

ALPSP 2017, Noordwijk, NL
Sept. 13–15, 2017

Tuesday, Sept. 12

Spent the day in Amsterdam with Rowly Lorimer, Suzanne Kettley, and Adrian Stanley (Digital Science) recovering from the overnight flight and adjusting to European time. Talked with Adrian about digital tools for research, publishing, and classroom use, particularly in the humanities. He referred me to the free/freemium ReadCube reference manager for organizing and annotating a personal library of readings as well as Altmetric's free bookmarklet tool for tracking the metrics on particular articles. He also told me about Joe Karaganis's Open Syllabus Project, which datamines online syllabi to collate information and generate metrics on what publications are being taught. More on this project below from Joe's presentation on day 2. Shared a cab to Noordwijk with Suzanne in the evening.

Wednesday, Sept. 13

Day 1 of ALPSP

Registered and sat for the opening conference lunch. Met Toby Green, who referred us to his article, [“We've failed: Pirate black open access is trumping green and gold and we must change our approach.”](#) recently published in *Learned Publishing*.

Keynote 1: Trust, Truth, and Scholarly Publishing

Speakers: Lex Bouter, Professor of Methodology and Integrity, VU University Amsterdam
Kei Koizumi, Visiting Scholar for Science Policy, AAAS (American Association for the Advancement in Science)

This opening session focused on some of the primary problems and challenges currently faced by researchers and publishers in the sciences. Bouter argued that the greatest threat to research integrity is not flagrantly dishonest practices such as falsification and plagiarism but the gray area of “sloppy science” or “questionable research practices,” including selective reporting, selective citation, and questionable supervision on the part of researchers, and the bias toward “positive findings” among publishers. These practices are all incentivized by a professional system that rewards “high impact,” “breakthrough” studies over “open and honest” research practices. In response, he called for greater transparency of methods, protocols, and data, and suggested a turn to publishing on the basis of the soundness of research questions and methods alone, not results, as a way of correcting the bias toward positive results.

Koizumi, formerly of the Office of Science and Technology Policy under President Obama (an office that has been, like many others, effectively gutted under Trump), spoke on the diminished role of data and scientific information, suspicion of expertise, and crumbling trust in learned institutions in the US under the current administration, which has resulted in the decline of research funding and the “decoupling of science from policy.” Noting that the declining trust in science, while enhanced under Trump, is a long-term trend in the US, Koizumi urged “scientific integrity” (the use and communication of scientific findings by the government and policy makers) and public access of research results as a way toward reversing this trend. With the White House uncommitted to fostering a public engagement with science, the AAAS has stepped in to normalize rigorous science as a public value.

On Twitter (where I followed, and occasionally contributed to, the #ALPSP2017 backchannel throughout the conference), Kent Anderson noted:



Plenary 1: Learned Societies: Navigating a Sea of Change
Speakers: Hetan Shah, Executive Director of the Royal Statistical Society
Alix Vance, CEO, GeoScienceWorld

This session addressed some of the changing structures and practices of learned societies in response to new sets of challenges in recent years—challenges including the expectations of members increasingly being set by large corporations (i.e. a certain commercial “slickness”), non-traditional competitors, demographic and cultural changes, and the impact of institutional sales on membership as well as library budgets on subscriptions. Shah described some of the changes in the Royal Statistical Society he has overseen in the past five years in order to further the society’s goals of strengthening the profession and discipline, promoting statistical literacy, and ensuring that statistics are used ethically. Under his leadership, the society has shifted from volunteer to professionalized staff and identified “growing areas for income” (i.e. commercialized services such as training provision). Initiatives for raising the society’s public engagement include the “data manifesto”—a document distilling the society’s conclusions that can be distributed to politicians and other public servants to increase the profession’s visibility and improve policy and public impact—and the appointment of “statistical ambassadors,” young people coached by the society on how to engage with the media. Vance described the way GeoScienceWorld bundles services for publishing and distribution for multiple societies in the earth sciences. By combining “best practices of both nonprofit and commercial worlds,” the company prioritizes society independence and society-led publishing in a way that yields increased revenues and publishing royalties. In a cultural landscape in which scholarly societies rely on commercial-level revenues to operate, a company like this one deals with the business of economic growth so the societies can “focus on science.”

