Interview with Martin Delahunty

For podcast release
Monday, August 26, 2021

KENNEALLY: Part business model, part ideology, open science emphasizes collaboration and transparency for research and research-related publications. Making science open, we'll see information about discoveries and related data become freely available under terms that enable reuse, redistribution, and reproduction. Welcome to Copyright Clearance Center's podcast series, I'm Christopher Kenneally for *Velocity of Content*.

Academic and scientific research moves methodically downstream from lab bench and laptop to data repositories and peer reviewed journals, then finally to public and professional audiences. This workflow shapes and is shaped by policies and mandates of private funders and government agencies. As open science principles increasingly prevail, what are the challenges? How will research be transformed?

Martin Delahunty, Managing Director of InspiringSTEM, an independent scholarly academic and scientific publishing consultancy, has recently undertaken research that explores academic research workflows in support of open science. Delahunty's findings identify important pain points, and the potential impact of such new initiatives. Martin Delahunty joins me now, from his office near London. Welcome to the program, Martin.

DELAHUNTY: Thank you for inviting me, Chris. I've been a avid listener for some time, so I'm delighted to join the conversation.

KENNEALLY: We thank you for joining us. And I suppose for our audience we ought to elaborate on this definition of open science. Tell us about the objective of this dramatic transformation to what's been a centuries-old model of research and publication.

DELAHUNTY: Well, open science, I think can be best defined as the practice of science across all science, technology, engineering, mathematical disciplines, such that others can collaborate and contribute, and where the research data and processes are freely available. This should also be under terms that enable reuse, redistribution, and reproduction. And under this broad umbrella of open science I include peer reviewed publications, data repositories, workflows, collaboration tools, and science policies and mandates. But what is most exciting for me is that the lexicon of open science is well beyond these components now to an even more intricate and complex landscape.

KENNEALLY: Indeed, Martin, it is intricate and complex, as you say. Open science a kind of an umbrella under which there are many other open acesses in the area of the publishing ecosystem that we have spoken a good deal about here on this program. But in your research, working with publishers, and on the research side of the equation, you've identified some priority challenges that academic institutions particularly need to manage so they can optimize the workflow for their researches. There are four of these, tell us about those.

DELAHUNTY: Yes. The first big challenge is the lack of interoperability of supporting workflow tools. And it's easy to see a divergence between the business and practice of research due to the proliferation of new, in many cases freely available research workflow tools. These are software tools. This creates a potential operational financial risk for institutions resulting from lack of interoperability between these freer point of use tools used by individual researchers and their institutions' data systems and repositories.

The second challenge that I see is the increasing pressure on university data repository infrastructures. What I mean by this is the conflicting priorities that arise where individual institutions need to provide their own repositories, but at the same time are now being asked to contribute to the development of national/international repositories. In the UK, for example, continued work is needed to find services supporting open research which requires a national approach. But whether for institutional or national repositories, the critical functions remain, that is discoverability, sustainability, and preservation.

Thirdly we have witnessed the extraordinary growth of open access funder mandates, most notably Clan S, (sp?) but there are many hundreds more and growing by the years. Whilst publishers adapt their publication outputs and systems, pressure is then brought to bear on research institutes and university systems to try and adapt to ensure compliance with these mandates. The benefits of reducing library subscription budgets will have to be balanced with the challenge of updating and sometimes creating new systems to efficiently and economically manage compliance.

And finally, there's the challenge of recruitment, training, and retention of increasingly specialized data management and open science skills now. And I think vendors have a role to play in supporting training and education. This is not just on technical skills, but on the broader mission of open science. With this need to attract highly skilled data experienced candidates, many universities are also producing more detailed open science strategies for data management and knowledge management, and for demonstrating the institutes open science role and its wider mission for science.

KENNEALLY: Well, looking more closely at that institutional role, Martin Delahunty, you see a transformation for the librarians there. From this metamorphosis, you predict, we're going to

see the emergence of data scientists. What repository would data scientists have? What will be the challenges they'll face?

DELAHUNTY: Well, absolutely there's a transformation and metamorphosis. These are vibrant, exciting, and at the same time challenging times for librarians and the custodians of the universities' research data. Increasingly I see recruitment advertisements for special staff to locate within libraries and infrastructures with titles such director of digital innovation and research, or open access and research repository specialists.

