion, and demographics. For Public Policy makers the topics are: open science, higher education accreditation, budgets, copyright, accessibility, net neutrality, and privacy and security. And for the research and learning communities the topics are open access, data, video, complex objects, text etc., cyber-physical systems, collaboration and continuous authorship, and personalized learning. These “topics of conversation” are all related; e.g. trust vs. privacy and security vs. accessibility, open access and open science, diversity and net neutrality, etc. And to be successful across the communities, the conversations need to be pulled together, and ARL hopes to facilitate just that. Their priority efforts are focused on the areas of scholars and research; diversity, equity and inclusion; and on data and analytics.

With regard to scholars and research they convened a meeting in December 2018 attended by learned societies, publishers, librarians, funders, and universities to see what initiatives could be established to advance open scholarship in the social sciences. By the end of the meeting, five group projects were proposed, with commitments from various participants to lead them. The projects will:

- (1) Conduct an authoritative investigation into scholarly society finances by a trusted third party, as the basis for financial and business model conversations with societies and external stakeholders.
- (2) Commission a paper on the role of scholarly societies and scholarly affiliation in a post-subscription environment.
- (3) Conduct a case study pilot on linguistics promotion-and-tenure (P&T).
- (4) Explore implementing peer review in SocArXiv and PsyArXiv.
- (5) Assess the impact of the reporting relationship between university presses and university libraries.

The full report can be accessed at <https://www.arl.org/wp-content/uploads/2019/01/2019.01.25-arl-ssrc-meeting-on-open-scholarship.pdf>.

With regard to diversity, equity and inclusion she noted their work with the University of British Columbia in collecting and curating scholarly material created by Indigenous communities. Annually, the librarians produce a list of courses with significant Indigenous content. For the winter session 2018–2019, there were one hundred and eighteen courses, from thirty-five different departments. The Xwi7xwa Library (see: <https://xwi7xwa.library.ubc.ca/>) at the university collects materials written from Indigenous perspectives, such as materials produced by, Indigenous organizations, tribal councils, schools, publishers, researchers, writers, scholars, filmmakers, and musicians. Its collections and services reflect Aboriginal approaches to teaching, learning, and research. She noted that the ARL-sponsored IDEAL19 conference - Advancing Inclusion, Diversity, Equity, and Accessibility in Libraries & Archives, will be held August 6–7, 2019 in Columbus OH [13].

On the final area of focus, data and analytics, she reported on the work of the ARL Assessment Program Task Force that met in December 2017. As a result of their work [14] five team-based pilot projects have been launched this year to look at the following questions:

- (1) How does the library help to increase research productivity and impact?
- (2) How do library spaces facilitate innovative research, creative thinking, and problem-solving?
- (3) How does the library contribute to equitable student outcomes and an inclusive learning environment?
- (4) How do the library's special collections specifically support and promote teaching, learning, and research?
- (5) How do the library's collections play a role in attracting and retaining top researchers and faculty to the institution?

The goal is to ultimately measure outcomes and results, not just things. Library assessment is really a new field and is becoming very important. She noted that outcomes will be examined within the goal of supporting and advancing research libraries and scholarly communication with an eye on creating the workforce of 2030 – only eleven years away!

Kennedy's slides are available on the NFAIS website and a paper based on her presentation appears elsewhere in the issue of *Information Services and Use*.

4. Virginia Tech and open science

The second speaker in this session was Julie Griffin, Associate Dean for Research and Informatics at Virginia Polytechnic Institute and State University. She opened with an overview of the institution, noting that there are over thirty-four thousand students and two hundred and eighty degree programs. It has a research portfolio of five hundred and twenty-two million dollars, a global education office, and has had a European campus in Switzerland for more than twenty years.

She went on to discuss the evolution of Virginia Tech (VT) University Libraries, with special attention given to program and service developments in data science, learning spaces, digital literacy, and open education. Griffin noted that VT Libraries facilitate knowledge sharing, use, and reuse through the provision of open publishing services and openly-accessible collections, with the goal of providing equitable access to information. She added that VT Libraries help make the scholarly record more open by embedding library services in the university research and learning infrastructures, by teaching digital literacy skills, by advocating for open access and open data, by integrating systems, and by providing services that enable institutional stewardship.

She noted that in 2018 VT entered an online publishing partnership with Ubiquity Press (see <https://www.ubiquitypress.com/>). However, her institution is not a newcomer to online publishing. VT began publishing online journals in 1989, and was one of the first universities in the country to require students to submit theses and dissertations electronically. This new three-year agreement with Ubiquity Press will allow the libraries to gain a state-of-the-art web platform that increases its capacity to publish freely-accessible scholarly research in a variety of formats, such as journals, books, conference proceedings, along with openly-licensed text, media, and other digital work used for teaching, learning, and research. According to the press release “The move to Ubiquity is part of a larger strategy by the libraries to build a publishing program - called VT Publishing - designed for the 21st-century digital economy [15].

In closing she also discussed VT’s support of the National Academy of Sciences (NAS) vision for Open Science by design – defined as “a set of principles and practices that fosters openness throughout the entire research lifecycle.” [16] Indeed, VT utilizes the Open Science Framework (see <https://osf.io/>) to ensure that the university’s faculty and researchers can openly share their work. (Note: in the Q&A period following the presentations in this session, Mary Lee Kennedy highly recommended that attendees read the free NAS report on Open Science by Design. She said that it is an easy read and that ARL is following its guiding principles).

Griffin’s slides are available on the NFAIS website and a paper based on her presentation appears elsewhere in the issue of *Information Services and Use*.

5. Publishers creating new value

The next session focused on how three publishers are creating new value for their content:

5.1. Institution of Engineering and Technology (IET)

The first of the three speakers was Vincent Cassidy, Director of Academic Markets at the Institution of Engineering and Technology, IET, who discussed how they have added value to the Inspec database via artificial intelligence and semantic tagging in order to create new services. He began by providing

an overview of IET. It is a global professional society as well as a Learned Society with more than one hundred and seventy thousand members. Headquartered in London, UK, it is a mid-size full-range publisher producing books, journals, magazines, standards, etc., as well as an abstracting and indexing database, Inspec. The organization itself is one hundred and forty-eight years old and is very traditional in its thinking, so one of the major challenges moving forward was to step back and try to look at things in new and different ways.

Inspec itself will be fifty years old this year. As of January 2019, it had eighteen and a half million records in the fields of physics; electrical and electronic engineering; computing and control engineering; and production, manufacturing and mechanical engineering. It includes material from more than forty-five hundred journals, and three thousand other publications from seven hundred and fifty publishers. Materials covered include books, journals, videos dissertations, etc. The database makes scientific research discovery easier through high-quality classification and indexing, through the curation of source material, and through monitoring science and research for quality and relevance. Inspec's main customer base is comprised of universities, research institutes and corporations. Cassidy noted that other A&I services in the audience can appreciate the effort that goes into creating and maintaining such a massive, curated database. It is very labor-intensive, especially in indexing.

So why change after fifty years of success? Well, as Cassidy pointed out, the research information environment has changed significantly. The advent of "good enough" alternatives has fundamentally changed user behavior - they prefer Google - not A&I services, and as a result usage is declining. And while Information professionals respect and value Inspec, there has been a decline in subscriptions in Inspec's core base. IET wanted to take action now, before things slid further. And most of what they are now doing required new skill sets and the hiring of data scientists.

Their goal was to apply semantic tagging to the eighteen million records, to develop a domain model from their diverse ontologies, to better understand user workflow and pain points, to build an MVP [17] (minimal viable product) platform from which to iterate, and to build and strengthen engagement with their customers. He said that incorporating agile product management skills into a one hundred and forty-eight year old institution has been "interesting" to say the least! He noted that the culture of the organization had to change - not the vision, and that they had the full support of IET's Trustees.

It has been a two-year process with frequent iterations for funding as well as to the market to confirm assumptions. They have taken the eighteen million records and created six billion concept relationships in a Knowledge Graph and this new product will be launched in the spring of this year as part of Inspec's fiftieth birthday celebration.

They faced several challenges during the process, beginning with their choice of technology partners; introducing effective agile project management and project momentum as noted earlier; understanding customer needs and the user workflow; and sustaining the core business through the two-year effort. But Cassidy has said that the effort has proven to be successful.

They have reengaged with customers and users and the new value-add has triggered increased usage and growth. It is the first new organic growth that they have had in years. Their conversations with their distribution channel partners have resulted in the emergence of new ideas and a fresh impetus on growth. And they have been able to refocus content curation and the inclusion of new content types. He said that they learned a lot as a result.

His closing message was that highly-structured, human-curated databases *can* be repurposed and recycled to retain relevance and to provide new value propositions with a resultant growth in usage and in impact. He said that as an organization IET is no longer "traditional," but rather has evolved into a one hundred and forty-eight year old "start-up" which is strange, exciting, and invigorating. He

emphasized that this “new life” is not about technology - it is about data and innovative uses of data. And he recommended that organizations who want to reinvent themselves to remain relevant to the new generation of information users get the skills required to make the most of the data that they have invested in compiling and curation over the years. Skill up!

Cassidy’s slides are available on the NFAIS website.

5.2. Association of University Presses (AAUP)

The second speaker in this session was Peter Berkery, Executive Director of the Association of University Presses (AAUP), who discussed how university presses are finding new ways to support the output of scholars in the Humanities and Social Sciences who have become increasingly comfortable with both digital research and publishing. In his presentation he highlighted four initiatives as follows.

5.2.1. Rotunda

Rotunda was founded by the University of Virginia Press in 2001 to apply traditional University Press (UP) strengths to the research that was beginning to emerge from Digital Humanities Centers (see <https://www.upress.virginia.edu/rotunda>), that strength being the UP imprimatur (e.g. peer review). The Press quickly discovered that these one-off boutique projects did not have the scale that libraries were looking for and it did a slight pivot towards documentaries which evolved over time into the Founding Fathers documentary editions, and again into literature and culture collections, such as Emily Dickinson’s correspondence, Herman Melville’s draft manuscript of “Typee,” etc.

Today tension persists among scholars between Digital Humanities as digital affordances versus a true Research and development effort; e.g., Natural Language Search, word frequency analysis, data manipulation, etc. The people at University of Virginia Press wrestle with this issue as well and they have found that sustainability issues still persist for boutique projects. They have found that their library customers want scalability: a perpetual license for what looks like content, and a small annual hosting fee for the platform (hosting, enhancements, browser compatibility). Going forward, new projects will incrementally expand Rotunda’s capabilities without compromising the Press’ core values and they will continue to do what they have done so well for almost two decades.

5.2.2. Manifold

Developed by the University of Minnesota Press, Manifold is a do-it yourself (DIY) web-based publishing platform for scholarly publishers, university departments, and scholarly groups (see <https://manifold.umn.edu/>). A Manifold project is composed of three parts: first, the base layer - the epub file or Google doc that a publisher uploads; second, the resources - the media and other texts that an author and editor will add to a project; and finally reader interaction comprised of comments, highlights, etc. There are also fee-based services for those organizations who do not want to implement Manifold on their own. Early experience has shown that a lot of groups do not want to DIY, even when it’s free!

5.2.3. Fulcrum

Fulcrum was developed by a group of campus-based publishers working closely with disciplinary faculty and information science specialists who recognized the changing nature of scholarly publishing in the humanities and social sciences (see <https://www.fulcrum.org/>). Initial development was supported by

a grant from the Andrew W. Mellon Foundation and implemented by the University of Michigan Library and Press working with partners from Indiana, Minnesota, Northwestern, and Penn State universities. It is a digital publishing platform and a set of publishing services that is committed to publishing scholarship in a flexible, durable, discoverable, and accessible form. It is flexible in that it connects to other open source tools and is responsive to the changing needs of digital scholars. It is durable because it has been built on a research library infrastructure that is a trusted steward committed to preservation and stability, Discoverability is supported by that fact that the system is interoperable with other publishing tools and has been integrated into the information supply chain. And it is accessible in that it is dedicated to inclusive services and content for all readers.

5.2.4. *.supDigital*

Developed by Stanford University Press, *.supDigital* is attempting something transformative at the research level by applying the rigors of traditional university press publishing to born-digital scholarship. The goal is to provide a formal, peer-reviewed publication process for interactive digital scholarship, allowing scholars to create a digital object that presents, explains, and displays their research. Once an overarching argument is embedded, the press can peer review and thus validate the work. Thus far, *.supDigital* has developed standards for presentation, metadata, a hosting platform, and archiving capabilities (see video at <http://blog.supdigital.org/>).

Berkery noted that the lessons to be learned from these initiatives are the following:

Going into a project know if the results are meant to be leveragable (standard format) or flexible and he used the Digital Einstein Papers website developed by Princeton University Press as an example (see <https://press.princeton.edu/einstein/digital>). He noted that the website is fantastic, but the infrastructure is not applicable to anything else. Which segues into the issue of sustainability as these projects are expensive and the investment needs to be leveraged in multiple ways. Also, peer review is a requirement and must be incorporated as University Presses develop new means to “publish” outcomes from digital humanities research. He added that surprisingly, do-it-yourself implementations have had limited appeal, And finally, he noted that definitional issues surrounding the term “Digital Humanities” remain.

