

Let's Talk: Bringing Many Threads Together to Weave the Scholarly Information Ecosystem

Laurie Goodman
GigaScience

Howard Ratner
CHORUS

Greg Tananbaum
ScholarNext Consulting

John Vaughn
Association of American Universities

T. Scott Plutchak
University of Alabama Birmingham

Follow this and additional works at: <https://docs.lib.purdue.edu/charleston>



Part of the [Library and Information Science Commons](#)

An indexed, print copy of the Proceedings is also available for purchase at:

<http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

Laurie Goodman, Howard Ratner, Greg Tananbaum, John Vaughn, and T. Scott Plutchak, "Let's Talk: Bringing Many Threads Together to Weave the Scholarly Information Ecosystem" (2014). *Proceedings of the Charleston Library Conference*.
<http://dx.doi.org/10.5703/1288284315613>

Let's Talk: Bringing Many Threads Together to Weave the Scholarly Information Ecosystem

Laurie Goodman, Editor-in-Chief, GigaScience

Howard Ratner, Executive Director, CHORUS

Greg Tananbaum, Owner, ScholarNext Consulting

John Vaughn, Executive Vice President, Association of American Universities

Moderated by T. Scott Plutchak, Director, Lister Hill Library of Health Sciences, University of Alabama at Birmingham

The following is a lightly edited transcript of a live presentation at the 2014 Charleston Conference. Slides and videos are available at <http://2014charlestonconference.sched.org/>

T. Scott Plutchak: Good morning. I'm always impressed with the intrepid nature of the Charleston conferees coming into these crowded rooms morning after morning. I'm Scott Plutchak, from the University of Alabama at Birmingham, and what I would like to do is welcome you to what I hope will be an interesting and engaging conversation for the next 45 minutes about some of the issues involved in creating the new scholarly communications ecosystem that we are all involved in.

I want to set the stage with what is commonly called the "Holdren Memo" released by the White House Office of Science and Technology Policy back in February 2013, which directed all of the federal funding agencies to develop policies and plans making the results of federally funded research publicly available. Now the memo addressed both peer-reviewed publications and data and it inspired many in the scholarly communication world to greater efforts and greater activity to try to address these challenges, which many people had been working on for a long time. Now we haven't heard a lot from the agencies since then but there's been an awful lot of activity across the systems since then. Much of which goes far beyond what was called for in the OSTP memo.

What we've convened here is a group of people who have a significant background and interest in these. What I hope we can do over the next 45 minutes is have some stimulating conversation,

bring many of you into that discussion as well. Now within a couple of months after the Holdren Memo came out, the SHARE and CHORUS initiatives were first announced. The two projects were developed independently and many people in the community saw them as being indeed competitive and there is some concern about who's going to win. But what we've seen over the succeeding months is that the people involved in those have really started to see a lot of opportunity for collaboration, a lot of shared interests, and there's an increasing amount of work being done together. And so to talk some about that, we have Howard Ratner who is the Executive Director for CHORUS, Greg Tananbaum who is working as a consultant to help shepherd the development of SHARE. Now I'm going to assume that most of you are basically familiar with the two proposals but I did ask Howard and Greg to give me just a little snapshot of where things currently are. And so, Greg describes SHARE as a "higher education research community initiative to ensure the preservation of, access to, and re-use of resource outputs. SHARE aims to develop workflow policy and infrastructure solutions that capitalize on the compelling interests shared by researchers, libraries, universities, funding agencies, and other key stakeholders to maximize research impact today and in the future. SHARE aims to make the inventory of research assets more discoverable and more accessible and to enable the research community to build upon these assets in creative and productive ways." And it's a joint initiative of the Association of Research Libraries, the Association of American Universities, and the Association of Public and Land-Grant Universities with funding from IMLS and Sloan.