Announcement of finalists for the ALPSP Awards for Innovation and Publishing:
[Escalex](#): database of food regulation and compliance information collated from multiple government agencies

[Publons](#): platform for collecting profiles and activity for peer review, that generates verified records of peer review contributions

[Delta Think](#): open access data and analytics

[INASP AuthorAID](#): support, mentoring, resources, and training in scholarly writing and publishing for researchers in the developing world, largely provides through MOOCs

[Springer Nature SharedIt](#): a content-sharing initiative that (ostensibly) allows full-text subscription research articles to be shared online freely and without limit in read-only form, without violating copyright

[SourceData](#): platform for making data content of published papers searchable

The winners, announced at the conference banquet, were Publons and SourceData.

Thursday, Sept. 14

Day 2 of ALPSP

Plenary 2: Rethinking Publication: The drivers, technologies, and financing behind bold evolutions

Speakers: Louise Page, Public Library of Science (PLOS)
John Inglis, Cold Spring Harbor Laboratory; BioRxiv
Liz Allen, F1000
Hannah Hope, Wellcome
Kristen Ratan, Collaborative Knowledge Foundation (Coko)

The day's first session addressed changing modes of funding, organizing, distributing, incentivizing, and assessing research outcomes in the sciences. Page opened by noting that while professional incentives are still driven by "journal brand" rather than individual articles, there has been lots of change in the publishing ecosystem in the past 20 years, notably the emergence of Sci-Hub, preprint servers, the increased agency of funders in making publishing choices and shaping career paths, and Open Science models. Inglis described BioRxiv, a large and growing archive of bioscience preprints. He defined a "preprint" as a "complete manuscript that has not been 'certified' via peer review," and claimed that the preprint server is not meant to replace published scholarship but to complement it, noting that it may help journals "take their time" with the peer review process by making the research data immediately available to other researchers to use in the meantime. Allen urged the audience to "rethink what a research output is" beyond the traditional article, since from funders' perspectives articles are "just one part of the funded work" and they would like to see more of the work shared. Some possibilities for moving beyond the article model include "open research" outputs (post-publication peer review; open peer preview) as well as "rethinking 'authorship.'" Hope addressed the question of whether funders should be their own publishers by describing Wellcome's focus on "outputs rather than articles," taking a stance against "the separation of data from published work, the time it takes to publish, etc." Finally, Ratan described Coko's movement to "break down platform silos" by providing interoperable open source tools "like Lego blocks for platforms" for users to reinvent platforms to their own specifications.

A consistent and mounting thread throughout this session was the implication that “traditional publishers” and the models on which they operate (anonymous peer review, article format, “journal brand”) were obstacles to the fast and free flow of information and the realization of funders’ goals (presumed to be the same as researchers’ goals), namely the unmediated distribution of, in Hannah Hope’s words, “*all* research outcomes.” As an editor, I found this thread of the discussion fairly alarming and, based on feedback from the audience and the Twitter backchannel, it seems I was not alone.

 **Pablo Palmeiro** @PabloAPNL 

Shocked to see this...😱😱 at #ALPSP17
"Publishing is a dumbing down of research to fit a format"
It's too bold of an argument



 **Clare Hooper** @ClareHooperLUP Sep 14
Hmmm.... #ALPSP17

Betsy Donohue @BetsyDonohue
Ratan: the reality is that we (publishing) are "dumbing down" what research is happening #alpsp17

👤 🔄 ❤️ 1

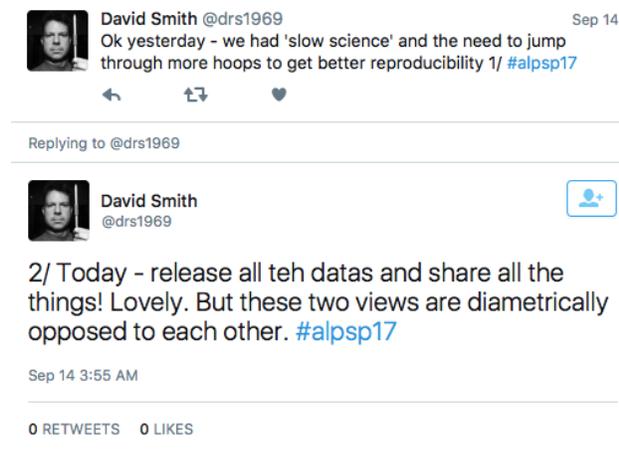
 **Joanna J Bryson** @j2bryson

I protested calling the time spent on top-drawer pubs “waste” — currated communication helps everyone, cf. [@twitter](#)