In closing, he offered the following quote from Alan Harvey, Director, Stanford University Press: “The goal is not to publish a book in digital form. The goal is to publish digital scholarship in its native form. That means embedding the scholarly argument within the digital object.”

Berkery’s slides are available on the NFAIS website and a paper based on his presentation appears elsewhere in the issue of *Information Services and Use*.

5.2.5. *MIT knowledge futures group*

The final speaker in this session was Catherine Ahearn, Senior Project Editor, MIT Knowledge Futures Group (KFG) - new joint initiative of the MIT Press and the MIT Media Lab. It is the first of its kind between an established publisher and a world-class academic lab devoted to the design of future-facing technologies.

Ahearn noted that the KFG’s mission is to transform research publishing from a closed, sequential process, into an open, community-driven one, by incubating and deploying open source technologies to support both rapid, open dissemination and a shared ecosystem for information review, provenance, and verification. It provides support for mission-driven publishers and brings like-minded groups and individuals together. Currently, there are four projects:

- PubPub: This is a free, open authoring and publishing platform initially developed as a Media Lab project. It socializes the process of knowledge creation by integrating conversation, annotation, and versioning into short and long-form digital publications. Currently there are more than one hundred PubPub communities (see <https://www.pubpub.org/explore>).
- Underlay: Protocol for data interoperability. It is an open, distributed knowledge store that is architected to capture, connect, and archive publicly-available knowledge and its provenance. The Underlay provides mechanisms for distilling the knowledge graph from openly-available publications, along with the archival and access technology to make the data and content hosted on PubPub available to other platforms.
- Prior Art Archive: The first free and open archiving platform for prior technical art for the entire IT industry, prototyped by MIT and Cisco. Its goal is to help fewer bad patents be issued, by giving USPTO examiners the tools they need to find old technology.
- Ecosystem map: A Mellon-funded environment scan to be published in June 2019.

Ahearn noted that there is value in experimentation. They are redefining the digital reading experience; supporting the development and evolution of new ideas; and introducing transparency and openness in the review process to provide greater value to authors and readers

She offered some interesting examples of what they have accomplished to date. The first was a collaborative reading experiment, Frankenbook - a special Edition for the 200th anniversary of Mary Shelley's *Frankenstein* that was published on PubPub in January 2018. This edition included additional annotations from the editors, multimedia embedded in the text and annotations; labels added to the annotations for tailored reading; special functionality for use in classrooms; and community discussions around the text - currently Frankenbook has four hundred and twenty discussions and more than seven thousand visits (see: <https://www.frankenbook.org/>).

A second example is Works in Progress (WIP). These are written works in early stages of development that would benefit from an open peer review process. It offers authors the benefit of community feedback in the development of their ideas, as well as the ability to publish a version of their work before more formal publication. After the open review period, authors may revise the work and submit it for consideration for formal publication. The MIT Press will have first right of refusal, and all submitted manuscripts will be subject to the Press usual rigorous peer review (see: <https://wip.pubpub.org/>).

Her final example was *Data Feminism* (see <https://bookbook.pubpub.org/data-feminism>), a contracted manuscript by Catherine D'Ignazio and Lauren Klein. It is available for peer-to-peer review and will be published by the MIT Press as part of their Ideas series. She noted that all titles in the Ideas series are open access and published on PubPub with support from the MIT Libraries. This manuscript had more than five hundred comments and more than three thousand visits at the close of open reviews. This was a successful manuscript that emerged from the WIP program noted above.

Ahearn said that while there have been many benefits from the experimentation that they have been doing, they have also faced some challenges, not the least of which have been funding models; creating new workflows and/or integrating with existing ones; and creating and communicating realistic goals.

In closing, she invited all attendees to consider creating a community (go to: <https://www.pubpub.org/community/create>) or joining an existing one and noted that all of their code is openly-available on GitHub.

Ahearn's slides are available on the NFAIS website.

6. The challenges to information discovery

The final speaker of the day was Tim McGeary, Associate University Librarian for Digital Strategies and Technology, Duke University Libraries, who gave one of the more thoughtful presentations at the conference. He discussed a major issue that all librarians have to address – how to provide state-of-the-art information discovery and personalization services while protecting user privacy.

He opened with the following question:

“How can a system be user-centric if it is not all about the user? And how can that happen without sacrificing some of our deepest values about privacy? Are our values incompatible with user-centricity?”

He noted that libraries have a unique position within the technology-driven marketplace - they are both an information-consumer and an information-provider. And, while libraries need to protect user privacy, they are competing for users with commercial services who are not so focused on user privacy and who gather a lot of user-information to provide very customized services - services that users also expect from libraries who do not gather such personal data.

McGeary looked at the past thirty years of information discovery - from the launch of web-based Online Public Access Catalogs [18] (OPACs) in the 1990's, Google in 1998, faceted [19] OPACs in the early 2000s, followed by web-scale discovery services [20] offered by companies such as EBSCO and ProQuest, and now index-based discovery services [21]. He noted that modern discovery services and advanced OPACs all use a single search box because that is what Google has always used and Google has (as we all know) totally reshaped user behavior and expectations.

So what do users expect? They want to find all of the information that they need in one place; they want to be able to access information online from any location and using any device; they want personalization and customized search; and they want just-in-time customer support. He noted that one of the biggest issues is being able to access information easily and noted that many academics who have legal access to content in their library actually use pirated material from Sci-Hub simply because it is easier to use than many library systems and he reviewed several of the new services that are attempting to alleviate this problem (see article by John Seguin that appears elsewhere in this issue as well as several discussions from the 2018 NFAIS Annual Conference [22]). He noted that Universities and libraries now realize what commercial and consumer providers have known for a long-time - that the data about their users is deep, untapped, and full of potential. He added that while libraries have long aimed to protect this data from being used harmfully, the era of constrained financial support, especially for libraries, requires much more intentional data-driven decision-making.

In closing, McGeary noted that the technology impact to user-centric discovery has created a new paradigm for libraries who must be willing to adapt to the user expectations of personalization, options, and convenience. Libraries should aim to serve their users better with the responsible use of data and be willing to go further than they have dared before in collaborating and obtaining informed consent from their users. But what remains constant is that libraries must continue to protect their users' intellectual freedom and privacy as a most foundational value, and the core to user-centric discovery.

McGeary's slides are available on the NFAIS website and a paper based on his presentation appears elsewhere in the issue of *Information Services and Use*.

7. An automated solution to information access

Day two of the conference opened with a plenary session given by Sabine Louët, CEO and Founder of SciencePOD, a start-up company in Dublin, Ireland that offers a digital content creation and publishing platform, designed to give people who are not familiar with the requirements of publishing, the help that they need to create high-quality content telling the story of their research as well as to provide experienced publishers with cost-effective tools to produce content (see: <https://sciencepod.net/#splash>). Louët was formerly a news editor at *Nature Biotechnology* and the Editor of *EuroScientist*. She said that her company has a mission and that is to explain the meaning of scientific research findings that are buried in the journal literature and in databases, and that accomplish this by translating complex ideas into simple messages. She opened by saying that there are a lot of misconceptions about where innovation is coming from. It is not, in her opinion, coming from large, established companies who are perhaps a bit risk-averse, but rather it is coming from start-up companies. She said that she was at a meeting in Berlin last week where someone did a study of one hundred and twenty start-ups and found that 77% were truly stand-alone companies, not offshoots of a larger corporation and twenty-three of the sample companies were ultimately acquired because of their innovations.

She said that as an entrepreneur she has learned to ask the right questions and one of the big questions today is what is driving innovation in the information industry? She believes that one of the major forces is Open Access (OA) and she referred to a presentation made at the 2017 NFAIS Annual Conference [23] by Deni Auclair of DeltaThink in which it was reported that OA accounted for twenty percent of the global content output, but only for ten percent of the revenue. OA has a ways to go, but it is having a major impact today and she commented on Plan S – the effort by the German government to accelerate OA by making it a requirement that all government-funded research in Germany be published in OA journals beginning in February 2020 [24]. She gave as an example the major deal that Wiley has entered into in Germany allowing seven hundred German institution access to Wiley journals and also allowing their researchers to publish in Wiley's OA journals. The multi-million dollar deal was signed on January 15, 2019 and a public version of the contract is available [25]. Louët said that she believes that more of these deals will be announced in the not-too-distant future. Many speakers throughout the conference reference Plan S and the Wiley deal.

She noted that the Impact Factor [26] still dictates journal and author recognition, and most journals with high Impact Factors are not OA (one of the problems facing the implementation of Plan S). She added that with social media and science networks we are entering an era where the author is becoming a “brand” more so than the journal. Because of this, Publishers are becoming more author-centric and are seeking to provide more author services, some of which are to assist authors in promoting their work. Which is where her company comes in.

Her company is focusing on taking bundles of articles (including OA articles) and, using Artificial Intelligence (AI), creating a “story” of the research in those bundles with the goal of shining a light on the authors of those papers. Put simply, they create “automated summaries.” She noted that the use of automation to “write” articles started in 2014 when the first new report on an earthquake in Los Angeles, CA was written by a robot [27]. As noted at the time, algorithms won't necessarily replace editors and reporters because they cannot generate proficient text, but automation can help speed-up the writing process.

In closing, she noted that the summaries created by a combination of SciencePOD editors, writers, and AI improve content discovery. She said that this combination makes complex topics accessible, using

crystal-clear English, so that information is easily understood and that it can save publishers editorial time and costs when creating content to raise the profile of their content and their authors.

Louët's slides are available on the NFAIS website.

8. Unconventional partnerships

8.1. Preprints and journals

The next session highlighted three speakers who focused on the importance of partnerships in the information industry. The first speaker was John Inglis, Co-Founder of bioRxiv and medRxiv, and Executive Director of Cold Spring Harbor Laboratory Press. The goal of his presentation was to demonstrate (1) the advantages that preprints offer individual scholars and their communities of practice, (2) the integration and collaboration that is possible between servers and journals, and (3) the ways in which upstream discovery and assessment of a preprint may help optimize the published version of record of the manuscript concerned.

He opened with the definition of a preprint - a research manuscript yet to be certified by peer review and accepted for publication by a journal that is loaded on an online platform that is dedicated to the distribution of preprints. He then went on to demonstrate the power of preprints using (with permission) the example of Kenji Sugioka Ph.D. who is now an Assistant Professor at the University of British Columbia. In January 2017 Sugioka published his first paper as a Post Doc at the University of Oregon. In June of that same year he published a preprint of a minor project that he was working on, followed in July with another preprint that discussed some more important research with which he was involved. The reaction to both preprints was very positive so Sugioka decided in October to start looking for a tenure-track position and due to his preprints he was invited to a lot of interviews. In January 2018 his minor project paper was published. By April he had accepted an Assistant Professor position at the University of British Columbia, and in August 2018 the paper on his major work was published. This, Inglis he said, is the power of preprints. Communication was done in parallel to the speed of research. The results were globally-disseminated freely; and the results were shared, evaluated, and commented upon quickly. Sugioka was able to clearly demonstrate his experience, knowledge, and value - and landed a tenure-track position in less than a year from when he started the search! Inglis noted that preprints:

- Uncouple distribution of results from certification through peer review.
- Enable community awareness of recent work and the chance to comment.
- Provide evidence of productivity before papers are published, and
- Accelerate the pace of research and the advancement of science.

He offered the following quote in support of the above: "If one preprint inspired the work of just two other people, biologists would see a five-fold acceleration in scientific progress in a decade" (Steve Quake, President, CZI BioHub, *Wired* July 8, 2017).

He noted that the number of preprint servers is proliferating across all disciplines and sub-disciplines, each with its own technology and policies (for an excellent overview of preprint growth see the summary of a presentation given by Shirley Decker-Lucke of Elsevier on Preprint Servers at the 2018 NFAIS Annual Conference) [28]. He then went on to describe the two preprint servers hosted by his organization, bioRxiv and soon-to-be-launched medRxiv. The former is a server for life science preprints. It is a five-year old not-for-profit old service, not a product, of Cold Spring Harbor Laboratory, and it is free for both authors

and readers. It is hosted by HighWire Press, supported by the Chan/Zuckerberg Initiative, and is publisher-neutral.

