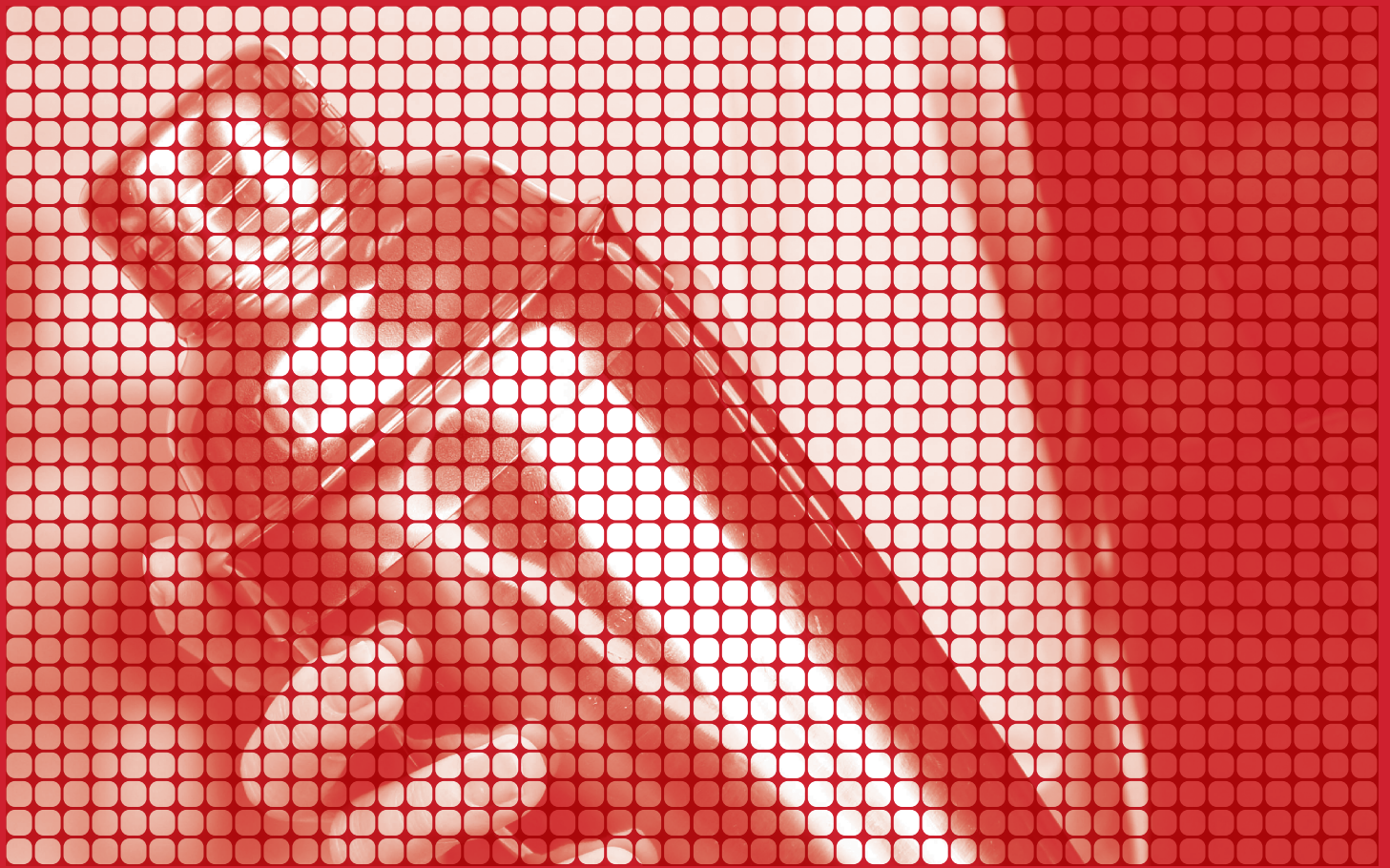




## **Workshop Report Series**

# Shaking Up Research

How To Thrive In – And Change – The Research Ecosystem

**Caitlin Trasande**, Digital Science

**FEBRUARY 2015**

#### **About Digital Science**

**Digital Science** is a technology company serving the needs of scientific research. We offer a range of scientific technology and content solutions that help make scientific research more efficient. Whether at the bench or in a research setting, our products help to simplify workflows and change the way science is done. We believe passionately that tomorrow's research will be different - and better - than today's.