Howard talks about the focus of CHORUS being "to efficiently advance public access to content, reporting on funded research and associated data. Services are available at no cost to funders, researchers, academic institutions and libraries, and the public. It's built on widely used technology thereby simplifying compliance, minimizing implementation costs, and enabling interoperability, text and data mining, and dashboard monitoring. It uses distributed access approach points to the accepted author manuscript or version of record in context on the publication site backed by a trusted archive. Systems currently in production with a commitment from the Department of Energy—they're in discussions with other U.S. and global funders. Project is managed by CHORUS as a 501(c)(3) membership organization with publisher members, affiliate members, funder partners, and academic supporters. It's growing and evolving through a membership drive, collaboration with SHARE, and exploration of connection to data repositories."

Much of the attention in the library and publishing communities has been on publications. The Holdren Memo lays equal stress on data. And it can be argued that the public benefit of access to data will be even greater than the public access to publications. There's been a flurry of activity in research institutions around the challenge of effectively managing research data. And on the publishing side, people really trying to figure out how do we connect publications to the data on which articles are based? To help us sort through some of those issues, we're joined by Laurie Goodman who's the Editor-in-Chief of the journal *GigaScience*. As Laurie pointed out in an email last week, there are basically five primary needs for data availability: release, accessibility, curation, tools for data manipulation, and permanent community-approved databases for all types of reusable data. Each of these has different levels of difficulty and different personnel and financial solutions and so we'll try to weave some of those issues into our discussion as well.

Finally, I want to welcome John Vaughn, whose experience on these issues is broad and deep. He's currently a senior fellow with the Association

of American Universities. He was, for many years, the Executive Vice President of that organization, worked on scholarly communication, intellectual property issues, has many interesting scars from the political wars involved in working on that association in DC for many years. I first got to know John when he was the Chair of the Scholarly Publishing Roundtable several years ago. The recommendations from the roundtable, many of which were incorporated into the America COMPETES Act and many of which really informed the Holdren Memo and those requirements. So we're very glad to have John here. I can attest to his ability to walk the tightrope of balancing the needs of competing communities and figuring out how to bring them together to work on common solutions. I want to start with John, and start by asking him, from his standpoint to someone who's worked with the higher education community for a long time, what you see as some of the key issues and points that those of us who are involved in developing systems like SHARE, CHORUS, data issues really need to be paying attention to meet the needs of those constituencies that presumably we are all trying to support?

John Vaughn: Well I, I think, as Scott has said, there's been a long running effort that I've been involved in on a series of versions to try to get these various, and I think it in the past has been appropriate to call them various warring factions, together because within the higher education community, you have multiple perspectives. The broad mission of higher education and of particularly of research universities is the discovery and dissemination of new knowledge, and we have a mission with students, with faculty . . . And I think in many respects, this scholarly communication system has been and is working very well. We produce high quality research. Publishers do a terrific job with peer review. There's broad dissemination. There have been surveys indicating that many scholars think that they have more access to information now than they ever have. But if you look at it from a different perspective, say research libraries, there's been an explosion of cost, and sort of an undercurrent of what's the source of these increased costs, which have really been quite

dramatic. There is some evidence in the past, and I think it is largely in the past, not totally, of exploitative pricing policies by publishers, commercial publishers making large profits, not-for-profit society publishers charging prices to generate revenue to run their societies. And that, many provosts see it as, well, tantamount to saying research libraries should subsume their responsibility of funding academic society. Commercial publishers have done a terrific job of providing high-quality publications and using the revenue in a variety of other ways that are very helpful to the community. The societies are extremely important to the university system, but essentially, some of the pricing policies have been saying, "Universities, you give us commercial publishers or societies some revenue that we will generate from the journals you buy and we'll decide how to use that money." And the decisions often have been quite good, but that has been a source of tension. The digital revolution has been seen by many of us as providing a way to increase access, reduce cost, and maintain quality. But as we work through this, and I should say that the increased cost to libraries and universities probably predominantly represents an increase in the explosion of research around the world. Me, I was astounded when we were working with the American Physical Society a few years ago that—and I think it's still the same—two-thirds of the authors of APS journals are outside this country. China, European universities, their governments are pouring money into them. They have seen the advantage to the economy that US universities have provided the increase and innovative capacity. So there's been an explosion in research. That's a good thing! But trying to figure out both in terms of cost and volume of this explosion of information how to manage it is really daunting and I think that's where we see, I think, less a way of reducing costs and more a way of expanding access using digital technologies. What becomes clear, and I think we will hear more today, is no group can do that alone. Universities can't do this alone. Societies can't. Publishers can't. We need a collaboration of publishers, universities, and their libraries, and even government. I say even government because I've over the years watched all too often good government intentions becoming bureaucratic ossification. But this OSTP