Submitted manuscripts are screened, not peer reviewed, and then posted within twenty-four to forty-eight hours. Submitters must declare (1) that they have the right and permission to submit, and (2) that the manuscript is unpublished. They must also register any clinical trials with a registry approved by the International Committee of Medical Journal Editors (ICMJE). Options soon to come are funder acknowledgement and access to data via a link to a repository. In-house staff check that the submission requirements are met; that the manuscript is of an appropriate scope (not non-science or pseudoscience), that it does not contain any obscenity, defamation, or plagiarism, and that it is, indeed, a research paper. Also, images of human subjects are not accepted for inclusion. Manuscripts are either made live or flagged for further review after consideration of the following questions: Is the content science? Is the manuscript a complete research paper? Is the content at a level appropriate for sharing with practicing scientists, regardless of quality or accuracy? Does the content have potential to harm (or prompt behavior that might harm) individual patients or populations; e.g., articles about dual-use research; articles about vaccine safety or infectious disease transmission; articles promoting or disputing specific drug regimens; and articles about the toxicity/carcinogenicity of common substances?

As of early this year there were forty-two thousand posted manuscripts; one hundred and seventy-seven thousand authors representing fourteen thousand seven hundred institutions in one hundred and ten countries. As of January there had been 3.3M page views and 1.3M PDF downloads. Usage, downloads, and manuscript submissions are all growing. He noted that use of Twitter to talk about preprints is popular and helps accelerate growth in readership.

He noted that bioRxiv partners with publishers so that authors can ultimately submit their final manuscript to be published - approximately one hundred and forty-two journals are covered by these partnerships - and some journals actively encourage submission of the preprints themselves. He noted that sixty seven percent of bioRxiv preprints are published on journals within two years, with a median eight months between posting and publication. Preprints are given a forward link to the published version and the published version has the journal DOI and a backward link to preprint.

Inglis briefly discussed three new services focused on preprints that have recently emerged. The first is preLights, a community platform for selecting, highlighting, and commenting on recent preprints from across the biological sciences launched by the Company of Biologists [29]. The second is PRE-reviews, a journal club for preprints that encourages scientists to post their outputs as preprints (see: <https://www.authorea.com/inst/14743-prereview>). And the third is Peer Community In (PCI), a non-profit scientific organization that aims to create specific communities of researchers that review and recommend, for free, unpublished preprints in their field (see: <https://peercommunityin.org/>).

In closing, Inglis noted that Cold Spring Harbor Laboratory Press, jointly with Yale University and the British Medical Journal, will be launching a preprint service for medicine (medRxiv) on June 25, 2019 [30].

Inglis' slides are available on the NFAIS website and they provide a lot of detail.

8.2. Blockchain, ARTiFACTS, and the Max Planck Society

The second speaker in this session was David Kochalko, Co-founder of ARTiFACTS, a relatively-new company that uses Blockchain technology to support the flow of research from start to finish. ARTiFACTS first came to my attention when Courtney Morris, the other Co-Founder of the company, spoke at NFAIS' conference on the use of Blockchain technology in scholarly communication last May [31]. The company

was launched in March 2018 and since then has entered into multiple partnerships, the most recent being with the Max Planck Society and the Bloxberg consortium [32]. The consortium is attempting to provide an infrastructure that will allow researchers around the world to use blockchain technology for their collaborative research. Their vision is to have sufficient representation from various scientific entities actively participating in the consortium, so that the blockchain network itself may replace the traditional scientific infrastructure, ultimately eliminating current challenges such as the closed-access publishing of research results, among others.

So what is Blockchain technology? According to a recent report from the National Institute of Standards and Technology (NIST), “Blockchains are immutable digital ledger systems implemented in a distributed fashion (i.e., without a central repository) and usually without a central authority. At its most basic level, they enable a community of users to record transactions in a ledger public to that community such that no transaction can be changed once published.” That same publication concluded that “The use of blockchains is still in its early stages, but it is built on widely-understood and sound cryptographic principles. Moving forward, it is likely that blockchains will be another tool that can be used to solve newer sets of problems... Blockchain technologies have the power to disrupt many industries. To avoid missed opportunities and undesirable surprises, organizations should start investigating whether or not a blockchain can help them. [33] The ARTiFACTS blockchain allows both individuals and organizations to get on the band wagon.

They have their own distributed ledger system (blockchain) that individual researchers can use, free of charge, to upload their research findings as they go through the research process all the way through to final publication. By doing so researchers have ultimate proof of their work and when it was done (entries are time-stamped). They can protect and manage their intellectual property while facilitating knowledge sharing, if and when they want to share; and they can get credit at any point for any type of research output - they do not have to wait until their research results are actually published. According to Kochalko, there will ultimately be a “deep historical archive of published and discovered findings” that will be accessible to the broader scientific community [34].

Kochalko’s slides are available on the NFAIS website and a brief paper based upon his talk appears elsewhere in this issue of *Information Services and Use (ISU)*. For more information on the use of Blockchain technology in scholarly communication, I refer you to the special issue of issue of ISU on this topic that was published last year (see: <https://content.iospress.com/journals/information-services-and-use/38/3>).

8.3. *The EBSCO partnership strategy*

The final speaker in this session was Nathanael Lee, Strategy Analyst at EBSCO Information Services, who discussed EBSCO’s philosophy of partnerships and how their partnership strategy has evolved over the years. Lee defined “partnership” as a relationship between two or more entities for the exchange of goods, services, and/or ideas in order to create outputs.” He noted that EBSCO started as a subscription business helping publishers get their journals into the hands of librarians around the world and as a result they built a really strong publisher network. As the information industry and technology changed EBSCO realized that librarians increasingly wanted digital information, they wanted to purchase “bundles” of journals, and they wanted more affordable and easy access to information overall. In the 1980’s EBSCO turned their attention to technology and acquired a CD ROM business, then, in partnership with members of their publisher network, they built EBSCO Host as a platform for the digital distribution of journals. The system, Lee added, is used by every major library in the world today. They then started buying major

abstracting and indexing services such as H. W. Wilson with the goal of building their own “indexes” to the journal literature and ultimately built a “discovery service” that libraries can use to access and search their holdings (see mention of such services in Tim McGeary’s presentation mentioned earlier).

Now EBSCO is entering the third phase of their business – they are offering services and solutions. They have formed a partnership with OpenAthens (see: <https://openathens.org/>) to provide easy user authentication/logins for EBSCO services. They partnered with Stacks to get a web-based content management system and have since acquired the organization. Now EBSCO offers the service to libraries together with login alternatives via OpenAthens (see: <https://www.ebsco.com/products/ebsco-stacks-library-websites>). A more recent partnership that Lee spearheaded is with StackMap (see: <http://www.stackmap.io/>) who offers a GPS-like tool that allows library patrons to locate an object; e.g. book, in the physical library. EBSCO has integrated that “service” into the search results of the EBSCO discovery service so users get both the information that they need and directions to its physical location in the library - cool! Another example is a partnership with FOLIO, which stands for the “Future of Libraries is Open.” It is a community founded in 2016 to develop a “reimagined” library services platform. While it is open source, EBSCO works with FOLIO to provide library support when needed (see: <https://www.ebsco.com/partnerships/folio>).

Lee said that EBSCO is a mission-driven organization and that their mission is to “Transform lives by providing reliable, relevant information when, how, and where people need it.” They look for partners with values that support the EBSCO mission such as thinking long-term and making society better off. They believe in supporting their partners in their shared goals.

He closed by encouraging organizations represented in the audience to consider partnering with EBSCO, with some of the reasons being that they have the largest sales force in the library industry, they offer proof-of-concept testing for new ideas, and are willing to share the results of what they learn. For an interesting history of EBSCO Industries (of which EBSCO Information Services is a part) see: <http://www.fundinguniverse.com/company-histories/ebsco-industries-inc-history/>.

Lee’s slides are available on the NFAIS website.

9. Library consortia - a changing world

The final speaker of the morning was Roger Schonfeld, Director of Libraries, Scholarly Communication, and Museums Program at Ithaka S+R who discussed some of the changes that the library community has been experiencing. He opened with a “pre-digital history” discussing the library dream of a universal collection delivered as efficiently and seamlessly as a system could enable – a dream that could not be fulfilled in a print world. Resource sharing was in the form of inter-library loan and union catalogs. Eventually, new library systems actually brought these functions into the digital environment. He used the example of what is now the Center for Research Libraries (CRL) when in March 1949, ten major U.S. universities entered into a formal agreement to establish the Midwest Inter-Library Corporation (MILC). Initially the Center’s chief activity was accepting and processing deposits of monographs, journals, and other materials that were transferred to it by member universities. In addition to accepting deposits, the Center also began to subscribe to U.S. and foreign newspaper titles not being acquired by members, thus establishing a collection that remains one of its enduring strengths. In the 1960s it expanded from a mid-west regional organization to one with a national scope and in 1964 became the Center for Research Libraries. As the organization grew its mission, out of necessity, changed because scope and mission, according to Schonfeld, are intertwined.

With the 60's and 70's came the digital world with shared cataloging and OCLC. Originally founded as the Ohio College Library Center in 1967 it quickly became the Online Cataloging Library Center and the first online cataloging by any library in the world took place in 1971. He discussed the regional membership-based, collaborative library networks that emerged; e.g. Solinet, Palinet, and others, and how they have either merged or disappeared over time as commercial entities emerged to provide sophisticated integrated library systems, and centralized repositories emerged to serve as centers for digital preservation. These networks or consortia wielded considerable power as “buying clubs,” and their members still do – even outside of the USA (Note: he referred to the Projekt DEAL in Germany in which libraries and research organizations get access to a set of Wiley journals back to 1997 and whose scientists can publish in Wiley Open Access journals without paying Article Processing Charges (APCs). He added that most consortia lack the scale to secure such huge deals).

Schonfeld said that today Library collaborative networks face three major challenges:

- Licenses vs. Open Access: subscriptions are slowly giving way to a growing number of OA models. Collaborative networks do not have the systems to handle this.
- Resource Sharing: Most collaborative networks were established in the print environment. Print has declined and cloud-based systems are the norm for handling digital content. With the rise of such cloud-based systems, print-based networks are no longer heavily-used and are becoming unsustainable.
- Funding: State support for higher education has declined, and continued scrutiny of library budgets has resulted in pressure to show value and differentiate against peers. Bottom line – what is the value in becoming a member of a collaborative library network in today's world?

Schonfeld said that such membership organizations are in the midst of a crisis. He noted that “trust” is the great intangible in networks. Members often want their organization to do almost everything involving collaboration, but the fact is that not every membership community infrastructure is well-suited to support a multi-purpose organization. Membership models are durable primarily due to a sort of peer pressure to belong - no one wants to be left out of the game. And it is also a fact that it is difficult to set an unambiguous strategic direction for a membership organization and it is a challenge to follow a single strategic purpose over a long period of time. He noted that many libraries belong to a number of consortia and their parent organization; i.e., the university, often is unaware of the memberships. Schonfeld said that libraries need to realign with their parent, directly contribute to the university's purpose, and integrate themselves into the university system. He noted that the essential transformations in libraries today are:

- Print vs. Digital
- Local vs. Shared
- License vs. Open Access
- General vs. Distinctive
- Collections vs. Workflows
- Selector vs. Enabler
- Provider vs. Partner

He added that all libraries need to accept these trends and restructure their operations and organizations accordingly.

In closing, Schonfeld said that the lessons learned are that:

- ✓ Every good idea does not require a new organization.
- ✓ For a new non-profit organization, grants should be used to establish a business model.

- √ Membership models are not well-suited to product organizations and marketplace competition.
- √ Open source solutions have an especially precarious balance between community governance and strategic agility.
- √ Startups have a precarious existence.

He noted that library consortia need to do some self-reflection and ask the following questions: What is our strategic role? Do we duplicate the work of others? Do we have the right partners? Is our governance model well-adopted to the strategic role that we envision?

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

10. Scholarly publishing rebuilt from scratch

The speaker for the Members-only lunch was Dr. Jon Tennant, identified as a Nomadic Paleontologist, Rogue Open Scientist, and an independent researcher and consultant who is working on public access to scientific knowledge. He asked the question: "What would Scholarly Publishing look like if it were built from Scratch?"

Is there a need to rebuild? Is there a problem? Tennant says it depends on your perspective and to whom you speak. Publishing today is either brilliant or full of holes and there is even a movie about it entitled "Paywall: the Business of Scholarship" (see: <https://paywallthemovie.com/>)!

Tennant said that in his opinion the answer to both questions above is "yes" and he purports that the problem is in many ways due to the business model - with access to information being "closed" and behind paywalls. There are efforts to change this and he mentioned some of the highlights from 2018: Sweden, Germany, France, etc. all taking a strong stand for Open Access; Plan S (see: <https://www.coalition-s.org/>); the Springer -Nature failed/delayed IPO [35]; the European Union Open Science Cloud (EOSC, see <https://eosc-portal.eu/about/eosc>), etc. He said that the current state of scholarly communication is that it is a 19th century process applied to a 17th century communication format that is slowly adapting to a 1990s-based web technology and he asserts that we can do better.