Visit [www.digital-science.com](http://www.digital-science.com)

This report has been published by Digital Science, a business division of Macmillan Science and Education.

The Macmillan Building, 4 Crinan Street, London, N1 9XW  
[info@digital-science.com](mailto:info@digital-science.com)

Copyright © 2014 Digital Science

# Shaking Up Research

## Executive Summary

New players in the form of start-ups, grassroots organizations, and visionaries are shaking up how research is executed. This workshop presented panels focused upon four themes:

- How can technology better capture the research lifecycle?
- Accessing scholarly content: what are the new business models?
- How do we measure the influence and impact of scholarly research?
- The “Article of the Future” - what does it look like?

## Introductory Remarks

[Geoff Bilder](#), Director of Strategy Initiatives, CrossRef

Before launching into his talk, Bilder acknowledged that many of the ideas he was about to present were “*formulated with [Cameron Neylon](#) and [Jennifer Lin](#) (both of PLoS).*” Bilder then introduced his theme: cyberinfrastructure, underscoring that infra means below and under (“*think roads and tracks*”). Cyberinfrastructure “*are fundamental processes and tools which are supposed to transcend individual institutions.*”

“*I am routinely asked to give talks about the future of scholarly communications. We talk about getting rid of the PDF, collaborative tools... inevitably the conversation descends into the mundane: storage, identification. **One of the problems that we have is that there is reluctance to address infrastructure as a first class problem in our industry.** Fundamentally, infrastructure is boring. The only time it is interesting is when it is broken and not functioning. When a DOI breaks. When a dataset is unavailable.*”

Bilder suggested that one of the problems with infrastructure was that “*if it is working you don’t notice it.*” However, a perhaps larger (perceptual) problem lies in the fact that infrastructure is “*associated with big bureaucracies. There is a general built-in hostility towards it.*”

On the theme of funding, Bilder was especially sharp in his observations about the pattern he routinely sees.

“One of the patterns I see is when we talk about infrastructure with funders. A funder will invite the best and the brightest about the future of scholarly publication. I get invited to one per month, at least. And what happens is almost entirely predictable. The top researchers talk about their projects and say ‘by the way we need infrastructure.’ Researchers by and large are interested in cutting edge and the last thing they want to think about is infrastructure. So here’s the pattern. The funder has funds A. Researcher is studying B. The researcher triangulates and they propose projects that address A through B.”

Bilder explained that the problem with this approach is that, while the perception might be that through this mechanism there is funding for infrastructure, the reality is very different. Bilder provided an example.

“Someone says, ‘I have a project looking at phenotype-genotype interactions and as a byproduct of funding this I will create an identifier system.’ Well, by and large this never works. What happens is that the researcher writes code, puts it on a server, but when the funds dry up, the ‘infrastructure’ is gone.”

Bilder further cautioned that there was a danger in using certain words when framing infrastructure solutions.

“If you use words like ‘open,’ ‘distributed,’ ‘lightweight,’ or ‘framework’ - a lot of those words are code words for saying ‘we lack trust in our ability to make decisions.’ So rather than build a solution, we’ll build a framework.”

He also observed a rule associated with infrastructure, “**distributed begets centralized.**” He suggested that “**the web would be useless without google (or someone else).**” To put it another way “**distributed is a punt.**”

“A lot of times people punt difficult decisions about infrastructure by using words like distributed. The danger is that if you don’t acknowledge from the outset that someone else can come along and build a better but closed system, they will win. The other issue is that when you ‘go distributed’ this is code for having trust issues. **Deep down things are about trust.**”

Bilder shared how he learned to detect trust as a core issue in infrastructure projects.