Dave Flanagan @DWFlanagan
[@j2bryson](#) speaks passionately on the benefit she derived from high-quality peer review *and* editorial on her recent Science paper #alpsp17

Sep 14 5:28 AM

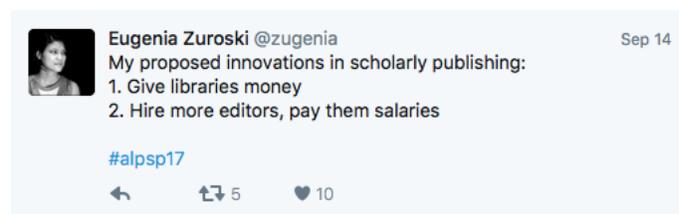
As David Smith pointed out on Twitter, day 1's opening keynote talk emphasizing research and publishing practices toward "scientific integrity" seemed diametrically opposed to the kinds of "evolutions" of distribution being championed in this plenary session:



Later in the conference, Johanna Bryson expressed a similar concern:



Altogether, it was a stimulating conversation that inspired me to think more about the values carried in traditional publishing practices ranging from peer review to copyediting that we may not want to lose in the move to a wider range of distribution platforms and the increased agency of funders in determining the shape, content, and reach of research outputs. Contrary to the panelists' suggestion that cutting out the "publishing" step and freely distributing anything and everything that results from research projects is a way to, in Hannah Hope's words, "empower researchers," I would argue that high quality peer review and editing are key components of transforming raw data into meaningful and communicable knowledge, and that taking the time to do them is essential to empowering the researcher by cultivating and protecting the integrity of their scholarship as it enters both academic and public circulation.



Parallel Session: Industry Updates
Speakers: Daniel Berze, Glasstree

Sam Bruinsma, RA21
Tasha Mellins-Cohen, Project Cupcake
Andrew Pitts, Publisher Solutions (PSI)

This session offered an overview of some of the new services being offered within research communities. Glasstree is an independent “publishing facilitator” (*not* a publisher, Berze emphasized) that promises to “return control to academic institutions and authors” by providing an alternative to traditional textbook publishing. Glasstree offers a lower price point for students buying books, higher author royalties, and a much faster timeframe for bringing out the published work (minutes to weeks, rather than months to years). Berze mentioned that they coordinate “review and editing” when necessary, but it was unclear how this was “facilitated” and to what standard. RA21 is a “resource access” initiative aimed at removing barriers to published content associated with IP-based authentication systems. Project Cupcake is an initiative collating APC information across publishers, presenting it in a standardized, organized manner in order to measure “quality of service provided by scholarly publishers to authors, funders, and readers.” Mellins-Cohen referred us to [the Scholarly Kitchen write-up of the project](#) for more details. Publisher Solutions is a security system designed to counteract endemic levels of hacking and theft published content from university servers; Pitts’s presentation focused on the dangers posed by Sci-Hub and similar “abusers.”

Parallel Session: New Technology/Traditional Values
Speakers: Jo Adetunji, The Conversation UK
 Heather Staines, Hypothes.is

This session focused on two sites devoted to presenting scholarly outputs to new and varied audiences and facilitating conversations and knowledge communities around them. The Conversation is a website that promotes “evidence-led voices” in public discourse by facilitating collaboration between academic experts and journalists to disseminate scholarly research and ideas to a wider public, building “trust in expertise.” They are a global network with sites based in the UK, US, Australia, Africa, France, Canada, and, most recently, Indonesia. It publishes all content under Creative Commons licensing, funded by HE institutions, which allows for international recirculation; its features include publishing companion pieces to academic publications “translating” them for a non-specialist audience; getting experts to fact-check public statements from politicians and news outlets; training academics to pitch and write ideas for public audiences; and (in the UK specifically) facilitating measurement of research impact. Hypothes.is is an open source, nonprofit annotation service, that allows users to build “layers of annotation” on published material across the web, to debate, fact-check, and correct misinformation; to facilitate conversations among experts, students, and other communities of readers; and to allow publishers to make notes on their own content for in-house purposes, among other uses. Staines spoke in particular about the pedagogical function of the service, allowing students and readers to cultivate media analysis skills by evaluating sources, adding context, searching debunking sites, and researching beyond digital sources by being linked to libraries through WorldCat. Currently grant-funded, the service is shifting to an earned income revenue stream, but remains committed to nonprofit status.