memo is just terrific and I think that is launching a really good collaboration.

T. Scott Plutchak: I think it has, and I want to pull in Howard and Greg here. There was a lot of work being done, but it, and I know from the publishing community really looking at where does open access fit, understanding the push towards from government mandate, certainly within the university and library communities trying to figure out how to really change the ecosystem, the OSTP memo really sort of lights a fire under everybody. And we start these two independent activities, which have gotten quite broad. Can you talk a little bit about your experience with those and when you started to see those guys are also working on the same problem in a different way and maybe we don't agree with all issues but we really need to start talking to them.

Greg Tananbaum: Sure. I can, I can take first crack at it. Certainly, informed by what OSTP produced but not necessarily a direct result, this notion of SHARE was bandied about in early 2013 and, and John has been an integral part of that since the beginning. And initially, we talked about, in a working graph that was meant to be provocative and evoke feedback, and get the community involved. We talked about the notion of a network of federated repositories to house this public access information that was going to be under the OSTP umbrella. And in talking to a wide range of stakeholders, not just librarians but provosts and vice chancellors to research, and individual professors and publishers, and the funding agencies, and nongovernmental funding bodies, it became apparent to us that discovery is important. Discovery's very important. But there are means to discover content. There are efficient paths to discover content. What was more pressing was understanding, getting a handle on who was writing what and on whose dime? And understanding that in a timely and comprehensive and structured fashion. So all the stakeholders that I just mentioned, in some capacity, want to know that. They want to know in a, again, comprehensive way who is doing what and who's paying for that?

T. Scott Plutchak: And, I don't think many of us, certainly in the library community, realize how big a problem that was.

Greg Tananbaum: Yeah.

T. Scott Plutchak: I think we just assumed the funders are funding stuff they know what's being published and . . .

Howard Ratner: No.

T. Scott Plutchak: . . . that's not it at all!

Howard Ratner: Yeah, no, it's not that way at all.

Greg Tananbaum: And just to conclude this point and I'll turn it over to Howard, the notion that we have all these systems . . . We live in an amazing age, right? There are all these systems. We have institutional repositories. We have CHORUS. We have grant management systems. We have editorial management systems. We have personal productivity tools. But there's this real potential that these become towers of Babel—that they don't talk to one another. So, to the extent that there's an opportunity and a pressing need here, I know the way that SHARE have used it is, if we can apply some rationality to the ecosystem, that will be a service that will be very, very valuable to all of the players.

Howard Ratner: So for me, spending about 30 years in publishing, built a career on interoperation, right? So when I was approached with this problem—and but I wasn't there at the very start, of course but certainly got involved very quickly—I saw . . . There's so much need for interoperation—along the same lines that Greg was just talking about and there are all of these identifiers that I helped create or helped cajole to move along from CrossRef's DOI to ORCID's to other all the things and now, including some of the work that's been going on with FundRef. But they're not being used and they're not necessarily being used as widespread. They're being used in pockets. What CHORUS is trying to do is really try to pull that all together and make all of these different things interoperate and most importantly, make it itself publicly accessible vis-à-vis transparency. Everything about CHORUS is about transparency. Our search is completely

