Open Science Talk No. 61 (2025): A realistic researcher's take on Open Science services: a computer-generated transcript [1]

00:00:12 Per Pippin Aspaas

Open science talk, the podcast about open science. My name is Per Pippin Aspaas. I'm joined today by Katie Smart, a colleague from UiT's University Library. Welcome to the podcast, Katie.

00:00:27 Katie Smart

Thank you, Per.

00:00:30 PPA

So you've been working here for three years in a couple of months and now you're leaving?

00:00:34 KS

I am leaving.

00:00:37 PPA

You came from the researchers' side and you're going back to research after this, so I thought I would like to hear from you some of the experiences that you've had during your stay here at UiT. But first, where do you come from?

00:00:52 KS

I am originally from Canada. I did my PhD in Canada about 15 years ago, I guess now – now I feel quite old! – before moving to Germany for a Humboldt postdoc and then ultimately taking up an academic position in South Africa for nearly ten years. And then I came to Norway and joined the library.

00:01:12 PPA

Right. So you've been on three continents and most of the time as an adult academic, so to speak, you've been on the researcher's side. And one of the things that I've learned from you is to always have the researcher's perspective – when we try and make services for open science, which is what we do here at our university library and many, many university libraries across the world these days. So I thought I would like to hear first from you about this *Open Access Week*, which is a digital event taking place in Norway each autumn, one week of webinars. And you took over for a while as the chair of this. What was important there, in the Open Access Week, in order to have that work well?

---

[1] This is a computer-generated transcript of the podcast episode Open Science Talk No. 61 (2025): https://doi.org/10.7557/19.8152. The automated transcript has been proofread by Per Pippin Aspaas. It is included here for the sake of Universal Design and improved discoverability by full-text search engines.

00:02:06 KS

Well, I would recall, Per, that actually you said I would find it interesting. So yes, I was a very willing participant to take it over. And from what I understood, it was basically a series of webinars that would span out over a week across Norway from different libraries and institutions that would discuss current issues in, like, Open Access publishing and data publishing. And looking at the previous years' talks, it became immediately obvious that it was very insular. It was sort of, like, librarians and open science people talking to each other, but not really talking to the crowd that needs to hear it, which is, like, the scientists, the researchers – the academic staff who actually are the ones who publish manuscripts and create lots of data. So it was a bit of a difficult endeavour, but the hope was to create sessions that would be interesting not only to people who work at libraries and are passionate about open science and publishing research data openly, but also to the researchers who are taking part in this open science movement. So the first year was maybe less successful, but by the second year I think we were targeting our audience, both audiences, a bit better, I think. With having talks about – one title of one session was particularly good, translated to English it was «Bad Guys in Open Science», like, basically talking about publishers who were maybe not fully coming to the table, which I think sparked some interest amongst the researcher crowd as well.

00:03:35 PPA

Yeah, you used the word «enthusiasts» when we talked about this before we went into this recording studio. You can't expect all researchers to be enthusiast, can you?

00:03:47 KS

No, absolutely not. Coming from a researcher, when I first joined the library, it's like I was coming to the open science religion. And everyone here was sort of, you know, worshipping at the altar of open science and saying how good it is, all the things you can do. And it is good, open science is good, but not really recognising that every open science practise that one wants to do – be it archiving data openly or creating a preprint or working in an electronic lab notebook that's open – it all takes time and researchers don't necessarily have extra time to do these things. So it's important to acknowledge that people might want to participate, researchers might want to participate in open science or open access practises, but the reality is they may not have the capacity to do so, and it may be not high enough up on their priority list, they may not recognise an immediate benefit.

00:04:36 PPA

One of the things that you've been doing is curation of data sets, and this is a word that I think most researchers don't know. They have no clue what is curation of a data set, but you know from the inside.

00:04:51 KS

Yes, I know. So if you have a data set – let's say you've written a paper and there's, like, a Table 1 full of values: in the olden days, like five years ago, you'd put that table of values just as a table in your PDF – like, as part of the PDF of your publication, or maybe you'd have it in a supplement. And that would be good enough, but now a lot of journals, and if you're writing a thesis, they would like to have that, the publishers might like to have that data archived or deposited on another website – on

like, a repository that's openly available. So people – instead of, you know, manually typing in the values from the table published in your journal, they can just click the link: Boom, get your table of values and away they go. So when we get a data set like that – if you want to publish it openly and FAIRly, and I bet you this will be a nice segway to talk about the FAIR Principles – if people want to make their data really open and accessible to people, you can't just dump in your data. It has to be looked at and evaluated to make sure it can be fully understandable and reused by the community, and that's what we mean by curation. You basically make sure you've added in enough metadata (like, data about your data) and that your file formats are right and you've chosen a licence and you've basically ticked all the boxes. So we look at your data set and make sure it's good to go, it's like a mini review process for your data set, and that's called curation. And it can be very easy and straightforward if the depositor has put some work into their data, or it can be quite a lot of work if we've just been, sort of, handed just one table of data with no explanations around it.