He added that our publishing failures are also due poor communication, with strong rhetoric overcoming the rational debate of issues such as Open Access and closed paywalls. We talk past each other rather than build bridges and he used an example of researchers responding to an announcement from Elsevier with regard to the four stated principles of their information system supporting research: source neutral, interoperable, transparent, and in the control of researchers. The negative Tweets against Elsevier went on for days (see: <https://twitter.com/ElsevierConnect/status/1090202327733227520>). He said that all too frequently the use of social media is reactive and not self-reflective and that the quality of social media discussions can vary in quality.

He briefly mentioned the Open Scholarship Initiative (OSI) which he believes is a brilliant idea. Their goal is to bring all of the stakeholders together to have a reasonable discussion. They assume that there is a reconcilable middle ground and that such a middle ground is worthy of attaining (see: <http://osiglobal.org/>). However, Tennant questions whether or not it is an ultimately doomed quest because tensions are rife and he suggests that perhaps there are times when reaching a middle ground is not necessarily a good thing. But through such an effort we can gain a more empathetic ground and come to better understand what each party really wants and needs.

He showed some charts about the growth of Open Access (OA) publishing and asked the audience if rapid growth of OA is a good thing. A show of hands indicated that the audience did not view it positively - a response Tennant did not expect. He also discussed briefly how publishing/peer review is being changed

to make the most of web technology and used F1000 Research as an example. This is an OA publish platform where the authors themselves control the peer review process (see an example of one of Tennant's articles at: <https://f1000research.com/articles/5-632/v30>). He calls this "constrained" innovation because most the new tools are still being built around a journal-based system and therefore depend on publishers for sustenance [36].

He noted that we all use networks and reviews to evaluate information on the Web; e.g., Trip Advisor, but we for some reason we do not do it in science. Why not use the power of professional networks to evaluate and communicate scientific results. He said that we should really give this some thought and move away from the world of journals and articles to focus on the power of networked technologies and version control. He asserted that research is a continuous process and should be communicated as such and supported David Kochalko's Blockchain efforts to get all of the research "objects" - data, videos, protocols, etc. into a place that is accessible by the entire community.

He asserts that the technology exists and said that a marriage of GitHub (see: <https://github.com/>), Stack Exchange (see: <https://stackexchange.com/>), and Wikipedia could form the foundation of the perfect platform having the required elements for quality control and moderation, for certification and reputation, and for engagement incentives. He went on to show how this "platform" could change our traditional methods of scholarly publishing. For example, publishing would no longer be organized around papers and journals, but rather there would be unrestricted content types and formats, and gatekeeping would be replaced by collaboration and constructive criticism.

He said that is still a place for publishers as no one denies the value-add that publishers bring to the process of scholarly communication. If publishers compete fairly as service providers, it is possible to move towards open scholarship with for-profits as part of the system. But some stakeholders will find it difficult to collaborate if publishers work against researchers by locking-in those profit margins.

Tennant said that the challenges to transforming scholarly communication are to ensure a shift to digital norms that will reflect the adoption of Web-based processes; to co-ordinate strategic and stepwise changes towards "open science"; to understand the changing roles of stakeholders such as editors, librarians, publishers, etc.; to reconcile changes across/between disciplines/communities having diverse norms, practices, and biases; and to resolve - with civility - the major tensions that exist between all of the stakeholders.

In closing he said that the ultimate goal is for science to be a public good for the betterment of society by pooling knowledge and resources to create a decentralized scholarly infrastructure based upon strong values, on the principles of Open Scholarship, and with communities as the focus.

Tennant's slides are on the NFAIS website. They contain a lot of information and useful links to additional worthwhile reading.

11. 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 Martin Kahn, Chairman, Code Ocean, and long-time Information Industry executive. Kahn opened by saying how surprised he was to get the award because he has not been associated directly with NFAIS through most of his career, although many of the companies with which has been involved, e.g., ProQuest, have had close relationships with NFAIS. So he was trying to figure out what he did right to be so honored and he did come up with a few ideas.

First, twenty years ago he did give a presentation at an NFAIS Annual Conference in Philadelphia and a lot of people liked it. They liked it because he took it almost word-for-word (with proper attribution, of course) from a great book by Kevin Kelly entitled *New Rules for the New Economy* [37]. Kahn then briefly reviewed his talk from twenty years ago to see what held up, saying that Kelly had ten rules that are as follows:

- (1) **Embrace the Swarm.** As power flows away from the center, the competitive advantage belongs to those who learn how to embrace decentralized points of control.
Kahn said that what Kelly meant was that the Internet is our future - we are connecting everything to everything (this was before smartphones and high-speed telecommunications). He said that there is power and insight in what he called massive dumbness.
- (2) **Increasing Returns.** As the number of connections between people and things add up, the consequences of those connections multiply out even faster, so that initial successes aren't self-limiting, but self-feeding and self-reinforcing.
Kahn noted that the power of networked connections has been enormous, leading to a multitude of opportunities.
- (3) **Plentitude, Not Scarcity.** As manufacturing techniques perfect the art of making copies plentiful, value is carried by abundance, rather than scarcity, inverting traditional business propositions.
Kahn suggested that a network economy runs on plentitude, leading to a multitude of opportunities – a world of zillions!
- (4) **Follow the Free.** As resource scarcity gives way to abundance, generosity begets wealth. Following the free rehearses the inevitable fall of prices, and takes advantage of the only true scarcity: human attention.
Kahn said that all things get cheaper as they improve and that the best way to reach ubiquity is to give things away for free.
- (5) **Feed the Web First.** As networks entangle all commerce, a firm's primary focus shifts from maximizing the firm's value to maximizing the network's value. Unless the net survives, the firm perishes. According to Kahn, members prosper as the Net prospers. It is now the rule of thumb to support the ecosystem in which we live along with other stakeholders. Closed systems, with some exceptions, have disappeared or are dying.
- (6) **Let Go at the Top.** As innovation accelerates, abandoning the highly-successful in order to escape from its eventual obsolescence becomes the most difficult and yet most essential task.
Kahn said that he believes that this is the most prescient rule - optimization precedes the demise of an organization, which is why it is much easier to start a new company than it is to change an older, established one. He referred to Vincent Cassidy's earlier talk on how IET had to change the culture of its one hundred and forty-eight year organization to survive and was awestruck at their ability to do so. Kahn noted that in times of change it is often the traits that were an organization's original strength that can be the cause of their downfall – and we are certainly in times of change! Kudos to IET for adapting and experimenting without ignoring their core competencies!
- (7) **From Places to Spaces.** As physical proximity (place) is replaced by multiple interactions with anything, anytime, anywhere (space), the opportunities for intermediaries, middlemen, and mid-size niches expand greatly.
Kahn noted that well before it happened, Kelly predicted the demise of "physical" spaces such as retail stores (Amazon versus Borders Book stores)

- (8) No Harmony, All Flux. As turbulence and instability become the norm in business, the most effective survival stance is a constant, but highly-disruption that we call innovation.
Kahn noted that the net causes turbulence and uncertainty - but that we must seek sustainable disruptive equilibrium without succumbing to it or running from it. Those who have thrived on the intellectual stimulation of change also are dismayed by where they are.
- (9) Relationship Tech. As the soft trumps the hard, the most powerful technologies are those that enhance, amplify, extend, augment, distill, recall, expand, and develop soft relationships of all types. Kahn said that this is the least realized – trust is still an issue, especially in consumer services that can misuse customer information as Tim McGeary discussed earlier.
- (10) Opportunities before Efficiencies. As fortunes are made by training machines to be ever more efficient, there is yet far greater wealth to be had by unleashing the inefficient discovery and creation of new opportunities.

Kahn suggested that we should not focus on problem solving, rather we should identify the opportunities that problems can often present. Kelly saw problem-solving as looking backward. He believed that doing the smart thing was better than doing an old thing better. Kahn said that this was the worst part of reliving that twenty-year old talk. Why? Because the example that Kelly used in the book was the industry's beloved Dialog online service – a service that Kahn acquired ten years after Kelly's book was published and an acquisition that was the worst one of Kahn's entire career. If anything, however, the revisit convinced him that future trends are identifiable by those with a strong intellect and thoughtful clarity in thinking, even if the time-frame is not.

Kahn said that the talk in Philadelphia was not the reason he received the Miles Conrad award. The real reason was his involvement with smart people at the right time and he provided several examples. First, he was involved with Bill Marovitz who saw the opportunity in the BRS online service to provide medical information to practitioners without the use of intermediaries. This relationship ultimately led to Ovid and Mark Nelson who saw that online services were to be temporarily supplanted by CD ROMs and that it was not the hardware that was essential, but rather software was the invaluable jewel. Nelson believed, correctly, that the price of hardware would eventually plummet and the two of them managed the business accordingly and ultimately survived (although at times barely). Mark had all of the ideas and strategies - Kahn said that he was simply the "front man" as Mark was painfully shy. Kahn's next venture in 2007 was at ProQuest where he was presented with a detailed proposal (codename "Magnolia") from Suzanne BeDell (now at Elsevier) and John Law for what would become the Summon Discovery Service. He jokingly said that he keeps it under his pillow as his own personal tooth fairy and good luck charm (Suzanne was in the audience!). This was the first of the library discovery services and, according to Kahn, the best thing that happened while he was at ProQuest

In summary, Kahn remarked that his journey to the Miles Conrad Award was due first and foremost to his involvement with really good people and due second to his involvement with really good technology. He was in the right place at the right time. He really had nothing to do with the creation of BRS, Ovid, or Summons - they were not his ideas nor did he write a line of code. He was simply able to put his faith in the right people.

He is now involved as Chair of Code Ocean, an organization that was one of the start-up companies in the shark-tank shootout at the 2017 NFAIS Annual Conference [38]. The company is early-stage and investor-owned. It is for-profit, but has yet to make a profit. Eventually it will bring in revenue from fee-based services, but a large part of the company will remain focused on open source software for the individual researcher who wants to create a computational environment for collaboration and

reproducibility (Code Ocean offers a collaboration platform for the creation of computational code and data; see: <https://codeocean.com/>). Today it has more than ten thousand registered users, thirteen thousand private development projects, and more than five hundred published projects linked to peer-reviewed articles - many of which are in IET journals with whom they have a productive relationship. He said that Code Ocean follows many of Kelly's principles for success and its founder, Simon Adar, is yet another remarkable person with whom Kahn has had the good luck to become associated.

In closing, Kahn said he still really does not know why he was given this honor, but he is proud to accept it and not at all ambivalent about doing so because of all that Simon Adar and the Code Ocean staff are trying to accomplish.

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

12. Unlocking the benefits of semantic search

The remainder of the afternoon focused on the value of semantic search. The first of the three speakers in this session was Bob Kasenchak, Director of Product Development at Access Innovations, Inc. and a popular speaker at past NFAIS conferences. His discussion provided an overview of semantic search (including the diverse definitions floating around) and explained some of the related technologies and applications. In many ways his presentation framed the two following talks.

Kasenchak opened by saying that semantic search goes beyond keyword searching to examine the context within which the search terms "live." Basic search fails for several reasons: (1) simple search only matches text strings; (2) by its very nature language is ambiguous; and (3), as we all know, we have amassed an enormous amount of content and continue to do so. Unfortunately simple string matching is completely inadequate for large, specialized repositories of content that have a unique technical language that evolves over time. He frequently used the phrase that "Google says things not strings," meaning that keywords are only an indicator of what the searcher is really looking for. One of the issues today is that search platforms vary and usually the default search option is the one chosen by the user - the more "advanced" options are usually ignored. And simple searching is just that - matching a string of words against the exact same words in a database, often ignoring synonyms and word variations.

He used the example of searching on Google for the term "horse" for which he received more than three million results and noted that he wondered how the results were prioritized, for the first hit listed was by an author whose last name is "Redhorse." Certainly the results were sub-optimal and there was no option to eliminate author names from the search. He modified the same search by pluralizing the search term to "horses." This time he received 1.7 million results that were completely different from those resulting from the first search. Based on the content of the second search result set, he assumed that Google Scholar ranks terms that match an article's title and author higher than those terms embedded in the text. He noted that Google Scholar does not appear to recognize that a singular and plural version of the same term may be related to the same concept! It only matches the exact term! (Apparently a search in Google itself, not Google Scholar, does recognize the relationship).

He did a second search in a service offered by a creator of scholarly services using the term "unmanned aerial vehicle" (drone) and received one hundred and seventy-one hits, but when he used the acronym "UAV" he received almost three times as many hits. Obviously in this case the search platform - at least in the default option - does not consider acronyms, so that users of the system are not getting all of the information that they are seeking. Simple searching fails over and over again because of its inherent nature of looking at exact term matching.

Semantic Search is any search that attempts to go *beyond* the text string in the box, the common denominator of which is that it tries to examine the context in which the text string lives in order to drive relevant results. The methodologies used can include the use of synonym rings, taxonomies, lexical variants, fuzzy logic, the geographic location of the searcher, the searcher's previous queries, previous similar searches, ontologies, knowledge graphs, and other strategies. Some of these are quite simple and others are quite complex. For example, some will allow you to search a database even if you are not familiar with the language, but if you are able to get part of a term correct it will bring up relevant hits. Granted, there will be some level of "noise," so the searcher needs to be careful when evaluating the results. These types of searches utilize the Levenshtein distance [39] or a similar measure to match misspellings and variants; e.g. "bob" is one variant away from "rob," and the greater the distance, the less likely the match.