open. Our dashboard's completely open. And so when we started talking to SHARE, I said, "You know, what's the problem here?" Really, you have some very unique pieces of data that publishers don't know about, "research events," as you call them. I think that's a great word. Publishers obviously know about publication events and therefore, they could feed into the system. Everyone wants to understand the data. And the thing that's missing, really, especially in the data space, is good comprehensive data which supplies context. Because it's great, you can hand me your Excel spreadsheet. That's great. It's a bunch of numbers. I have no idea what context it was developed in. I have no idea who you are. Right? I have no idea what experiments were done with it. And so, context around that, again, metadata around the data is key. And so, you're absolutely right that when we started talking to the funding agencies . . . And this got actually started with my ORCID work first. It was so clear to me that they had no idea what happened to the research after they gave the grant. And this is the genesis of ORCID, quite honestly. It's because we said, "This is a huge problem." And first, we thought it was a publishing problem. But then, we opened it up and we called a summit and we had lots of people come from the university, we had lots of people come from the government and we said, "Wow! This actually isn't a publishing problem." The publishing problem is maybe 10 or 15% of the problem.

T. Scott Plutchak: Yeah.

Howard Ratner: It's a huge problem outside of it. And that's why, ORCID in particular, became such a collaborative effort amongst all the different stakeholders and CHORUS is doing the same. But the biggest difference, CHORUS is also about efficient workflows. There's already existing workflow that can channel most of this. That funnel already exists. The researcher pays attention to a lot of metadata when they go to submit a manuscript. So why not reuse that? Why not maximize that? Why not leverage that so you could reduce the burden on the researcher and then, up compliance?

T. Scott Plutchak: Okay. When you talk about efficient workflows and I think about my life in academia . . . Academia's not known for efficient workflows. So again, it is something that is desperately needed. And then of course, you talk about data and as I said, I think, many of us looked at the publications because we've been so focused on open access as the publications. But you start to get in the data, and if you look at the work that Laurie's been doing at the journal, *GigaScience*, it really makes open access to publications look easy. And data is just a mess. The context issue, for example, that the physicist in the previous session said, "Oh, well, you know, we don't need librarians and we don't really need to make our data open and nobody would understand it anyway because it's all just ones and zeroes." Um . . . how do we start to address that practical issue but then also that mindset?

Laurie Goodman: Yeah. There's actually a number of things when you talk about mindset that is where people are like, "Nobody can understand my work." And I'm like, "The whole point of your work is for people to be able to use it and understand it." And in the publishing field, I've certainly had people go, "People have to stand at my side at the bench in order to understand this." And I've said, "Well, that's great for you to put out an advertisement that your lab is doing nice stuff but you have to explain this for people to use it." But data curation, being able to organize the data so that people who are not collecting this, is a major, major issue! And when you talk about metadata, that is the . . . When I look at your zeroes and ones, this metadata tells me what each zero means and what each one means. And librarians are curators of information. And that's really what data is. We were talking last night about how, you know, putting money aside and all that, I'm a strong believer in if it's important, we can figure out a way to fund it. There might be arguments but we can figure it out. Librarians are in a perfect position as already understanding the issues of organizing information and information is data that you can be a driving force for curating this information, for having initiatives within your universities to educate the researchers. You would think the researchers would understand that they have to curate it! But in fact, researchers

say as soon as I put this into the computer and I'm done with it, if you ask me three months from now or next week, I wouldn't be able to tell you what this data means. Well, if you're funding this with your taxpayer dollars, or you have people who are giving money to charities where their children are dying from medical diseases, do you want all of that work just put away in, into a computer and . . . ?

T. Scott Plutchak: As you say, they're focused on their particular thing.

Laurie Goodman: Yeah.

T. Scott Plutchak: I would say, by the way, we'd be happy to, any of you who would also like to join the conversation, come up to the microphone and I'll try to recognize you. You mentioned the money piece and many people who've looked at this said there's enough money in the system to do what we need to do but the challenge is of moving it. You've dealt with universities and moving money. How realistic is it to think that we're going to be able to persuade the people who run the funds in universities to make the kinds of decisions that need to be done to actually address some of these challenges?