Parallel session: Insights from Teaching and Research: Maximizing the value of data

Speakers: Joe Karaganis, Open Syllabus Project
Ian Mulvaney, Sage

In this session I was particularly interested in Karaganis's Open Syllabus Project, which Adrian Stanley had mentioned to me, and which I'd discussed with Joe a bit over lunch prior to the session. The initiative collects information by datamining online course syllabi to generate a new publication metric: how often a published text is taught in university classes. There is much work out there, Karaganis told us, that is infrequently cited but frequently taught. In addition to helping publishers, institutions, and peers recognize the contribution of particular authors by generating classroom metrics, and allowing instructors to see examples of how particular subjects are being taught across the field, the project will be able to provide things like the Open Textbook Network with "demand data," and will allow us to analyze trends such as gender distribution among assigned texts by discipline or field. Mulvaney, addressing how "data must be cleaned and processed to maximize its value," all of which is performed by software, presented a software compression model based on "containers" as opposed to the currently predominant "virtual machine approach," where the containers use significantly less space and energy, and can solve problems of software dependencies and allowing live coding to be delivered over the web.

Parallel Session: Artificial Intelligence: two publishing case studies
Speakers: David Smith, IET
Marcel Karnstedt-Hulpus, Springer Nature

Smith walked through the narrative of IET's development of INSPEC, a database covering content in engineering, computing, and physics going back 40 years. Along the way, they faced the problem of keeping up via their manual system with all the content that needed to be indexed, and they invented an AI to increase efficiency. Their goal, Smith said, was to save money without sacrificing the quality of human work by "moving the human effort further up the value chain"—extending human capabilities rather than replacing them. Through a mix of heuristics, language processors, and other reading functions, organized through a mode of directed learning, the AI learned to read text—and learned to do it very well, as well (perhaps) as human readers. Karnstedt-Hulpus described Springer Nature's initiative to move beyond "2-D search engines" that display linear lists of results based on keywords, into a new realm of "smart searching" that facilitates browsing across scientific communities through AI-enabled "community detection." The technical vocabulary is beyond my ability to reproduce, but I understood the idea to be that the technology would be intelligent enough to have a much more nuanced sense of what kind of thing a particular user was looking for by being able to interpret "browsing clusters" in terms of communities of like-minded searchers.

Plenary 3: The Disruptive Consequences of HP's Victory Against Reprobel ... and Reasons to Be Optimistic
Speaker: Sarah Faulder, PLS

Faulder provided an overview of the history of equipment levies in the EU, established in the 1980s, and made the case for them as an important source of revenue for publishers and other rightholders. A recent court decision in Germany threatens this revenue stream by having

determined that publishers are not “rightholders.” According to the decision, the publisher was not only disqualified from future levies but was ordered to pay back levies received over the past few years. As a result, distributions were frozen, significant revenues were lost, author-publisher relations were damaged, and publishers’ status as rightholders under EU law was undermined. Among the “reasons to be optimistic” was that the German President had recently released a statement that “journalists, publishers, and authors” should be paid for the copying of their work, and that the Belgian courts are poised to revisit the case, considering whether publishers should be considered rightholders not in terms of copyright but in terms of investment in product.

Friday, Sept. 15
Day 3 of ALPSP

I didn’t attend plenary 4: Maximizing the Value of Research Value and Data: Cross community innovation, but followed it (to the best of my ability, as someone situated firmly in the humanities) through the Twitter stream. The most traffic was generated in response to the talk by Marcus Munafò (Prof. of Biological Psychology, University of Bristol). [It drew on his article “A Manifesto for Reproducible Science” \(Nature, 10 Jan 2017\)](#), which makes the case for open practices as a way of improving reproducibility and improving quality in published science. This article struck me as a good companion piece to the Wednesday keynote by Lex Bouter.

Plenary 5: The Robots Are Coming!: Ethics, work, and scholarly communication in a radically different future
Speakers: Joanna Bryson, Reader in AI, University of Bath and Princeton Center for Information Technology Policy
Sebastian Huemphreys, Echobox
Volker Hirsch, Venture Partner, Emerge Education