Some engines "parse" queries; e.g., if Kasenchak searches on "Harrison Ford" he will get one set of hits, but if he asks the question "When is Harrison Ford's birthday?" the string is parsed and Google returns the appropriate hit. One is a keyword search, the other is not. He also discussed contextual searches that are frequently based upon geographic location or prior searches; e.g. a search on the term "pizza" brought up a number of places near him where he could buy pizza (this can be done if the search engine recognizes an IP address based upon a GPS locator). He also discussed the Google Knowledge Graph that pops up on the side of the search results' page that alerts the searcher to things that may be of interest to that specific person based upon all of the information that Google has gathered on that person as a result of their past searches; e.g. he did a search on "jaguars" thinking that it would bring back results on the animal, but rather it brought back information on a sports team that he follows on a regular basis. The Google Knowledge Graph connects search with known facts about entities and this is becoming quite common across search engines today. This "Graph" presents the user with a lot of diverse information and provides a much richer search experience. To see what I mean, go to Google and search on "Empire State Building." (Note: the underpinnings of this type of search is actually a very large ontology).

Kasenchak then moved on to the use of taxonomies, tagging, and controlled vocabularies by content providers. Such use allows the context of terms in their databases to be more easily identified. He reminded everyone that while the searcher is interested in concepts, all that a search is based upon is words - they are the "window" to the content and all that the search engine has to go on. Unfortunately, words can be ambiguous - hence there is an absolutely essential need for really good subject metadata. A search engine can be "tuned" to search tags before searching the free text and the use of tags, taxonomies, and controlled vocabularies not only allows users to query, but also to browse. Such use also allows search engines to suggest topics using type-ahead or "did you mean" and to leverage synonymy to deliver the same relevant results from various inputs.

He gave the example of the Public Library of Science (PLOS) where such taxonomies are used. It has a thesaurus of more than nine thousand preferred terms and they apply about eight terms per article. It also has about four thousand synonyms. These are applied to documents and are actually exposed when users browse so that the user can click on the term for another search. They are also used to redirect search queries for synonyms. If the user believes that a term has been incorrectly applied they can click on it and notify PLOS - they use crowd-sourcing to keep their taxonomies honest!

A second example came from JSTOR - specifically JSTOR labs - that actually uses the full article as the query using a combination of traditional taxonomies and naïve Artificial Intelligence topic modeling. The searcher simply drags or drops a document into the search box. It can suggest related content for research and for bibliographies. Go take a look. It is experimental, successful, and really something quite novel (see: <https://www.jstor.org/analyze/>)!!!

Kasenchak closed by saying that the practical implementation of semantic search can be easy or complex. For the first route a content provider can use simple things such as already-existing search software tuning/options, enable fuzzy matching, use weight fields, document types, etc. where appropriate, and use dates to deliver recent results. This can be done directly if they have their own search software or via their distributor. The next level is more complicated and involves the creation of taxonomies and tags, the building of Knowledge Graphs, and a better understanding of the users of their data via the creation of user profiles, tracking, and keeping a history of user behavior, and other targeted means. Kasenchak then ended and turned to the remaining two speakers to go into specific examples.

Kasenchak'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*.

13. Making semantic search work for you

The second speaker in this session was Travis Hicks, Director of Web Operations at the American Society of Clinical Oncology (ASCO), who discussed the steps that organizations can take to optimize their digital content for discoverability when it is indexed by a semantic search model, including content structure and metadata strategy. In addition, he looked at the steps an organization can take to better understand how far they should dip their toe into semantical waters when it comes to their own internal search engines, including the benefits of structured thesauri.

Hicks opened by saying that content providers must understand all aspects of search because external search engines are the number one method of choice for those seeking to discover unknown content. He added that the wrinkle is that search queries take many forms. As a result, all types of search queries need to be taken into consideration when constructing content - it needs to be optimized for external discovery. (*Note:* this is related to the concept of "Mental Models" discussed by the next speaker).

He noted that good content is discoverable content and, as Kasenchak mentioned earlier, Google parses content, not only for keywords, but also for phrasing in order to provide additional insight into relevancy (the Harrison Ford vs. When is Harrison Ford's birthday? example above). In doing this Google is able to identify terms that are semantically-linked in its analysis of billions of documents and has dis-incentivized poor content that is limited to keywords. He added that content with a high-user value is discoverable content and that there are ten steps that content providers can take to boost search results. These are as follows:

1. Focus on your users and their intent in content
2. Write clean, concise copy
3. Create links to reliable, high-quality internal and external content
4. Use structured schemas that leverage updated html tags (see: <https://schema.org/>)
5. Use bullets and organized lists (these are machine-preferred)
6. Ensure that your content is mobile-friendly
7. Utilize taxonomies
8. Make sure site performance is strong - fast loads rank higher
9. Ensure that the physical address is machine-readable
10. Pay for it with Google Ad Words (then again, maybe not - this can get expensive!)

He noted that it is important to understand user intent and to do this you need to review internal analytics; e.g.

- What types of queries do your users use (e.g., keywords vs. natural language)?
- What are the topical buckets of your queries (e.g., what terms/concepts are users generally looking for)?
- What terms are searched, but do not produce any results?
- Are your users more likely to search internally or externally?
- If you index content from multiple sites, do users search to get to those sites?
- If you have facets, do users actually use them?

He added that it is equally important to conduct user research; (e.g. what is the overall level of satisfaction with search results and if satisfaction is low, what trips them up) what areas or content types are the least discoverable and what facets may be helpful in enabling the users to access content they are seeking?

He agreed with Kasenchak in that he encouraged all content providers in the audience to create and support their own thesauri which he defined as a hierarchical classification system of terms. A good thesaurus will employ synonyms and facilitate the identification of semantic relationships. The addition of metadata will also enhance discoverability, and he added that synonyms can be utilized by search platforms such as Solr (see: <https://lucene.apache.org/solr/>). Indeed, the use of a thesaurus will enable a content provider to identify previously-unknown relationships across their content.

Hicks cautioned that it is very important to understand both the potential upside and the limitations of your internal search engine - it is unlikely that it is equivalent to Google. Know what it can and cannot do, and if there is functionality that is not being leveraged, identify what needs to be done to unlock the system's capabilities. Other reality checks that are equally important are to know the staff resources that you actually have available: Do you have internal development staff? Do you have individual(s) that are involved with ongoing system evaluation? Do you have a search governance plan for your search platform; and is there someone ultimately accountable for search performance - where does the buck stop?

He used his own organization as a case study. Their website provides information to their members as well as to the general public. There are about 4.5 million page views on an annual basis and about one hundred and eighteen thousand unique searches were executed in 2018. The site has two primary user groups - "heavy" users such as board members, volunteers, and internal staff, and "occasional" users, such as the average member, meeting attendees, pharma staff, and the general public. The major challenge that they were facing was the fact that they use Solr as its search engine and the specific model and relevancy algorithm was last updated in 2014. In fact, while there had been some ongoing maintenance, no major work done had been done since the last update. The primary issues that needed to be addressed were a general dissatisfaction by all user groups with search result relevancy - indeed, relevancy tuning had not kept up with content and business needs; and the fact that user expectations had shifted because of their experience with Google - they needed a new user interface, new technology, and new functionalities.

An examination of their search logs told them that searches were primarily keyword - there were only twenty-two searches that could be considered natural language searches and only eleven that could be considered advanced. They found that they lacked content for some queries and that other content needed to be refocused. They also found that more than half of the search queries involved basic cancer types, treatments, etc. So what did they do with the information?

Their plan for improvement included, among other things, adjusting proximity parsing, enhancing taxonomies, adding new/improved search models, implementing a relevancy score floor, and improving the user interface. The relevancy score floor was implemented to reduce the number of overall results (noise). If content doesn't reach a certain relevancy threshold, it will not be presented as a result and the threshold is determined based on testing after initial changes were made (e.g., where do results tail off?).

With regard to the user interface they added a type-ahead, contextual searching (Did you mean?) and they distinguished actions such as paying dues from search.

Success was measured via user experience testing that provided feedback from end users on the search experience over time as improvements were implemented. The first step was to have users complete search-related tasks with the search functionality and existing user interface in order to document usability of search/satisfaction with the current state. The second step was ongoing testing: as major milestones were met to improve search, users were asked to perform the same set of tasks that they completed during the baseline testing. By comparing the results of each round of testing, they confirmed that improvements had a positive impact on search usability and identified additional areas for improvement. The third step was to rinse and repeat.

For the long-term they will look at machine-assisted relevancy optimization; use SOLR logs and analytics to aid indexing and relevancy; consider the implementation of personalized search that could initially be based on simple user history, including profile data, later adding known online user behavior and individual click data. They will also consider voice-assisted search.

In closing, Hicks offered the following conclusions: first and foremost, understand the expectations of your users, both internal and external; develop a strategy for content optimization and stick to it; high-value user content is, by its very nature, findable and discoverable; there is no silver bullet for improving internal search - you must constantly reach out to your users; and finally, all efforts need ongoing evaluation and iterative testing of improvements.

Hicks' slides are available on the NFAIS website.

14. Information design considerations for effective semantic search

The final speaker in this session was Duane Degler, Strategic Design Consultant for Design For Context, LLC, a consulting firm that specializes in usability and user-experience design (see: <https://www.designforcontext.com/>). Degler opened by saying that semantic search seeks to enhance the meaning in content and to more closely align the searcher with the available information resources. As a result there needs to be a strong user-centered aspect in order to unlock the benefits.

Degler noted that users are highly- knowledgeable and that their search experiences are longer. Today, documents are larger, more topics are covered, and there is a lot of repeated information so it is difficult for users to know what is and what is not relevant to their search. He talked about "Mental Models" and I was not sure what he meant, but apparently that is user experience design jargon. Basically, a mental model describes how you or I would do something. It's our thought process. It's our expectations [40].

Degler said that searchers have a variety of Mental Models. These are: survey ("something about..."), targeted ("find something I know..."), exploratory or archival ("do not miss anything"), routine ("regular searches"), or collaborative (working with multiple authors). The use of semantics can be very helpful when it comes to these "Mental Models," but we must understand our users and use semantics optimally, such as in navigation and term relationships.

He asked if the meaning of a search term includes context and commented that a user's task has its own language. The task (an element of user context) shifts perception of what matters in content, so we need to ask if our sites and content need to respond to tasks. Usually, if users are clear on the relationship between information and task, then their focus on the information subjects should be sufficient and they can draw meaning based on their context. But in dense information environments, users may need help,

and the use of semantics can provide it. In particular, the use of semantics is extremely useful when it comes to search, content navigation, and term relationships.

We need to know if an internal search is a first stop, a next step, or a last resort and need to consider the user's journey and path analysis. We also need to consider the semantics of our site as the language used for navigation, links and headings sets user expectations of relevance. Semantics plays a role in focusing search, in providing the pathways for browsing, and in building relationship structures across terms and concepts.

In closing Degler noted that as we model content, we must recognize that its character, structure, and context all matter.

Degler's slides are available on the NFAIS website. Readers may also be interested in slides from some of his related talks that can be accessed at: <https://www.designforcontext.com/insights/search-mental-models>.

15. Democratizing data - whose job is it?

The final day of the conference opened with a plenary session by Dr. Daniel Barron, a resident psychiatrist at Yale University. The focus of his presentation was on the challenges that we face today as more and more personal data is captured - what information researchers need, how others use it, what the implications are, and what policies are needed. His presentation on the need for open science and data sharing was refreshingly logical and non-combative.

He noted that he is a clinician - both a resident physician, clinical neuroscience trainee, and that he is also a researcher. His dissertation was on brain damage in patients with temporal lobe epilepsy. He uses open science tools every day, but has been warned that science is competitive business and to prove a point he gave the example of Jack Gallant, a neuroscientist, who published research in *Nature*, but did not share his data for fear of competition. Barron noted that what followed was "trial by twitter." Gallant was publically shamed for not sharing his data and for a period of three weeks there were multiple twitter threads that included an attempt to get him banned from *Nature* and to lose his research funding (see: <https://twitter.com/gallantlab/status/1014932001734864897?lang=en>). Gallant himself was a user of social media, but at the time of the NFAIS conference he was relatively quiet.

Barron went on to say that there are many ways of sharing data. For example, one is to deposit in a publically-available repository or posting on a website. He noted that the latter is not a common practice in his discipline because of the data file sizes. They use MRI scans (6-7 mg each), usually eight per subject, and a study could include a hundred patients. Neuroscience data usually resides in a non-public repository that requires a credentialed login and password, or a credentialed login with certain strings attached such as citation requirements, authorship agreements, and peer-to-peer collaborative agreements.