John Vaughn: Well, part of the job of universities is to enable bright, energetic, creative researchers to do their work driven by their understanding of the discipline, investigate or negotiate their work, but there has to be . . . There's a point at which you want to enable researchers to follow their own logic but there's a responsibility back, most university research is publicly funded. We have got to be able to make a connection between this work and an eventual advance of knowledge. It may be years away, but unless we can connect what we are doing to societal benefit, we're going to . . . I think most of you are aware of the battles that are going on in Congress now: the assault on NSF and peer review and social science. But universities are really struggling with finding the resources. Research funding in this country has been flat for some time. We've been talking about ways of moving toward open access. That gets you into author pays. Where does that money come from? I think there are funding challenges and that is where I go back to this notion that we've

got to have collaboration among the key players. I think the way CHORUS and SHARE are working together is exceptionally important. The government is now working, effectively launched by OSTP. But, there's a public responsibility that we collectively have to work on.

T. Scott Plutchak: Yeah. To what extent in the SHARE and CHORUS discussions do you all get into those human factors thing? I know you're focused on the systems and the workflows but we were talking last night about some of the data challenges, and so much of it is getting to change the way that people think about this stuff.

Greg Tananbaum: I think one of the things that Howard and CHORUS have done well, and he alluded to it, is their capturing information at a point where it's in the author's interest to provide it. Right? The author's never more captive or the researcher's never more captive than when they're submitting a paper for potential publication. You know! Just by a show of hands, how many people here work in an institution that has a repository? An institutional repository? Okay, keep your hands up. Now, of you, how many of you have had success in getting authors submit directly to your institutional repository? [laughter] So when we talk about data, uh, for example, this is why, as a fundamental principle, CHORUS knows or SHARE knows or we do together as part of a larger set of activities, capturing that information as organically in the workflow as possible is critical. Because as you said, the moment they're done with it, they're done with it. And trying to get them to retroactively, the researcher, to retroactively go back and put out this information, it's simply not going to happen. So that's a challenge and it's a challenge that we collectively have to face.

Howard Ratner: But there's also capturing and using the data. I mean, publishers have been capturing data and institutions have been capturing data for years. It's not until you start to use it that you see the value it or the inaccuracy of it. So one of the interesting things is because of CHORUS, we're actually pulling that data all the through CrossRef and all the way through our workflow and actually showing it up in our search interface, people are saying, "That's not right!"

And why isn't the Department of Energy showing up there? We know that that's about physics—why isn't it there? And so, the publishers that are part of my group have really been tweaking those interfaces and working with FundRef. And FundRef actually, when it started, had all of—what?—about 2,000 entries? Okay? So this is identifying all the various different funding agencies and departments around the world. It's now up over 8,500 terms worldwide! And again, why is this happening? Because it's being used! If it's not going to be used, it's just going to be dormant and then, no one really knows if it's good. And by the way, that's the other thing that we're doing. It's like we worked with Portico. And for the first time ever, we're saying, "Okay. You know, the publisher says that that DOI or that work is actually archived at Portico." But we actually asked the question. We say, "Here's a bunch of DOIs that we're actually monitoring. Do you actually have this?" And we're making that public to all of you.

T. Scott Plutchak: A question from the floor.

Gail Clement: Hi. I'm not sure it's a question. It may be a comment. And it's also a little plug for a lively lunch presentation later today. I'm Gail Clement. I'm at Texas A&M University in the Schol-Comm office and I work directly, my colleagues and I are the ones that would actually be working at worm's eye view. Where y'all are at about either space shuttle view or . . . you're very high level. And I think when we start to wrestle with how we implement stuff to support our campus authors to be aware and ultimately to comply with policies, mandates, and practices, you know, what I feel is there's this ginormous missing link. That's what we're going to talk about in the Empowering Data (lively lunch session). By the numbers, at the end of the day, the average researcher on campus who knows the most about that data—because they're the ones collecting it, curating it, and best qualified to document it in a rich way so it could be made reusable by others—are the graduate students and the postdocs. You know, I'm going to share this in our lively lunch, but there's 29,000 faculty out there in the US. There's 1.7 million graduate students and there's about 90,000 postdocs. So the scale of not closing