And so what are the challenges? Well, firstly is simply getting to grips with these new open science workflows. These represent opportunities for service providers and vendors, but increase the need to be truly customer focused. A common approach I'm finding amongst vendors is to be a development partner with institutions, taking up thought leadership and a very practical role around open science and sharing insights, case studies, and practical recommendations with existing and potential customers.

The second challenge is training. For example, the UK's Open Research Data Task Force recommends alongside enhanced incentives, coordinated efforts are also needed to improve researchers' skills in handling and analyzing data to deliver increased capacity in data science and to provide specialist support services.

Third and lastly is the open science leadership, itself. Universities have a direct role to play in promoting and leading in open science principles, policy, and practice. A recent comprehensive University College London published study is a really nice case study based on how libraries can engage with and offer leadership in open science. This was part of a wider U-funded research (inaudible) project on research data management.

And so UCL identified four leadership roles for universities – open science and open access publishing, research data management, key infrastructures, especially the European open science clouds, and finally (inaudible) science.

KENNEALLY: Martin Delahunty, in the background of all this change, of course, is technology, especially technology standards and the identifiers that help these systems, these workflows operate. Tell us about the role identifiers play and how important they are to this change.

DELAHUNTY: They're important, Chris, and critical. I can say happily real progress has been made with the adoption of persistent identifiers, especially ORCIDs, DOI, and funder identifiers. These provide consistency and transparency via machinery of the links between research and identities, for example researchers disciplines, their affiliations, institutions, funders, and publishers. I think a great example is the UK institutional repositories' adoption of the RIOXX

Metadata Application Profile which is required to comply with the Research Council in the UK open access policy.

However, there's a pressing need for a community-wide solution for institutional identifiers. The ability to standardize disambiguation of (inaudible) is a major obstacle to automated solutions. Although service providers like Ringgold and GRID provide part solutions, there is a move to establish an organizational ID registry with international community governance.

Overall the greater take-up of metadata (inaudible) schema applied to all points research where (inaudible) is required, and where discoverability of content and repositories will also be improved includes clarity around vision of record. This will require community-wide and coordinated efforts, and this might include establishing publisher agreements for metadata and accepted manuscripts, a full integration of ORCID IDs into systems by institutions, funders, and publishers, publishers to share DOIs at point of acceptance, and to establish standardized contributor role taxonomies, for example, credit.

KENNEALLY: Well, Martin Delahunty, it's interesting and ironic, isn't it, that in this open science environment that you describe, and despite all this proliferation of free software tools and applications, you still expect to see an enterprise-level reliance on paid for institutional service providers. So why would that be the case?

DELAHUNTY: Yes, Chris, I think maybe a little surprising, but a clear and consistent message from market research that I've done and conversations that I'm having is the value still attributed by research libraries and institutes for paid-for, and what they might call professional enterprise level services – databases, systems, and associated services. For example those (inaudible) we mentioned includes IET's Inspec, Ex Libris, Elsevier's ScienceDirect, ProQuest, (inaudible) and Digital Science. But there's a continued (inaudible) gap between what librarians recommend and what researchers use. Research is increasingly being directed with (inaudible) and adopt free software tools are easy to sign up to, but frequently lack interoperability with institutional systems. So I do still expect a continued reliance upon enterprise level and paid-for institutional services. These paid-for services come with contractual service level agreements, user support, and best practice guidelines. I feel that this engenders trust and long term reliability for customers who are most concerned about securing and making sure they have persistent management of the organizations' intellectual property. It's also about equipping their researchers with the best tools to advance science, of course, to be competitive.

KENNEALLY: Martin Delahunty, Managing Director of InspiringSTEM the UK. Thanks for an insightful look at the evolution of open science.

DELAHUNTY: Thank you again for inviting me. It's been a pleasure, as always, and I look forward to our next conversation.

KENNEALLY: Likewise. Martin Delahunty, thanks again.

Our co-producer is Jeremy Brieske of Burst Marketing. Today's recording engineer is Rob Simon. You can subscribe to the program wherever you go for podcasts and follow us on Twitter and Facebook. I'm Christopher Kenneally, thanks for listening. Join us again soon for another *Velocity of Content* podcast from CCC..

END OF FILE