Researchers share data to facilitate transparency and reproducibility and to avoid duplication of studies and bias in reports. Also many funding agencies around the globe demand that the results of government-funded research be made accessible. He noted that in the U.S. clinical trial data is considered confidential. The U.S. only requires that a summary report (including the questions asked) be filed. However, if the FDA needs to see the results they can request and have the right to see the details. As of 2014, the European Union has said that clinical trial data is not confidential. Ownership of data depends upon where you are located.

Researchers also share because data is expensive and he gave the following example. Between 1990 and 2011 there were about twenty-two thousand MRI studies published. He said that if you take a conservative

estimate that these papers represent twelve thousand studies with twelve patients per study, the time required to do the MRI scans would be one hundred and forty-four thousand hours (one hour per MRI). At Yale the cost per MRI (excluding ancillary labor costs) is six hundred dollars, so the cost of those studies - just for doing the MRIs - was \$86.4M.

Researchers share to exploit a limited and expensive resource through thoughtful stewardship. He noted the open fMRI group at Stanford that facilitates the uploading of these data sets to make them accessible. The data is put in a standardized format and tagged for ease of retrieval (*Note:* Open fMRI no longer exists is it now Open Neuro (see: <https://openneuro.org/>), but all files uploaded to fMRI remain available (see: legacy.openfmri.org). Sharing is essential because science is collaborative. Barron noted that he relies on the assistance of computational data scientists with whom he has to share data sets. Also, researchers build on the work of their peers - yet another reason why they share data - they share to expand and advance science.

If sharing is so important, why the battle to control data? One of the reasons Jack Gallant did not share his data was that he did not want his work to be pre-empted as his grant funding could be jeopardized. Others may be seeking patents for themselves or their organization and future revenues are on the line. Clinical data often includes personal data that is protected by law such as HIPAA (the U.S. Health Insurance Portability and Accountability Act of 1996 - see: <https://www.hhs.gov/hipaa/index.html>). Also, if a researcher has an idea for a potentially- valuable clinical, significant investment will be required to bring that idea to fruition and investors need to be shown that the intellectual property that the basis of the idea is safe and protected; i.e., that the research results are indeed proprietary. Investors expect a return on their investment.

Barron went on to look at who actually owns data and noted that there *are* legal owners. Sometimes it is the person or organization who paid to collect the data such as a pharmaceutical company. Sometimes it is the institution where it's collected such as a university. And sometimes ownership is specified in a legal contract. He said that usually someone or something owns data and that data ownership is a legal issue not an ethical or moral one.

He noted that in a world of open science there are many discussions about who should control ownership. He said that it is not an issue of who *should*, but who *does*. In academia it is usually specified in a contract and he said if researchers do not like the current state of affairs they need to change the system and he believes that is happening right now. The Twitter storm against Jack Gallant is just an example of the passionate feelings surrounding the sharing of data. But he views that to be a bottoms-up approach to change and he strongly believes that a tops-down approach such as the proposed legislation from NIH, is better [41]. Why? Because there are many types of data, some of which needs to be proprietary (clinical data) and some of which needs to be quickly shared. Also, there are many stakeholders, and rationale and constructive conversations need to be held (sounds like Jon Tennant's statement that reasonable and civil conversations are more effective than reactive and disruptive social media rhetoric).

In closing, he referenced an article that he just published on Open Science [42] and a podcast response to that article from ORION Open Science [43].

Barron's slides are accessible on the NFAIS website.

16. Shark tank shoot-out

The second session of the day was a "shark tank shoot out" in which four start-ups competed.

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 on a future NFAIS Webinar). This session has been a “tradition” at the conference since 2016 and can be quite entertaining – and informative as well. This year was no different.

The judges for the session were: Kent Anderson, CEO, Redlink; Ann Michael, President and Founder, Delta Think, Inc.; and Jignesh Bhate, Founder and CEO, Molecular Connections.

16.1. A maritime vessel index

The first presenter was Peter McCracken, Co-Founder and Publisher, ShipIndex.org (<https://www.shipindex.org/>), who spoke about a database that he has created for both maritime and genealogical research. He noted that the database can help you learn more about the ships that interest you – and it will lead the user to the books, journals, magazines, newspapers, CD-ROMs, websites, and online databases that mention them. Simply put, ShipIndex is, at its core, an index.

The business model offers both free and fee-based services. To provide access, the database uses a “freemium” model, in which over 160,000 citations are freely-available to all, without any registration or payment. The full database, currently at nearly 3.4 million citations, is available through subscription. Individuals can subscribe for a set period of time - from \$6 for two weeks of access to \$65 for a year of access - or for a recurring monthly fee of \$8, with coupon codes available to bring that down to \$6 per month.

McCracken admitted that throughout its ten-year existence, ShipIndex has faced a number of challenges and that these continue, but that he and his team are committed to seeing the database succeed. He also alluded to a new business relationship that will help him reach his objective.

I found the database very interesting, but as even McCracken said – it is a very niche product. However, he has a history of success. In 2000, he co-founded Serials Solutions with his brothers and a high school friend and the company was acquired by ProQuest in 2004. Who knows where ShipIndex may end up?

None of the slides from this session are posted on the NFAIS website. However, a paper based on McCracken’s short presentation appears elsewhere in this issue of *Information Services and Use*.

16.2. Information streaming

The second presenter was Violaine Iglesias, CEO and Co-Founder [44] of Cadmore Media, founded in June 2018, a company that provides video and podcast streaming for scholarly and professional organizations (<https://cadmore.media/>). She opened by asking the audience to do a show of hands (1) if they ever watched a video on YouTube (everyone), (2) if they ever wished that the longer videos were shorter and got to the point more quickly (everyone), if they ever wished that the content of the video was available to read (just about everyone), and if they ever stopped watching because the video was either too long or too boring (everyone). With that, she said that everyone should be interested in her company because they are in the business of making video and audio material snappy and engaging while being informative at the same time.

She noted that YouTube is the second most popular search engine and that all content providers need to be visible on that site. Video streaming [45] is growing and is very popular with the younger generations and it is also widely-used as a teaching tool. She added that even surgeons use it to learn and fine-tune their skills and added that she would feel better as a patient, if such teaching videos are curated and peer reviewed. She said that watching a video as opposed to reading a document can be more challenging – they cannot be quickly scanned, the viewer may have difficulty if the speaker has an accent or uses unfamiliar

terms, etc. Her organization actually supplies manuscripts of the videos, creates metadata, and provides workflow tools and a media player to help content providers make their content usable and findable via streaming media.

Iglesias gave a brief demo of the video player and showed how a transcript is displayed (transcripts can be provided in any language). The viewer can go to any part of the transcript to quickly access that part of the video – an asset for lengthy teaching tools. The demo was impressive (see: <https://www.cadmore.media/playerdemo>).

Their major competitor is YouTube who does not offer as many services as does Cadmore Media such as the creation of transcripts, metadata and the assignment of DOIs. The service has a tiered subscription model for large and small organizations and is based upon content, volume and usage.

A paper based upon Iglesias presentation appears elsewhere in this issue of *Information Services and Use*.

16.3. Credit reports for responsible science

The third presentation was by Rebekah Griesenauer, Data Engineer, Ripeta and Leslie McIntosh, CEO and Co-founder of Ripeta (<https://www.ripeta.com/>), an organization that is focused on improving the reproducibility of science by creating what is basically an automated credit report for responsible science.

They said that the Ripeta software identifies and extracts the key components of a research report/article, thereby shortening and improving the lengthy publication process while making the methods more easily discoverable for future reuse. This is important because scientific research is based on the principles of the scientific method, which assumes that research can be reproduced in future experiments. Yet, an estimated fifty to eighty-five percent of research resources are wasted due to the lack of reproducibility. This not only wastes resources, it reduces the confidence in science as a whole. Ripeta's goal is to provide researchers, publishers, and funders with a streamlined method for assessing the quality and completeness of the scientific research. They do not judge the quality of the science that is being reported - they are evaluating the robustness of the report. Does there appear to be sufficient information to reproduce the experiment(s) contained therein?

Digital Science (<https://www.digital-science.com/>) is a partner in this effort and a demonstration is available at: <https://demo.ripeta.com/> - definitely worth a view! Basically a paper can be uploaded and scanned to see if best practices for reporting research have been followed and a "credit report" is automatically generated. A new software release is planned for April 2019.

A paper based upon their presentation appears elsewhere in this issue of *Information Services and Use*.

16.4. Using artificial intelligence to identify concepts

The final speaker was Nicole Bishop, Founder and CEO of Quartolio (<https://quartolio.com/>), a company founded in 2016 and one that has created a cross-discipline research platform that transforms scientific documents into research intelligence with Artificial Intelligence (AI) [46]. She noted that researchers read only about one percent of scientific articles and one can only wonder what key research is going unread. Part of the problem is the paywalls to journals as well as the length of the publication process and she believes that the open access movement will resolve these issues. The number of research publications is another part of the problem which she believes can be solved using AI. Her organization is attempting to leverage the growing body of open access research publications to create an ontology of scientific research through the application of AI.

At Quartolio they are using natural language processing to extract information from millions of OA articles to connect the dots across scientific disciplines and identify concepts that are then aggregated into “folios” in a digital library. It is a cloud-based platform that uses library science and AI so that researchers can “discover, manage and curate research” in the folios. Each is a collection of all the documents that the company has compiled on a single scientific concept.

They also provide a service to organizations to do a similar analysis on their internal documents. Clients upload their documents and the system applies the AI to “connect the dots.” They can process about five thousand documents in six hours. To date they have \$290K investment funding and as of the NFAIS conference they were in the process of closing a \$1.5M round.

The judges closed the session and later in the day announced the Shark Tank winner: Cadmore Media.

17. Lightning talks

The final session of the morning was a series of lightning talks (six minutes duration each). There was no specific theme. They were to be “short, concise presentations that include sensible solutions to problems, overviews of new projects, news, etc., with the goal of sparking ideas and debate.

17.1. Discovery

The first speaker was Mark Gross, President, Data Conversion Laboratory (DCL, <http://www.dclab.com/>) who talked about the challenge of having your content discovered. For some searchers Google delivers – for some it does not. He used a search on Brian May as an example - the results do not bring up any of May’s scholarly articles (refer back to Kasenchak’s presentation for more examples). He referred to an article in *The Guardian* [47] where it said that it takes an average of fifteen clicks for a researcher to access an article! He noted that Scholarly Publishers have a complex web to navigate as users seek content through many channels and from many locations, using a diverse array of devices. There is no “standard” XML format across the discovery vendors (EBSCO, OCLC, ProQuest, etc.). He noted that the challenges for publishers is for them to bring together their content and metadata and deliver to all of their vendors at the right time. But how can they do that?

Gross then went on to described the services that DCL offers that can content providers to meet that challenge. They offer digitization, taxonomy creation, semantic enrichment - all of the key discovery factors that were mention earlier by Kasenchak, Hicks, and Degler. They have been in the business since 1981, have seen the industry evolve, and have the knowledge and experience to be of assistance.

17.2. Facilitating computational peer review and research reproducibility

The second speaker was Pierre Montagano, Director of Business Development, Code Ocean. If you read the section on the Miles Conrad Lecture you know a bit about the organization as Marty Kahn, the 2019 Miles Conrad Lecturer is the Chairman.

Code Ocean (see: <https://codeocean.com/>) is a cloud-based, open, online code execution platform that integrates with any scholarly platform. It provides researchers and developers with an easy way to share, discover and run code that is published in academic journals and conference proceedings. It allows users to upload code, data, or algorithms and run them with a click of a button. The platform enables reproducibility, verification, preservation, and collaboration without any special hardware or setup. Code

Ocean provides next generation tools to facilitate digital reproducibility, where users can access a working copy of a researcher's software and data, configure parameters and run it regardless of the users' operating systems, installation, programming languages, versions, dependencies, and hardware requirements.

Montagano reported that Ocean has recently launched a new service that provides publishers with a private link for executable code uploaded by the author during submission. Using container technologies, code execution is agnostic to programming languages, versions, or operating systems. The link can then be shared with peer reviewers who can easily change parameters, modify the code, upload data, run it again, and properly vet the submission. This will hopefully empower reviewers to conduct a more rigorous review of the science and help ensure reproducibility. This is being done in partnership with a number of publishers, including Nature and Elsevier.

Code Ocean has a history with NFAIS. Simon Adar, Code Ocean Founder and CEO, actually participated in the NFAIS 2017 Shark Tank Shoot out the same year that he launched the company. At that time he noted that more and more research results include actionable data or code, but that the dissemination of that code relies on individuals to set up environments to reproduce the results. Code Ocean did not win, but based upon Kahn's earlier comments, it looks as though it is doing quite well.

17.3. Author choices in an open access world

The third presenter was Serena Tan, Senior Editor, Publishing Development, John Wiley & Sons, Inc. She opened with an overview of major Open Access [48] (OA) initiatives that have taken place over the last two decades.