this gap . . . And when we start looking at practices around how we treat graduates—I'm going to, specifically with graduate scholarship is what we're going to look at later today but . . . The point is they're being mandated to put their scholarship behind a pay wall. They're being mandated to roll everything up into big, fat, chunky PDFs because that's the most operationally manageable way for most of our campus repositories, whether it's bepress or dSPACE. If you actually look at the problem space for trying to get our campus authors to meet you somewhere in the middle and be able to meet your high level objectives, we've got to close that gap. And I'm not sure our campuses can handle new formats. And it's . . .

T. Scott Plutchak: Yeah. Well, I think some of that is Howard's emphasis on workflows and existing stuff to try . . . And I think SHARE at the same time also looking at how do we leverage those existing things on the ground level? And figure out how to link 'em up? But it's, it's huge and I see Laurie's nodding as you're . . .

Laurie Goodman: I must answer! You know, I see problems as sort of an exciting moment to how can we answer this? I love problems. And one of the things . . .

T. Scott Plutchak: You're in right field.

Laurie Goodman: Yes! One of the fields . . . Well, all fields have problems. One of the things that you're talking about is the grad students are the ones who know what they're collecting. Now, they know what they're collecting but they don't know how to write it down. There are no tools. None! Available for people to properly curate. What if you're out in the field? In the jungle collecting stuff? You don't have a, you know, a PDF maker out there and PDFs aren't searchable. Librarians interacting directly—don't wait for it to come to you! Create initiatives at your university where librarians are engaged with the researchers where you understand what is going on. You don't need the details of exactly how this experiment is going to work but you need to be able to engage in a conversation at the time that they are developing these experiments. Don't wait until they're publishing. I have a bio-curator with my journal

because by the time it comes to us, that data is in a disastrous format and we have someone who walks them back through it. You guys are in that place! You could do it at the beginning.

Gail Clement: Well, can I just do a Part B follow up?

T. Scott Plutchak: Yes.

Gail Clement: Because I really appreciate what you're saying, and I think many of us here are right there with you. The concern and the challenge we face is that in the case of graduate stuff, we know what we would like to see happen. But the problem is that when it comes graduate scholarship in particular, and I would argue probably also postdocs, 1) these are transitory linkages to the data set. They are not the persistent curators over time. So that wouldn't be enough if we don't catch the PI, their faculty adviser, because they have the long term. Postdocs, we now know that "Scholarly Kitchen" Phill Jones piece on the state of postdocs and in some cases, there's more po-, you know. In chemistry, most research is being done by a PI as a postdoc, than as faculty. These are people that don't necessarily have a long-term relationship so we won't get there. The other thing is that administrators are driving the bus. We know what we would like to reach out to them, and many of us are very engaged in outreach and learning, and teaching and learning efforts. But in many of our institutions, those early career people have minimum agency over their choices. They are out of power. And we can talk to them about what it would be nice to do, but I've been on a task force at A&M for now two years just to hash out the IP rights—not only in the dissertation but in all that constellation of research output, you know, the underlying data, and authorship epics. So, until our institutions honor the agency of our early career researchers that are doing the most heavy lifting around creating research, I'm not sure how effective a librarian, who may be also a faculty member in many of our cases, we can go toe to toe as faculty members, but as long as there's no agency with the users . . .

Greg Tananbaum: And some of this, some of this is certainly . . . I mentioned at the outset, workflow infrastructure and policy. And on the

policy perspective, if you look at PubMed Central for example. When NIH stepped up enforcement and made it clear that if you didn't comply, you were at risk at not getting future funding, suddenly that became, the numbers jumped up. So there is a policy component. This isn't front and center of what SHARE is doing at the moment but certainly, we talk to lots of folks at the university administration level and at the funding agency level about the impact of existing and potential policies in attracting compliance.