“When we were starting up ORCID - which started as a CrossRef project - the idea was to create ‘author DOIs.’ When this project started, I interviewed stakeholders and we quickly learned that we couldn’t provide TRUST. A lot of the stakeholders were concerned that because CrossRef was publisher-run that there couldn’t be trust.”

Bilder shared that from his perspective, that “**the message was, great idea, but has to be trustworthy.**” Leveraging his stakeholder interviews with colleagues, he developed guiding principles for the then-nascent ORCID. Bilder then shared his views on some principles that could be used to found cyberinfrastructure, organized below by theme.

## Coverage

“Research infrastructure should transcend disciplines, geography, institutions and stakeholders. Chemists need to cite chemical structures. Biologists need to cite biological molecules. Funders are regional, so one will say ‘we will create a giant database’ and other governments have to create their own ‘version’ and now you have multiple versions. We have this repeated thing that we create these little solos.”

## Governance

*“Infrastructure should be stakeholder governed. If you have an organization that is simply run by libraries, the publishers will not be so keen to participate. One of the things we have to do is try to capture the full spectrum of stakeholders. You should have a policy of non-discriminatory membership. If someone is interested enough they are, by definition, a stakeholder.”*

## Transparent Operations

*“The organization cannot lobby. This is a controversial one. Lots of people fear organizations are just trying to keep themselves alive. We’re not saying the membership cannot lobby, but not the organization itself.”*

Bilder’s main theme was about sustainability, and he provided details about how sustainability needed to be built in tandem with the very mission of the infrastructure-building organization.

## Sustainability

*“If you are talking about infrastructure you have to consider what will happen if all things go wrong. What happens to the data? How can it be accessed? Who owns it? How is it all documented? You need to ensure sustainability. You have to be concerned that the thing is going to be there in the long run. There have to be other revenue streams like services. Bear in mind that endowments and/or time-limited funds are dangerous. You can’t run an electricity grid on grants. The goals should be to generate a surplus. Sustainable is not break-even - that is a very fragile position. If you don’t generate a surplus you are brittle. They need to create a contingency fund. Generate revenue based on services not on data. One of the things we often see is that we’re building new infrastructures to handle open content and open data. We do not want to be in a position where we need to make money by closing access to that data. We want to provide services. Revenue generations should be consistent with the mission. CrossRef’s mission is to enable linking between*

*papers. CrossRef's idea, was to charge publishers to resolve citations to DOIs. That was against our mission and it slowed our growth. Once we resolved DOIs freely, our growth exploded.”*

## Insurance

*“Do we really want one organization to support something so critical? ORCID has the ultimate insurance: forkability. Anyone can download the source code for Apache and can create their own versions. Forking a major project is actually a major event. But sometimes it is critical. So making sure the fundamental infrastructure is open source, that the data is available – these are the ORCID insurance policy if all else fails, if we turn evil, you can take the code and data and start up your own system and - critically - bring the community with you. You will win.”*

Bilder summed up his thesis for the audience:

*“Effectively I'm trying to say we've got to stop asking researchers how to fund infrastructure. Funders, institutions, publishers and other stakeholders need to recruit people who know how to run infrastructure. **As a community, we need to get away from this fear of infrastructure as a line item or cost center. There are many ways to make infrastructure sustainable.** It cannot be a byproduct of other research efforts, it must be a first class part of the research enterprise.”*

## Panel I: Capturing the Research Lifecycle

[Courtney Soderberg](#), Statistical Consultant at the Center for Open Science, led a panel discussion of the challenges associated with organizing research data, disseminating and discovering new research tools. Panelists included: [John Lees-Miller](#) Co-Founder of Overleaf, [Lars Holm Nielsen](#) Software Engineer at Zenodo, Arfon Smith of GitHub, [Robert Seigel](#) of Pubiscize, [Tim Gardner](#) Co-Founder of Riffyn, and [Rion Dooley](#) of Texas Advanced Computing Center.