- 2000: PubMed Central is founded and BioMed Central formed in 1998 becomes the first commercial "born Open Access" publisher
- 2002: The Budapest OA [49] initiative is signed; PLoS is founded and funded
- By 2006: Elsevier, Wiley, Taylor and Frances, the Royal Society of Chemistry, the American Chemical Society, and more all launch hybrid OA programs
- 2008: NIH mandates a twelve-month Green OA policy

She noted that since then OA has really gone mainstream and publishers have been developing new models; e.g. Wiley launched a full Gold OA program. PLoS One has become the world's largest OA journal. The USA expanded Green OA to all federal funders. Funders in China and India have also launched Green OA policies. The Gates Foundation launched a Gold OA mandate, and this year Wiley signed a Read and Publish deal in Germany that covers the cost of OA publishing and subscriptions for their hybrid titles. The European Union has announced OA 2020 [50] – and of course there is the ever-present Plan S (see: <https://www.coalition-s.org/>).

Tan said that Wiley believes that publishers and scientific societies play an essential role in enabling researchers to do their best work and that Wiley embraces Open Access. It continues to develop sustainable new business models that support author choice and accelerate open access and they innovate to create new products and services around open access and open research in order to meet researchers' needs.

She noted that authors want to choose how they disseminate the results of their research; that not all authors have the money to spend on Article Processing Charges (APCs); and that they want publication outlets that match their desires and needs around scope, audience, and speed. As noted several times throughout this conference, in order to meet these needs Wiley became the first publisher to partner with Projekt DEAL for a countrywide deal in Germany to better address the growing research market and evolving needs of researchers. Projekt DEAL is a consortium of German libraries and research institutions

commissioned by the Alliance of Science Organizations in Germany, represented by the German Rectors' Conference, the HRK. DEAL represents more than seven hundred mainly publicly-funded academic institutions in Germany, including the most important science and research organizations.

Tan went on to say that the agreement preserves author choice and promotes greater access while supporting great science and scholarship. It provides German institutions with access to read Wiley's entire portfolio of electronic journals (e-journals) back to the year 1997. In addition, corresponding authors from DEAL institutions will be able to publish in Wiley's Gold open access journals without worrying about their ability to pay the APCs. They will also be able to publish open access articles in Wiley's hybrid journals, at no additional charge, if they so choose.

In closing Tan said that Wiley supports author choice by: providing both Green and Gold open access options for authors; converting previously-hybrid titles to Gold OA when sustainable; supporting OA titles that select manuscripts based upon their potential to advance thinking in the field; launching sound science titles that support the reporting of incremental findings; supporting authors without access to APC funds; collaborating on sustainable solutions that best enable researchers to take advantage of these publishing options; and innovating in open research to support the aspirations of the diverse communities we serve.

Tan's slides are available on the NFAIS website.

17.4. What is a Library?

Scott Livingston, Executive Director of Library Management Systems at OCLC, was the next speaker and he gave quite an entertaining and informative presentation. He opened by saying that in the USA when people think "public library" they automatically think "books." The figure he quoted was seventy-five percent of all Americans have that reaction. He added that library usage is declining - people just don't use them as much as they did in the past, and that library circulation is declining as a result. So as a public librarian what do you do?

Livingston said that the heart of a library is not comprised of books, it is comprised of people, and that libraries really are about community engagement. But since the interests of communities are quite diverse due to factors such as geographic location, industries, etc., not all public libraries are equal. Hence, OCLC offers a new library system, Wise, that is actually a community-engagement system that allows librarians to more easily and quickly gauge the interest of their current and potential patrons. According to Marshall Breeding, "OCLC Wise has built-in tools to help the library develop and manage its collection in response to use patterns. The product relies on its internal, real-time data rather than having to rely on exports to a third-party service... These automated processes are based on policies and thresholds set by the library, which can be updated as needed and these processes result in a customer-driven collection development strategy [51].

Livingston noted that Wise uses key functions, such as circulation and acquisitions, to analyze usage data that can inform librarians how their patrons use or do not use their collections. It is designed around people, both library users and library staff, with the goal of delivering great public library experiences and enabling libraries to evolve as their communities evolve. And it allows library collections to reflect the preferences, not only of the entire community, but also at an individual branch level.

In closing, he said that he is Scott Livingston - and he is not a book!

Livingston's slides are on the NFAIS website and a brief paper based upon his presentation appears elsewhere in this issue of *Information Services and Use*.

17.5. *The future of information access*

The next speaker was John Seguin, President, Third Iron, LLC, a library technology company (<https://thirdiron.com/>) who discussed some of the technology approaches to making information access easier while in parallel reducing the piracy rampant in the information industry. Specifically he addressed RA21 (<https://ra21.org/>), Google's Campus Activated Subscriber Access (CASA - <https://www.igi-global.com/librarians/casa/>), and Third Iron's LibKey (<https://thirdiron.com/libkey-discovery/>).

He said that in order to make the information access process simpler, the systems must work with and understand the authentication mechanism that an institution (university, library, etc.) uses. They must understand what access a user is entitled to before they authenticate as well as understand the access rights to content – is it open access material or subscription-based content. If the latter it not available, the system must route the user to the institution's fulfillment mechanisms. The goal is to generate links as close to the content item (PDF) as possible.

He added that all three of the above systems have the same goals, but that they have different starting points and that all have limitations. CASA is only available when a searcher uses Google Scholar and it is IP-based. For RA21, an institution must use Security Assertion Markup Language (SAML) [52] and he noted that not all publishers participate in RA21. Also RA21 is not aware of entitlements and this can produce confusion for users. LibKey must be supported by the institution, but is not limited to searches via Google Scholar nor to participating publishers. He added that they have ten million users across more than seven hundred institutions in twenty-five countries.

In closing, Seguin said that because these technologies overlap in numerous ways, but do not compete with each other, users can easily utilize the technologies that are best for them to simplify their access journey. It should be noted that at the time of the NFAIS conference RA21 was still being developed. The final recommendation was just released by NISO on June 21, 2019 [53].

Seguin's slides are on the NFAIS website and a brief paper based upon his presentation appears elsewhere in this issue of *Information Services and Use*.

17.6. *A research data management academy*

The final speaker in this session was Jean Shipman, Vice President of Global Library Relations for Elsevier, who discussed the launch of a new free, online Research Data Management (RDM) Librarian Academy. The main audience for the academy training is librarians of all types: academic, medical, special, public, government, and school. However, the key audience is the practicing librarian who is unable to leave the workforce to obtain additional formal training in RDM principles and best practices. As an aside, she noted that researchers can also benefit from participation. The training modules will be available to anyone around the world who has internet access. There are eight total units within the curriculum and each unit may be taken alone or if all the units are completed, continuing education certificates will be issued to those who want them.

The "academy" is being developed by a team that includes librarians from Harvard Medical School, Tufts Health Sciences, Massachusetts College of Pharmacy and Health Sciences (MCPHS), University, Boston University School of Medicine, Northeastern University, Elsevier, and Simmons University. Simmons University will be the institution who will grant continuing education credits, but a fee will be required to cover the costs of providing such documentation. Shipman noted that this is a unique partnership between librarians, library educators, and a publisher.

The need for this training was demonstrated through interviews and surveys that identified gaps in current training offerings and highlighted what skills librarians and researchers need to contribute to their RDM success. An inventory of existing courses was prepared as well.

In closing, she noted that the Academy will be launched later this year.

Shipman's slides are on the NFAIS website and a brief paper based upon her presentation appears elsewhere in this issue of *Information Services and Use*.

18. Designing user experience

The final presentation of the conference was a very visual and humorous talk by Willy Lai, Vice President of User Experience at Macy's. He gave example after example of designs fated to provide dreadful user experiences. One was a hotel room safe that could only be locked by using your credit card – but then the safe was locked and your credit card was outside the safe. He showed photos of bathrooms where the toilet paper roll was totally unreachable – too far away. Same thing for a bank drop off window – far too high! The visuals were extremely humorous and he more than made his point.

Lai said that the definition of User Experience (UX) is “a person's perceptions and responses that result from the use or anticipated use of a product, system, or service.” [54] A bad UX-design can hurt business and Lai said that studies show that seventy percent of customers abandoned a purchase due to the amount of frustration derived from trying to make that purchase online. About sixty-seven percent of customers said that a bad online purchase experience leaves them with a negative impression of the brand. He used Zulily as an example. On their old website a customer had to fill out a form before they could even see what was being sold. Now you can browse all you want and only need to register when you want to make a purchase (*Note*: I used the site to check it out, and registration takes a second – not the hurdle that it once was!)

Lai said that good UX design has the opposite effect - it is really good for business (as Zulily learned). He said that a well-designed site can have as much as a two hundred percent higher visit-to-order conversion than one that is poorly designed. He added that some studies have shown that every dollar invested in ease-of-use can return anywhere from ten dollars to one hundred dollars and a side benefit is that people are less likely to abandon a purchase. He suggested taking a look at VWO (visual website optimizer) at <https://vwo.com/> and noted that the site has some good reading material on UX design (I checked it out and it does!).

Lai said that a good design needs three components: it needs to be business-viable, technically- feasible, and user-desirable. He added that users want everything, but actually need less - you need to learn what is at the core of what they want and he used Henry Ford as an example. Ford's customers did not know that they really needed a car - they said that what they wanted was faster horses. Lai said that you should not bring users in for validation of the final product. The strategic approach is to involve them from the very beginning. You need to know and understand their unmet needs and it gives you an opportunity to know what you don't know and to see the user's world. Lots of feedback brings clarity and if there is ever confusion – just dig deeper.

Lai closed with some guiding principles for UX design:

- Design for your target audience.
- Provide all the essential information at the upper part of the site “above the fold” (80% of the time that users spend on a site is near the top – they do not scroll down).

- Promote helpful information and make it look like relevant content (just like Amazon).
- Shrink or eliminate forms to be filled out, which will result in significantly more conversions and increase order values (learn from Zulily).
- Users do not read digital sites; they scan them first and then read what interests them.

His closing comment was remember that Frank Lloyd Wright said that he never designed or built anything until he visited the site and got to know the people who would be living in “his” house. Get to know your user!

Lai’s slides are not available on the NFAIS website.

19. Conclusion

When reading the title of the 2019 NFAIS Annual Conference, *Creating Strategic Solutions in a Technology-Driven Marketplace*, I thought that the speakers would focus on technology itself, but the real message, perhaps from my biased perspective, was that in order to adapt to changes in today’s information community, content - both new and existing - should and must be the focus of our attention. Technology’s chief role is to increase the value of that content - through enhancements, increased ease of access and use, all the way to the creation of new services and perhaps new forms of content.

From the opening keynote that discussed the use of the Internet-in-a Box to store and disseminate relevant, high-quality medical information for offline utilization in rural areas, to the rejuvenation of IET’s fifty-year old Inspec database via the use of semantic tagging - it was content that reigned center stage. The key message is that stakeholders in the information community *must* adapt to change in order to remain relevant.

The same can be said of almost every presentation. How University Presses are using technology to enhance the value of their content in addition to improving the ease of accessibility and use of that content. How libraries are using technology to offer the services that their users have come to expect from commercial services such as Amazon. How new entrepreneurial start-up companies are using technology to make the audio component of traditional videos readable and searchable. They all focused on improving the relevance of their content and services through the use of appropriate technologies with one additional caveat - all applied technology with the needs and expectations of their user communities in mind. The conversation was very informative and demonstrated that today’s stakeholders are not standing still wondering *how* to move forward - they *are* actually moving forward. To sum up, I rephrase from Vincent Cassidy’s presentation. He said that as an organization IET is no longer “traditional,” but rather has evolved into a one hundred and forty-eight year old “start-up” which is strange, exciting, and invigorating. He emphasized that this “new life” is not about technology - it is about data and innovative uses of data.

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 continued the tradition. What was new was the fact that more than one speaker requested an end to “unproductive” rhetoric from those on both sides of the Open Access movement and less use of social media that often provides a knee-jerk reaction to a situation rather than the expressing of a rationale, well-constructed viewpoint. Even Jon Tennant who is very much for Open Access, stressed that we need “to understand the changing roles of stakeholders such as editors, librarians, publishers, etc., reconcile changes across/between disciplines/communities having diverse norms, practices, and biases; and resolve - *with civility* - the major tensions that exist between all of the stakeholders.”

In closing I leave you with the following quotes:

“Our technology produces a state of chronic revolution” - Aldous Huxley - 1956.

So even if today’s “revolution” ends, be prepared for the next, and to help you prepare, keep in mind the following recommendation:

“After you’ve done a thing the same way for two years, look it over carefully. After five years, look at it with suspicion. And after ten years, throw it away and start all over.” Alfred Edward Perlman (<https://www.quotes.net/quote/16343>).

There is no information on a potential “NFAIS” conference in 2020. NISO indicated that they are considering the continuance of this traditional and well-respected event, but no decision has been announced. Watch for details on the NISO website at: <https://www.niso.org/>.