T. Scott Plutchak: Yeah, you know, we're talking here about the need to get these larger communities working collaboratively, but the problem on individual university campuses is huge because they tend to be so siloed. As you were talking, I was thinking, two threads—one is so much of the work that Howard has done even prior to CHORUS is developing those kinds of linkages that can create that persistence through CrossRef, through ORCID. But then also looking at what John has done from the association level is the associations of universities have got to put pressure on their members to recognize that these needs require different interactions on our campuses and there has to be room to develop those extra kinds of agencies that you talk about.

Fred Dylla: Fred Dylla, American Institute of Physics. I want to pick up on Greg's last point. The beauty of the OSTP memorandum was its highly nuanced language that allowed many solutions to boom. The danger of the OSTP memorandum is the same language. And I think the nightmare scenario we all have—whether you're an university, university administrator, a librarian, a publisher—is we'll have 41 different mandates from various federal agencies. I think one possible way—and I'd like to hear your comments on this—to guard against this chaos is for our colleagues in the federal agencies to see the universities and the publishers and the data management community working together to establish a rational way as to move forward. What are your comments on the ability . . .

Greg Tananbaum: I would say, Fred, it's difficult to coax and cajole graduate students and post-docs to do what you want them to do. I don't think it's much easier to coax and cajole agencies

to do what we want them to do. I wish it were, but I think there's a sense—and you can weigh in here, obviously, too—the agencies are going to do what they're going to do. Universities have tried to inform the process by which they've set forth their answers to the Holdren Memo, and publishers have tried to weigh in on there as well. But the publishers are going to do what they're going to do. And of course, that's a, potentially a tremendous burden for many people in this room. I mean, you think about, as you said, from an institutional perspective, if the vice provost for research office, president for research office has to make sure, as a condition of funding, that these mandates have been fulfilled or have been complied with, and there are 41 different ones, and people get funding from multiple sources, the coordination problem is a mess! It's absolutely a mess. So I agree with you. To get out in front of it is important. But ultimately, I'm not sure how much influence we potentially have to move their policies.

John Vaughn: Let me add something there. There was one line that John Holdren put in his memo that to the extent feasible, agencies should try to coordinate their policies. They're aware of this. And I know a goal of SHARE, and I think of CHORUS, is to first . . . There are a lot of behind the scenes conversations with agencies to try to minimize the nightmare you talk about, Fred. And I think there'll be some success there. But the other responsibility we have is to try to protect our researchers from having directly to spend time responding to all these different compliance rules. There was a study recently that indicated that of the faculty members' time in research, 40% of it is spent in administrative compliance. You could argue that that's a generous percentage that is being taken away from created work into bureaucratic work. And I think part of the job of CHORUS and SHARE is to try to minimize that. So we've got to have multiple conversations. There will necessarily be agency differences, sometimes driven by disciplinary differences. But I do think we all need to focus on trying to avoid that kind of nightmare disparity and I think we will have some progress on that.

Fred Dylla: I think that the agencies have admirers. The fact that FundRef started, ORCID started, DataCite started; you ask the typical graduate students down in the trench, they think ORCID's a flower. They never heard of CrossRef. [laughter] So these have to be, as both Greg and Howard have mentioned, behind the scenes thing that take care of that magic moment when a data set is being sent in to DataCite to get an identifier or the manuscript's being sent in. We need to do more of that.

Greg Tananbaum: And to that point, I'll just give a plug for you here. If you are going to walk away with one action item in the audience, evangelize ORCID on your campus. It makes life easier. It makes life easier for publishers, for funding agencies, for institutions. We need to be able to know who is doing what.

Howard Ratner: Right. The . . . But, but ORCID is one thing. And that's definitely a grease in the wheels and that's great. We want more grease in the wheels. We want to make things move along. But actually, with what we're talking about today, much more important than ORCID actually is FundRef. Because if we can't identify the content, actually, the whole system breaks down—and that includes everything along the way. So we need to be able to identify this content. So I would say . . . Yeah, ORCID's great. Do that. Please do that.

Greg Tananbaum: And while you're talking . . .