[Courtney Soderberg](#), Statistical Consultant at the Center for Open Science began the lightning-style presentations.

*“Over the past couple of years there has been a growing movement to increase the openness and reproducibility of research. At the Center for Open Science, we want to allow researchers to continue to work with the tools that they use and like and connect with us.”*

She shared how the Center for Open Science is reaching out to the scientific research community through workshops and other forms of engagements to serve their mission. Next up, [Tim Gardner](#) Co-Founder of Riffyn described how a focus on quality, inspired by the best manufacturing quality practices, can support research efforts.

*“I’ve been in biotechnology for 20 years. I’ve seen how the adoption of rigorous quality practices, practices that lower the noise in experimental data, can dramatically improve research productivity. Reducing the noise not only reduces false positives, more importantly it reveals experimental signals that would otherwise go undetected (would-be false negatives).*

*The lessons of manufacturing quality are to start with a process design, then measure its performance, continuously improve it, and share it. With quality we can turn R&D into a supply chain of high-quality research processes, data and knowledge products.”*

Next up, [Arfon Smith](#) of GitHub shares a bit about how the massively successful start-up serves as an increasingly important tool for researchers.

*“At GitHub, we’re interested in tools that help people do research better together. Our scale is such that there is potential to make a large impact. We have over 10M GitHub users. Our motto is ‘Build better software together’ GitHub’s collaborative pull-request based model helps make this a reality.”*

Smith described how GitHub was born from the challenge of managing the source code of the Linux kernel. He explained that one significant challenge is that software is not recognized as a creditable research product.

**“There is a fundamental problem that you don’t have a good career if you write a lot of code in academia. We need a way of giving code academic credit.”**

Smith described some of the fundamental principles of running OS projects.

**“When we talk about open source lots of people think ‘open or die.’ That means open communications – everything having a URL – and, overall, working transparently and lock-free.”**

Another theme Smith commented upon was that of reproducibility. Smith shared that he considers reproducibility to be a workflow and cultural problem.

[Rion Dooley](#) of Texas Advanced Computing Center, next described the Texas Advanced Computing Center at UT Austin as a ‘Science-as-a-Service’ platform **“to allow you to build your digital lab.”**

Dooley described how the Texas Advanced Computing Center has been funded by the NSF for 10 years. It is free to use, 100% OS and available to fork on GitHub or BitBucket.

According to Dooley, educators as well as researchers found the platform helpful. For example, educators use it for discussion forums.

[Lars Holm Nielsen](#) of Zenodo began his presentation with an image of ‘lost briefcase’ flyer with **“10 years of data”** stored in a laptop gone with it. **“We want to keep data out of lost briefcases on buses,”** he announced.

Critical to this is using tools where researchers are **“already working.”** Zenodo is working with GitHub, for example so that **“the researcher does not have to think too much about it.”** In another example, researchers can **“export GitHub software to INSPIRE (high energy physics repository)”** seamlessly.

[John Lees-Miller](#) of Overleaf introduced the product he helped build, originally called WriteLaTeX, as **“online collaborative editing.”** He disclosed that Overleaf has helped over 150 thousand authors create over 1.5 million documents.

Lees-Miller shared how the project was inspired by EtherPad, a basic precursor to Google Docs. He felt that the application was good for text but not for figures or equations, as there was limited styling functionality. His goal in creating Overleaf was to **“make the scientific editing process neater and more integrated - moving people off email and attachments.”** Moreover, with Overleaf, **“when you are finished - you can publish”** to arXiv, figshare, F1000Research, PeerJ, and nine other publishing partners.