Note: If permission was given to post them, speaker slides used during the NFAIS 2019 Conference are embedded within the conference program at <https://nfais.memberclicks.net/2019-conference-program>, and if they are available, the term “slides” appears highlighted in blue next to the title of the presentation.

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), with subsequent studies at INSEAD in Fontainebleau, France. Contact: 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-one years has provided a forum in which to address common interests through education and advocacy.

At this conference it was announced [55] that NFAIS would possibly be merged into NISO, pending membership approval of each organization. This approval has been attained and the merger became official on June 30, 2019 [56]. It marks the end of one era and the beginning of a new one!

References

- [1] Proto-writing, *Wikipedia*, <https://en.wikipedia.org/wiki/Proto-writing>, accessed July 16, 2019.
- [2] The History of Paper, *Wikipedia*, https://en.wikipedia.org/wiki/History_of_paper, accessed July 16, 2019.
- [3] Printing Press, *Wikipedia*, https://en.wikipedia.org/wiki/Printing_press, accessed July 16, 2019.
- [4] The History of Computers, http://www.softschools.com/timelines/computer_history_timeline/20/, accessed July 16, 2019.
- [5] History of the Internet, *Wikipedia*, https://en.wikipedia.org/wiki/History_of_the_Internet, accessed July 16, 2019.
- [6] B. Marr, A Very Brief History of Blockchain Technology Everyone Should Read, *Forbes*, February 16, 2018, <https://www.forbes.com/sites/bernardmarr/2018/02/16/a-very-brief-history-of-blockchain-technology-everyone-should-read/#3651718e7bc4>, accessed July 16, 2019.
- [7] B. Lawlor, An overview of the NFAIS 2018 annual conference: *Information transformation: open, global, collaborative, Information Services and Use* **38**(1-2) (2018), 26, <https://content.iospress.com/journals/information-services-and-use/38/1-2>, Accessed June 20, 2019.
- [8] *Internet-in-a-Box*, <https://meta.wikimedia.org/wiki/Internet-in-a-Box>, Accessed June 17, 2019.
- [9] S.C. Grover et al., Comparison of the impact of wikipedia, UpToDate, and a digital text book on short-term knowledge acquisition among medical students: randomized controlled trial of three web-based resources, *JMIR Medical Education* **3**(3) (2017), <https://mededu.jmir.org/2017/2/e20/>, Accessed June 17, 2019.
- [10] S. Harrison, Why Wikipedia Medical Content is Superior, Future Tense, January 28, 2019, <https://slate.com/technology/2019/01/wikipedia-doctors-medical-knowledge-study.html>, Accessed June 17, 2019.
- [11] A. Gomez, Exploring Offline Access to Wikipedia: Dr. Samuel Zidovetzki on Wikipedia's role in rural health initiatives, Wikimedia Foundation, September 26, 2018, <https://wikimediafoundation.org/2018/09/26/wikipedia-offline-access-samuel-zidovetzki/>, Accessed June 17, 2019.
- [12] E. Schulze, Everything You Need to Know about the Fourth Industrial Revolution, Davos World Economic Forum, January 17, 2019, <https://www.cnbc.com/2019/01/16/fourth-industrial-revolution-explained-davos-2019.html>, Accessed June 17, 2019.
- [13] IDEAL -19, <https://library.osu.edu/events/ideal-19-advancing-inclusion-diversity-equity-and-accessibility-in-libraries-archives>, Accessed June 17, 2019.
- [14] Assessment Program Visioning Task Force and Athenaem21 Consulting. ARL Assessment Program Visioning Task Force recommendations. Washington, DC: Association of Research Libraries, December 4, 2017. Available from: <https://www.arl.org/wp-content/uploads/2017/12/2017.12.04-AVTF-PublicReport.pdf>, Accessed June 16, 2019.
- [15] J. Boone, New library publishing program launched to support faculty scholarship, September 13, 2017, <https://vtnews.vt.edu/articles/2017/09/unirel-ubiquitypress.html>, Accessed June 17, 2019.
- [16] National Academies of Sciences, Engineering, and Medicine 2018. *Open Science by Design: Realizing a Vision for 21st Century Research*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/25116>. It can be freely-downloaded at: <https://www.nap.edu/catalog/25116/open-science-by-design-realizing-a-vision-for-21st-century>, Accessed June 17, 2019.
- [17] *Minimal Viable Product*, *Wikipedia*, https://en.wikipedia.org/wiki/Minimum_viable_product, accessed July 2, 2019.
- [18] *OPACs*, *Wikipedia*, https://en.wikipedia.org/wiki/Online_public_access_catalog, Accessed June 20, 2019.
- [19] *Faceted Search*, *Wikipedia*, https://en.wikipedia.org/wiki/Faceted_search, Accessed June 20, 2019.
- [20] *Discovery Services*, Library Technology Guides, <https://librarytechnology.org/discovery/>, Accessed June 20, 2019.

- [21] G.P. Khiste and R.K. Deshmukh, Discovery services: an overview, *Library Research World* 2(2) (2017), 27.
- [22] B. Lawlor, An overview of the NFAIS 2018 annual conference: *Information Transformation: Open, Global, Collaborative, Information Services and Use* 38(1-2) (2018), 6–7, <https://content.iospress.com/journals/information-services-and-use/38/1-2>, Accessed June 20, 2019.
- [23] B. Lawlor, An Overview of the NFAIS 2017 Annual Conference: *The Big Pivot: Re-Engineering Scholarly Communication*, <https://content.iospress.com/journals/information-services-and-use/37/3?start=10>, accessed June 21, 2019, *Information Services and Use* 37(3) (2017), 285.
- [24] Based upon comments made during the public review process, implementation of Plan S has been postponed until 2021, <https://www.coalition-s.org/rationale-for-the-revisions/>, accessed June 21, 2019.
- [25] Publish and Access Agreement, Projekt DEAL and Wiley, https://pure.mpg.de/rest/items/item_3027595_7/component/file_3028230/content, accessed June 21, 2019.
- [26] Impact Factor, Wikipedia, https://en.wikipedia.org/wiki/Impact_factor, accessed July 2, 2019.
- [27] Robot Writes *LA Times* Earthquake Breaking News Article, BBC News, March 18, 2014, <https://www.bbc.com/news/technology-26614051>, accessed June 21, 2019.
- [28] B. Lawlor, An Overview of the NFAIS 2018 Annual Conference: *Information Transformation: Open, Global, Collaborative, Information Services and Use* 38(1-2) (2018), 8, <https://content.iospress.com/journals/information-services-and-use/38/1-2>, Accessed June 21, 2019.
- [29] K. Brown and O. Pourquié, Introducing preLights: Preprint highlights selected by the biological community, posted February 23, 2018, <http://thenode.biologists.com/introducing-prelights-preprint-highlights-selected-biological-community/news/>, accessed June 21, 2019.
- [30] J. Kaiser, Medical preprint server debuts, *Science*, June 5, 2019, <https://www.sciencemag.org/news/2019/06/medical-preprint-server-debuts>, accessed June 21, 2019.
- [31] C. Morris, Powering research reputations: using real-time reputation building as an incentive to aid researchers in the sharing and discovery of knowledge, *Information Services and Use* 38(3) (2018), 149–151, <https://content.iospress.com/journals/information-services-and-use/38/3>, accessed June 24, 2019.
- [32] The Novel Blockchain Consortium for Science: bloxberg@Blockchain Munich Meetup, October 19, 2018, see: <https://www.mpl.mpg.de/en/about-us/news/488-the-novel-blockchain-consortium-for-science-bloxberg-blockchain-munich-meetup.html>, Accessed June 20, 2019.
- [33] D. Yaga, P. Mell, N. Roby and K. Scarfone, *NISTIR 8202 Blockchain Technology Overview*, National Institute of Standards and Technology, U.S. Department of Commerce, January 2018.
- [34] Editor, Dave Kochalko Interview: How Technology could restore trust in science publishing, *The SciencePod Magazine*, March 20, 2019, <https://sciencepod.org/2019/03/20/dave-kochalko-interview-how-technology-could-restore-trust-in-science-publishing/>, accessed June 24, 2019.
- [35] R. Schonfeld, Why was Springer Nature’s IPO Withdrawn? *The Scholarly Kitchen*, May 15, 2018, see: <https://scholarlykitchen.sspnet.org/2018/05/15/springer-nature-ipo-withdrawn/>, accessed June 24, 2019.
- [36] J. Tennant, Academic publishing is broken. Here’s how to redesign it, *Fast Company* (2018), <https://www.fastcompany.com/90180552/academic-publishing-is-broken-heres-how-to-redesign-it>, accessed June 24, 2019.
- [37] K. Kelly, *New Rules for the New Economy: Ten Radical Strategies for a Connected World* Viking Press, 1998, <https://kk.org/mt-files/books-mt/KevinKelly-NewRules-withads.pdf>, accessed June 24, 2019.
- [38] B. Lawlor, An overview of the NFAIS 2017 annual conference: The big pivot: re-engineering scholarly communication, *Information Services and Use* 37(3) (2017), 291, <https://content.iospress.com/journals/information-services-and-use/37/3?start=10>, accessed June 24, 2019.
- [39] *Levenshtein Distance*, Informally, the Levenshtein distance between two words is the minimum number of single-character edits (insertions, deletions or substitutions) required to change one word into the other. Wikipedia, see: https://en.wikipedia.org/wiki/Levenshtein_distance, accessed June 27, 2019.
- [40] K.K. Berg, Information behavior and mental models, *Search Engine Land* (2010), <https://searchengineland.com/information-behavior-mental-models-40032>, accessed June 30, 2019.
- [41] *Request for Information (RFI) on Proposed Provision for a Draft Data Management and Sharing Policy for NIH Funded or Supported Research* (Notice: NOT-OD-19-014). Closed December 10, 2018. See: https://osp.od.nih.gov/wp-content/uploads/Data_Sharing_Policy_Proposed_Provisions.pdf, accessed June 28, 2019.
- [42] D. Barron, How freely should scientists share their data?, *Scientific American* (2018), <https://blogs.scientificamerican.com/observations/how-freely-should-scientists-share-their-data/>, accessed June 28, 2019.
- [43] Good Scientists Share Data, ORION Open Science Podcast, see: <https://orionopenscience.podbean.com/e/good-scientists-share-data/>, accessed June 28, 2019.

- [44] The other Co-founder is Simon Inger, a well-known Information Industry consultant and past speaker at several NFAIS conferences.
- [45] A Landscape Review of Streaming Media, Renew Publishing Consultants (2018), see: <https://renewconsultants.com/wp-content/uploads/2018/09/Streaming-Landscape-2018-Renew-Publishing-Consultants-FINAL.pdf>, accessed June 28, 2019.
- [46] See short video from EdTech Week 2017 at <https://www.youtube.com/watch?v=BqfiqAVB0jk>, accessed June 28, 2019.
- [47] Scientists Should be Solving Problems, not Struggling to Access Journals, *The Guardian*, May 21, 2018, <https://www.theguardian.com/higher-education-network/2018/may/21/scientists-access-journals-researcher-article>, accessed June 28, 2019.
- [48] *What is Open Access?* <https://www.springer.com/gp/authors-editors/authorandreviewertutorials/open-access/what-is-open-access/10286522>, accessed June 28, 2019.
- [49] <http://www.budapestopenaccessinitiative.org/read>, accessed June 28, 2019.
- [50] *New Initiative to Boost Open Access*, <https://www.mpg.de/openaccess/oa2020>, accessed June 28, 2019.
- [51] M. Breeding, *Smart Libraries Newsletter*, April, 2018, <https://librarytechnology.org/document/23500>, accessed June 30, 2019.
- [52] *Security Assertion Markup Language*, Wikipedia, https://en.wikipedia.org/wiki/Security_Assertion_Markup_Language, accessed June 30, 2019.
- [53] *Recommended Practices for Improved Access to Institutionally-Provided Information Resources. Results from the Resource Access in the 21st Century (RA21) Project*, NISO, June 21, 2019, <https://www.niso.org/standards-committees/ra21>, accessed June 30, 2019.
- [54] International Organization for Standardization (2009). *Ergonomics of human system interaction - Part 210: Human-centered design for interactive systems (formerly known as 13407)*. ISO FDIS 9241-210:2009, see Wikipedia, https://en.wikipedia.org/wiki/User_experience, accessed June 30, 2019.
- [55] T. Carpenter, NISO and NFAIS announce plans to merge, *The Scholarly Kitchen* (2019), <https://scholarlykitchen.sspnet.org/2019/02/14/niso-and-nfais-announce-plans-to-merger/>, accessed June 30, 2019.
- [56] Merger of Major Information Industry Associations Finalized, NISO press release, July 1, 2019, <https://www.niso.org/press-releases/2019/07/merger-major-information-industry-associations-finalized>, accessed July 16, 2019.