Howard Ratner: And while you're talking to them, encourage when they submit their manuscripts to fill out those screens that are on every single manuscript tracking system now that say tell us about where you got your grant from. Okay? It's very simple! Usually it's two fields and in some experiments or some articles, it could be as many as five or what have you. But the more you fill that out, the more transparent information we have and data so we can actually make some logical, informative decisions. Because one of the things that I've experienced by talking to these agencies . . . And you're right, they're tough to influence. Right? I'm not even trying to influence them one way or the other. I'm just trying to present them with data that there is data available that they are not even aware of. They

don't know what happens on the institution campus. They have no idea of the publishing workflow. They don't know what CHORUS is. They don't know what ORCID is. Some of them do know about FundRef—the larger agencies do. But we need to really get out there and speak to them so that they understand our language and they understand what we've already built, what we've already worked together on so that they don't redevelop it.

T. Scott Plutchak: But again, part of the challenge on the ground, when you're talking about getting them to fill out the form. On an institute like mine, which is heavily invested in big center grants, any particular paper that author is not entirely sure what he or she is being funded for . . . And again that comes back to those of us on the ground in the institutions doing that educational work to try to keep that going. Question, comment here?

Marilyn Billings: Marilyn Billings from UMass-Amherst and kind of a comment and a suggestion. We've done a lot of work working with our graduate school and with our office research, but one of the comments I hear from our VC for Research is, "I don't see what librarians bring to the table. I don't see what the added benefit is. Why do librarians know about data? Or any of this kind of works?" So, comment to AAU and ARL also is that when we have our programs for our chancellors, our provosts, the research officers, it would be very useful to bring in a panel that would talk to them about all of these elements and what librarians are bringing to that table.

Laurie Goodman: I actually want to speak to that because we're at this meeting where you're all here. And I think what you need to have, absolutely, is each of you've brought up a particular problem. There should be sessions where everybody who has that problem comes together and says what can we do about it? How can we move our academic organization forward? What powers do we have and who do we have to engage? When everybody raised their hand with a database and then lowered it . . .

Howard Ratner: That's the problem.

Laurie Goodman: That's because it's hard to do. But what was interesting was how many people raised their hand for a database. All of you people together have expertise in knowing, creating a database, and not being able to get people to populate it. You're all smart. You should have a group that comes together. I mean, librarians are the ones who created the cross-sharing between libraries! You guys are already organized on getting people to share. And that expertise, I think, is lost a lot of times because it's not directly addressed when every library in the world is here.

T. Scott Plutchak: Right. We're close to end of time. John.

John Vaughn: Just, just quickly to respond to. I think that few of the UMass-Amherst administrators are somewhat idiosyncratic. AAU's now worked with ARL on a task-force for a couple of years that grew out of precisely the meeting you were talking about. We had the head of ARL and the librarian and the publisher meet with our chief academic officers at their annual meeting about three years ago. That led to a task-force that is, some of you heard about this yesterday. We're focusing on trying a new way of handling book publishing. But we also had an initiative on scholarly journals that was turned into SHARE. So I, I think on most campuses, provosts are acutely

aware of the key role that research libraries have played and will play even more so in the future. But again, it really is a challenge on these enormously complicated research universities just to get people within a single institution talking to each other.

T. Scott Plutchak: Right.

John Vaughn: But we're working on it.

T. Scott Plutchak: We're out of time. We could go on like this for a long time. I want to thank all of you. I hope what we have done is given you a little bit of a way of thinking about how these things interact, what our various responsibilities are to reach out, as Laurie said, to each other and take advantage of that, but also to reach out beyond into those other communities which are absolutely essential if we're going to get it to where we want to go. Before I end, I need to thank a couple of people who are not here who were very helpful in putting this together. My other colleagues from the roundtable: Fred Dylla who is here, Crispin Taylor, Judy Ruttenberg from ARL helped put this together, Alice Meadows, Liz Ferguson from Wiley, and a big hand for my conspirators up here for participating in the discussion. All right. Thank you all very much.