[Robert Seigel](#) of Publiscize shared how he was motivated by the fact that **“interdisciplinary audiences and public are poorly served”** by most forms of scientific communication. In his view there are two key ways to shape research communication for a wider audience: **“reduce complexity and reduce jargon.”**

Users of Publisize “**re-write their research article in 300-400 words.**” Publisize serves as the editor, helping to shape the article for a broader audience. “**We then help the scientists publicize their works.**”

Following these lightning style publications, the floor was opened for questions. The first concerned preventing “**lock in**” within any particular tool. A consensus among panelists: allow users to export their data using standard formats. An audience member asked, provocatively, how all of these tool should be linked together; suggesting it was confusing, rather than helpful, to researchers to have so many different tools and services available. The panel responded as a unified front that no one thing could be “**all things to all researchers.**” The key, rather was to “**solve your problem well and play nice with others.**”

The next question drew different responses from each panel member. How does each project/service reach out to its audience? Publisize’s Seigel found it productive to reach out to researchers via academic listservs. GitHub’s Smith suggested that “**academics have been the most abusive of our platform in the nicest way possible.**” Zenodo’s Nielsen admitted that while they were “not great” about spreading the word to users, nonetheless some universities had requested clones of the service for members of their academic community.

## Panel 2: Accessing Content: New Thinking / New Business Models to Accessing Research Literature

Chair [Alex Hodgson](#), Director of Marketing for ReadCube, led a panel discussion about the current challenges around the accessibility of scholarly content and the rise of innovative new models aimed to tackle them. Panelists included [Robert McGrath](#) CEO of ReadCube Access/Checkout, [Eric Hellman](#) Founder of Unglue.it, [Marguerite Avery](#), MIT Libraries / ex-MIT Press, [William Gunn](#), Head of Academic Outreach, Mendeley and [Jennifer Farthing](#), Director, Consumer Products, Ithaka

First up, [Robert McGrath](#) discussed ReadCube Access/Checkout and described highlights from a myriad of conversations with academics.

*“The biggest feedback was how to improve access. A lot of people hit paywalls and get frustrated. Users try to go to Google Scholar or Microsoft Academic Search, email the author or contact their librarian.”*

McGrath described ReadCube as a platform with desktop and mobile environments which featured reference management, an enhanced PDF reading experience, the ability to discover, annotate and share new content, and view usage metrics.

*“We are looking at how we can broker relationships between users and publishers. [The ReadCube service] ‘Checkout’ was the first access solution we implemented. At the point of hitting the paywall we introduced different price points. Many academics would not pay US \$35 for a paper. But if there*

*were lower price points would academics pay? It turns out when we introduced the price point, non-institutional users opted for the lowest price pt (\$6). Alternatively, the price can be paid on demand by the library with access provided by the university.”*

Next up, [Eric Hellman](#) of Unglue.it described his organization succinctly as “**a bookstore for free ebooks**.” His vision of creating a digital bookstore quickly met with reality, as he explained:

*“Book publishing has an extremely well-developed supply chain. The idea of Unglue.it is to connect readers and authors directly. However, a lot of [publishing industry] players survive on taking a percentage of the purchase price. Even libraries hate free ebooks!”*

Hellman then launched into a case study via a book about oral literature in Africa which “**put oral literature on the scholarly map**.” But he reported that

*“African scholars didn’t have access because it was too expensive. So we crowdfunded a campaign to make this a free ebook.”*

Hellman described a series of other books he targeted: a textbook for scientific computing, a cook book, a novel by a graduate student in mathematics, an English translation of a French book. The latter example set the stage for a discussion of fair use. Hellman capped his presentation with different funding models for free ebooks (pay-what-you-want, purchase, pledge).

[Marguerite Avery](#), MIT Libraries /ex-MIT Press shared with the audience that multiple factors are “**barriers to access**” of the scholarly literature. While price is a main driver, so are time (namly how long it takes to retrieve an article) and the format of the article. Avery noted that there are over 100 university mission-driven non-profit presses. While historically university presses published long forms (books, monographs), they were increasingly engaging with shorter format works. Avery highlighted “**Bits**” from the MIT Press, as well as “**Briefs**” from Stanford University Press as examples.