The final session saw the conference out on an exciting note, first when the first speaker (Hirsch) failed to show until the other two speakers had presented; next when Bryson set a breakneck pace with her extremely rich and informative talk; and finally when ideological sparks flew between Hirsch and Bryson, who come at the field of AI from opposing sets of interest (the venture capitalist vs. the expert in regulatory policy). Bryson opened her talk by asserting that “Robots are not the issue. Sustainability is the issue. Inequality is the issue.” She broke down some of the key principles of the nature of intelligence that is crafted in AI, insisting that it is an artefact (something made by humans and requiring human structures of responsibility for it); that, like natural intelligence, it is not general but always operates and furthers itself through learning within strict parameters of preconditions; and that it enhances human power. In response to the fear that AI threatens human employment, she argued that jobs are threatened not by intelligent machines but by an unhealthy economy in which wealth doesn’t circulate. Huemphreys described the social media AI at Echobox, which decides what story, image, or lede to show which user at an optimal time for engagement and resharing. His account was consistent with Bryson’s description (as well as David Smith’s on day 2) of AI’s qualities and strengths, enhancing human capabilities, and freeing human agents up to focus on “big issues” over automated work. When he finally arrived, Hirsch delivered a TEDtalk-style presentation that Bryson explicitly objected to as alarmist and unscientific (using images of robots with red eyes from the *Terminator* movies, for example). Describing how the value of networks is defined by

the number of nodes they contain, he defined an intelligent machine as a “data node.” Responding to the previous panelists’ suggestion that AIs do not pose a significant threat to human workers, he noted that while they won’t replace high-end employees, they can and will replace the lowest-paid, least specialized workers, as well as certain mid-range skilled workers (he used the example of the law firm Clifford Chance replacing its junior associates with AI). Much of the discussion during Q&A focused on questions of ethics, with Joanna insisting that “we need regulatory frameworks to hold corporations accountable for their AIs,” and cautioning that “data is the new oil, and oil explodes: data is an asset, but also a liability.” While AI may be, as Huempher suggests, a solution to managing the massive quantities of data that, as Hirsch argued, define and will continue to define our world, Bryson urges us not just to “keep up” with data but to responsibly manage it (including, vitally, not circulating or even keeping much of it). Finally, while her efforts to turn the conversation to the implications for publishing specifically were not very successful, Bryson did offer one piece of valuable advice to publishers:

← Replying to @zugenia and 1 other



Eugenia Zuroski @zugenia

Sep 15

More advice from @j2bryson, to scholarly publishers: Don't pander to audiences. Publish with integrity! #alpsp2017



1



4

More from the Twitter coverage:



Jo Wixon @JoWixon

Sep 15

Fantastic enlightening talk by @j2bryson showing AI is not a threat to us, rather we need to look at our own behaviour/decisions #ALPSP17



1



4



Ashley Farley @ashleydfarley

Sep 15

Fantastic definition of AI. An artefact means we own the responsibility of care & feeding #alpsp17



4



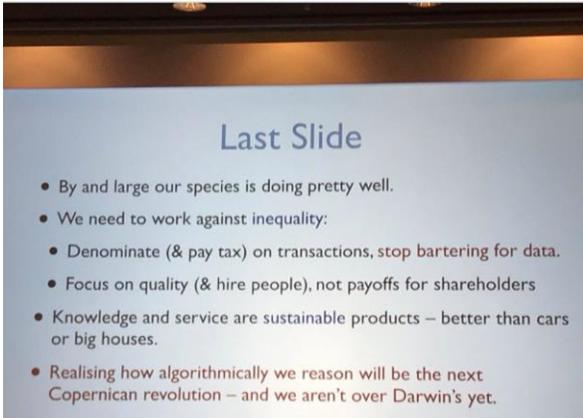
3



Jasmin Lange
@JasminLange



To conclude. @j2bryson #alpsp17



Sep 15 5:20 AM

4 RETWEETS 5 LIKES

Eugenia Zuroski Retweeted



Kent Anderson @kanderson

Sep 15

My new hero when it comes to economic policy in the Age of Algorithms is @j2bryson. Hear her, and heed her words. #alpsp2017

4 8



Adrian Stanley @AdrianStanley13

Sep 15

The lively Volker Hirsch @vhirsch giving some great pictorial data and AI examples, glowing eyes and all #Alpsp17



1



Adrian Stanley @AdrianStanley13

Sep 15

Passionate panelists, great way to wrap up #Alpsp17 great job staging and planning this final session @IsabelT5000 and @drs1969



2 3

Conclusions:

This was a very exciting and full conference. In addition to the sessions detailed above, I had the opportunity to spend time and have longer conversations with a number of people, including (in addition to Rowly, Suzanne, and Adrian already mentioned), Rose Nottebaum (Aries Systems), Satyajit Rout and Pablo Palmeiro (Editage), Heather Staines (Hypothes.is), Kristen Ratan (Coko Foundation), Joe Karaganis (Open Syllabus Project), and Joanna Bryson (Bath/Princeton), and I made some ongoing connections with other delegates on Twitter.