Next, [William Gunn](#), Head of Academic Outreach at Mendeley plunged into the topic of search as a barrier to access, “**search doesn’t work**.” Gunn suggested that search could be reimagined as “**push over pull**” and put the spotlight on “related papers” (papers topically related to the paper the user was viewing) as a mechanism to “**push**” search to the user. Gunn shared that at Mendeley, “**in our initial tests, researchers like recommendations**.” Gunn also suggested that technical enhancements to search, such as autocomplete and suggested keyword terms, could facilitate discovery of scholarly content.

The final presenter, [Jennifer Farthing](#), Director, Consumer Products, Ithaka, described the JSTOR program, “**Register and Read.**” She shared that JSTOR has programs to help non-affiliated researchers access content. Register and Read provides “**free, limited access for everyone.**” The program initially enabled access to 76 Journals and allowed participants to access up to three articles every two weeks. Several years later, the program has been a “**success story.**” Currently participants can access content from 1600 journals, 900 publishers, and features over 2M users. Who are these users? Farthing shared data slides showing that users are undergraduate and graduate students hailing from outside the major centers of scholarship.

## Lightning Demos - Round I

Workshop attendees next viewed parallel lightning-style demos given by Charlie Wong of [BioRAFT](#), Julia Hawks and Kelsey Rosell of [Symplectic Elements](#), Kathy Christian of [Altmetric](#) and Thomas Crouzier of [Connected Researchers](#).

## Panel 3: Measuring Content: Evaluation, Metrics and Measuring Impact

Chair [Gregg Gordon](#), President, Social Science Research Network, led a panel discussion on evaluation and impact “**beyond citations**” for articles, researchers and institutions. The panels stated goal was to challenge the way we think about metrics and provide some insights into their future. Panelists included: [Ashlea Higgs](#), Founder of UberResearch, [Mike Taylor](#), Research Specialist, Elsevier Labs, Elsevier, [Andrea Michalek](#), Co-Founder of [Plum Analytics](#).

[Mike Taylor](#) began things with a declaration:

*“I’m altmetrics number one fan... Metrics must be: community driven, fit-for-purpose, validatable, accurate, be built with open methods and data sources and lastly business model agnostic.”*

Taylor next suggested that the model for metrics and the most famous metric in scholarly publishing - the impact factor - has “**been distorted beyond all recognition.**”

He suggested that there is a “**big gap**” between what we’ve seen and what we hoped we would see in the development of new metrics.

*“We’re really at the beginning of understanding altmetrics. We hoped altmetrics would give a broader picture of the citation universe - it doesn’t. Alternative metrics are actually very narrowly focused. It is the responsibility of the researchers to make these caveats and point out their faults and limitations.”*

Taylor concluded by highlighting that, “*there isn’t a generally accepted definition for what social impact is. What questions does the community want? Academics have questions, deans have questions, funders have questions.*”

Next up, [Ashlea Higgs](#), co-founder of UberResearch focused on metrics from “*the funder perspective.*”

Higgs shared stories from UberResearch’s experience working with research funders globally and presented the problem statement around funders + metrics + impact.

Not only are funders asking the same questions as others, but also they are being asked to answer even more questions that reach into system-wide questions and impact on health outcomes, economic development, and more.

Only recently have the means – capabilities, data, know-how – started to catch-up with the demand.

By way of example, recently from a government health funder:

*“From my perspective the days are gone when publications are sufficient in and of themselves in reporting on the accountability of funded research. More and more we want to look at health impact and social impacts and organizational impacts and system impacts.”*

[Andrea Michalek](#), Co-Founder of [Plum Analytics](#) began her talk with a ‘*tour*’ of classic and new metrics. She touched upon Journal Impact Factors, Eigenfactors, SCIMago journal rank, SNIPs as well as Google journal metrics. She suggested that “*no single tool captures all metrics.*” Michalek suggested that Plum was positioned to capture all metrics, even “*data exhaust*” such as views on Amazon. Only this holistic picture can “*tell the stories of research impact.*”

Michalek shared that Plum measured the following major categories of metrics: usage (clicks, downloads, views, library holdings, video plays); captures (bookmarks, forks, fav, readers, watches), mentions (blog posts, comments, reviews, wikipedia links), social media (pluses (+), likes, shares, tweets) as well as citations (classic references in a published article). She suggested that social media is a great measure of “*how well promoted something is*” whereas citations, though the most lagging indicator, were “*the best measure of research.*”

She shared that Plum measures “*beyond the journal article*” and includes video, data sets as well as conference presentations. Because some work produces multiple versions, Michalek detailed how Plum unifies the object (OA version, published version, institutional repository version).

Michalek next launched into a case study featuring the University of Pittsburgh’s Plum dashboard.

*“Metrics provide a feedback loop. They are giving you data you can react to in real time and respond to what’s happening.”*

Her next case study was the organization “*Autism Speaks.*”

*“Autism Speaks used Plum Analytics to tell their donors what they are doing. They fund different proposal areas: biology, treatment & prevention, screening/diagnosis as well as dissemination.”*

Michalek struck an optimistic tone about the future and utility of measuring impact.

## Lightning Demos - Round II

Workshop attendees next viewed parallel lightning-style demos given by Mitar Milutinovic of [PeerLibrary](#), Ben Lewis of [WorldMap](#), Griffin Weber of [Harvard Profiles](#) and Eleni Castro and Elizabeth Quigley of [Dataverse 4.0 Beta](#).

## Panel 4: Publishing Content: The Article of the Future

Chair [Alyssa Goodman](#), Professor of Astronomy, Harvard University led a panel discussion on the research article of the future. Why is the article of today lacking? How can journal articles feature linked open data? How can we use the way researchers collaborate to create authoring tools? What does the article of the future look like? Panelists included [Tiffany Bogich](#), Co-Founder of Standard Analytics, [Tim Clark](#), Co-Founder of Annotopia; Assistant Professor of Neurology, Harvard Medical School, [Alex Johnson](#), CTO at plot.ly, [Dan Valen](#), Product Sales Manager at figshare and [Alberto Pepe](#), Co-Founder of Authorea.

Goodman opened the panel session by sharing that though her “**day job**” was being an astronomer (“*That’s a joke, being an astronomer during the day.*”), her interests extended into scholarly publishing.

*“My main interests are in making the research process much less painful than it is now, so that people will take that extra step in their research. We in astronomy like to think we’re the vanguard because our data have no economic value.”*

She introduced [Alex Johnson](#), CTO at Plotly. Johnson suggested that “*the figure is the most compact representation of the data in a research article. It pulls the readers in as well as ties meaning to the data. Put another way, “visualization is the core of the article.”*” Johnson suggested that

*“The key to the “**article of the future**” was to maintain the network of data sources, acquisition/analysis scripts, related information, and references that generated the figure, to let each reader interact with the data in their own way. We want to get to a place where the data stays with the figure. But also the network where the data came from, the scripts, everything you think is important [can be accessed].”*

Recalling the theme of “*solving your problem well and playing nicely with others*” (Panel 1), Johnson laid out a major goal of Plotly.

“One of the things we’ve tried to do with plot.ly is to make a universal translator for plotting programs people like: matlab, python, r, julia, javascript.”

Johnson described Plotly’s user base as diverse, including researchers from many different disciplines as well as students in high school and even middle school.

Next up, [Alberto Pepe](#), Co-Founder of Authorea began by suggesting that today researchers are

“doing 21st century research, written up on 20th century tools, packaged in 17th century format.”

He shared a series of figures (drawings) from an article written by Galileo and suggested that because Galileo shared his “*raw data*” he was “*doing open science*.”

Fast forward to today and the reality is that “*most of the time the data is too large, and doesn’t fit into the scope of a paper anymore*.”

Pepe suggested that we should include all the raw data in the paper itself – not as a supplementary material but as a repository.

[Tiffany Bogich](#), Co-Founder of Standard Analytics became “*obsessed*” with a question during her postgraduate studies. “*If you could redesign the publication submission process, what would it look like? What would the ideal format contain?*”

Bogich suggested three main features: machine-readable metadata, cross-platform and fully accessible. She suggested that the pieces of the puzzle are there, “*we are adding RDFa into CSS/HTML, with HTML and RDFa we have discoverability*.”