My sense is that everyone understands that the publishing ecosystem is not only vastly different than it was only a decade or two ago but also changing rapidly, which requires a certain level of attention to and engagement with how knowledge is being produced and circulated, and a willingness for various kinds of stakeholders to collaborate on forward-thinking approaches. I found certain initiatives particularly promising in terms of knowledge-sharing and community-building, such as the Open Syllabus Project and Hypothes.is, and I also see promise in the way initiatives like Open Syllabus Project and Publons are organizing and generating metrics for conventionally “invisible” but essential academic labour like course design and peer review.

I continue to hold significant reservations about the industry’s uncritical embrace of values like speed, unmediated access to everything, and “innovations” that rely on transferring the burden of assessment, design, communication, etc. to “users” rather than professional specialists. For example, Coko’s modular, open source arrangement (like OJS) may offer users the flexibility to make the perfect platform for their purposes, but figuring out what kind of platform that is, building it, tweaking it, and maintaining it is all work that becomes absorbed into the job of publishing. The same could be said of the kind of “fact-checking” and expertise communication enabled by a site like The Conversation—on the one hand, better integrating scholarly voices into public discourses seems vital to reestablishing trust in learning and expertise; on the other hand, public relations and scholarly research are two full-time jobs that draw on distinct sets of skills and energy; neither one happens effortlessly or affords much “spare time.” Having the platform and an invitation to participate is not enough for a sustainable conversation. I did not see any explicit discussion of how these new kinds of work that emerge in the course of the “evolution” of publishing are going to be organized, accommodated, and compensated (except for the suggestion that maybe AIs would become largely responsible for them).

I also think that publishers, researchers, professional societies, and academic institutions should continue to insist on working relationships with funding agencies, relationships in which we are able to articulate and assert the conditions of ethical and rigorous knowledge production as we understand them, lest those agencies completely eclipse our own agency in the discourse on what knowledge is and is for. The discrepancy between the opening keynote’s call for scientific integrity through more vigilant and careful research methods, and the funding foundations’ call for quicker, less mediated access to the research data they’ve invested in, signals a profound failure of communication on the value of the contribution editors and publishers make to what we might call “quality control” in the information industry. Part of the reason initiatives like The Conversation and Hypothes.is have emerged to provide tools for correcting public misinformation is that public discourse, having been so thoroughly “decoupled” from scholarly expertise in the current media landscape, is flowing fast and freely without reliable mechanisms

in place to qualify and adjudicate it. We have the ability to address these problems further up the production chain, but that ability is compromised when financing structures are set up to work against rather than with publishing experts.

Finally, as an academic and editor situated in the humanities, I did feel like something of a lone wolf—not in an unpleasant way!—at a conference (like the publishing industry) heavily oriented toward the sciences. It’s not immediately clear to people at a gathering like this what someone with an expertise in eighteenth-century literary and cultural studies might bring to the table. But, at least in England, the eighteenth-century was a period largely defined, like our own moment, by radical shifts in the media environment, namely the institution of modern publishing—which is to say, we eighteenth-centuryists are trained to take a broad historical view of complex cultural phenomena just like what ALPSP-goers kept calling “this moment.” Humanities scholars in general are trained to think about what kinds of relationships are possible at particular cultural moments; how innovations both material and conceptual generate unexpected outcomes both good and bad, sometimes immediately, sometimes at a “slow burn” only palpable in retrospect; and how what counts as meaningful knowledge changes according to how we organize information and exercise our social values. So, at a time when humanities fields are being decimated as an effect of predominant financing structures that do not recognize the broader value of their contributions, I’d like to put in a word for the importance of having people with these forms of expertise interacting with other stakeholders at all levels of the conversation on how we generate and share knowledge. I think we have a lot to offer!



Matthew Gabriele 
@prof_gabriele

Literally, the Humanities.

Neil deGrasse Tyson @neiltyson

In school, rarely do we learn how data become facts, how facts become knowledge, and how knowledge becomes wisdom.

Sep 5 9:57 AM

9,343 RETWEETS 26K LIKES

Submitted by Eugenia Zuroski, McMaster University, *Eighteenth-Century Fiction*