She suggested that the article of the future would be “*interactive, scriptable, and feature truly extendible behavior (web components)*.” In her view, automatic generation of semantic tags and peer review with certified reproducible results were critical. However, we are missing the “*end-to-end infrastructure*” to provide this.

Next up, [Tim Clark](#), Co-Founder of Annotopia; Assistant Professor of Neurology, Harvard Medical School introduced Annotopia by describing the key to his open annotation services platform. The WC3 open annotation data model (OADM) is different because, “*in the old annotation model, you add comments and they ‘stick’ to the text. In the new model annotations are held separately from the article*.”

He said annotations should be considered to be a first class object on the web - no longer a derivative object attached to the primary object. A more “**granular security model (e.g. just for lab vs whole world)**” would also be a boon to researchers.

Wrapping up the lightening presentations, [Dan Valen](#), Product Sales Manager at figshare suggested that we should “**differentiate between recommended, mandated, and enforced**” when it came to making research open.

*“Many funders are now creating policies in which funded projects must be open. At figshare, we are trying to capture data at every point of research. From the research management plan, experimental plan, data, figures and papers.”*

Valen stressed that while funders were mandating that data derived from funded research was to be made open, that this did not automatically mean researchers would comply. He then touch upon the central theme of the panel.

*“What’s essential for the article of the future? Repositories are essential. APIs are essential. Persistent identifiers are essential. Open access and liberal licences are essential. We need funders to mandate it, journals to enforce it, web-based data to be enabled and data to live in linked, discoverable repositories.”*

## About the Workshop

The **Shaking It Up** workshop was organized by [Amy Brand](#) (Digital Science), [Caitlin Trasande](#) (Digital Science), [Alex Wade](#) (Microsoft Research), [Christopher Erdmann](#) (Harvard-Smithsonian Center for Astrophysics), [Margo Avery](#) (MIT) and [Patsy Baudoin](#) (MIT Libraries). Held on November 24th, 2014 at Microsoft's NERD Space in Cambridge, MA, the workshop included a cross-section of participants to share success stories and challenges of innovating in the research space:

- What are the most prevalent pain points for researchers today?
- How are innovators in the research space reaching out to their target communities?
- How is the practice of research changing as new tools and services become available to researchers?
- How do these initiatives launch, grow and scale?
- The flora and fauna of the research ecosystem include scholars, libraries, universities, funders, governments and companies: how are the rules of interaction changing?
- What new partnerships are emerging?

The organizing group invited a cross-section of members of the community to the workshop. Over 150 people registered in person or virtually. The themes and demos proved to be very attractive to a spectrum of different members of the research ecosystem. Workshop attendees included:

**Universities** (Boston University, Brown, Harvard, Massachusetts General Hospital, MIT, Northeastern, Tufts, Temple, University of Massachusetts-Dartmouth, University of Massachusetts-Amherst, University of Texas at Austin, University of Pittsburgh, Worcester Polytechnic Institute, Yale).

**Research Institutes** (Cold Spring Harbor Laboratories, Marine Biological Laboratories, CERN).

**Startups and companies** (Aries Systems, Authorea, Elsevier, Emerald Group Publishing, figshare, GitHub, Mendeley, Microsoft, Overleaf, Plotly, Plum Analytics, PubScize, Quintesscent, ReadCube, Riffyn, Symplectic Ltd, ÜberResearch, Unglue.it, Vertex Pharmaceuticals).

**Organizations** (Association of Research Libraries, Boston Library Consortium, Center for Open Science, Connected Researchers, CrossRef, Data & Society, DuraSpace, GEM Consortium, Social Scientist Research Network (SSRN), Standard Analytics, PeerLibrary).



# Work smart. Discover more.

Our tools and services help at every stage of the research cycle.



**BIK>RAFT**



**labguru**



**Overleaf**



**SYMPLECTIC  
Elements**



digital-science.com