00:06:23 PPA

Yeah, because you need explanations, don't you?

00:06:25 KS

Absolutely.

00:06:26 PPA

You need to, in order for it to be reusable to others --. But what's in it for the individual researcher – you called it depositor, I guess that's the researcher who is submitting the data set?

00:06:36 KS

Yes, what's in it for them? I'm sure they'd like to know that too. Well, there's a lot of things in it for them. First of all, they may have to do it. So some journals these days, they have a data policy and that data policy says: «Hey, you want to publish here? Fabulous! You better put your data in an external archive with a DOI link. Otherwise, we're not going to take your paper!» So they might have to do it – all right, so that's one thing. They also may think: «How many times have I sat in front of my computer screen, looking at a published paper and I want to use their values that they've published, but since I can't easily extract them, I have to sit here and type out all the numbers into my own excel sheet.» So if I make my data available in an external data set, which – I can just click a link and grab the data, it makes it so easy for people to use my data again. I don't have to sit in front of my computer and go type, type, type, type, type. So that makes my data easier for people to reuse and work with, which – if you're sort of thinking about your citations, that might increase your citations and your work being more well known. You also may work for an institution – here, like at UiT, where you actually have to make your data externally available on a repository if it's possible. So what's in it for you? Well, it's just nice, first of all. You may have to do it because of a stipulation by the publisher and if you are, if you think it's really nice to share data openly and make it easier for your community to use your data, then you'll want to do it just to, you know, play fair with the community.

00:08:06 PPA

Speaking of community, we also need to meet each other, don't we? And another thing that you've been involved in is to try an convene – with the librarians on one side, the IT department and to meet, then, the researchers – the *UiT Data Days*.

00:08:25 KS

Yes, the UiT Data Days! UiT Data Days was born from frustration with finding a way to effectively communicate how to practise open science and handle research data nicely – how to effectively communicate that with the research community here at UiT. Previously, I think probably born out of COVID we had these Zoom webinars which were horrifically boring – «black screens of death». So UiT Data Days, we wanted to treat it like a conference. You come to a conference – you know, you get some free coffee and some cake, and then you learn something cool. So we tried to market UiT Data Days as a place where you could come for one day – come for all, come fore some – learn something about open science and handling research data, learn some good tools, meet the people who you want to meet so you know who to contact if you have problems in the future. And instead of using, sort of, open science speak or research data speak – like, you know, how to archive your data on a repository – we tried to use titles and keywords that were more catchy to researchers, like: «Need a DOI for your data? Come get one!» «Messy data? We'll show you how to clean it!» «You have to write a DMP because your grant says so? Come on in and we'll tell you how to do it!» So we tried to market it in a way that researchers might be like: «Yeah, actually I have to do that! Cool. I'll go learn how to do that!» So that was Data Days, so it was sort of this, this mini mini working conference, I guess you could call it.

00:09:48 PPA

And how was it received?

00:09:50 KS

I think the first time we ran it was last year, and it was good. We had a really good response. We had maybe upwards of 50 people attending and they were anywhere from PhDs, postdocs, professors, technical staff, engineers – we had a whole wide range of people, which was great. And people were actively working in our sessions because the sessions were built as, like, sort of mini lectures and then followed by active working, so you'd actually get to work on your problems, you could walk away with something done. And I think it was really well received. We had lots of networking and communicating happening at the lunchtime and that was really great. The second time we ran it was also very good, lots of personal connections were made and I think it's well received. People who want to get something out of it will definitely get something out of it.

00:10:39 PPA

I know some institutions internationally, they they have a system of Open Science Ambassadors that they nominate, some members of faculty, of the different disciplines – to go out there and be the go to person. Data Days doesn't really – it's not rigged that way?

00:10:59 KS

We don't have that yet. That's something – behind the scenes we're trying to do is to think about getting members of the academic community here at UiT to come and say: «This has been my experience with open science, and this is why it's worked for me». It's hard to identify those people. We can identify them in house. I mean, Per, you gave a poster about how it worked for you in your role as an academic, but it's been difficult to, as yet identify those champions, so to speak. But I think that would be very powerful if we could get Open Science Champions from science, from social sciences, and what not, to say: «Look, guys, this is what I've done. This is how it's worked. It's been great. You know, here's like a use case of how it can work for you too». And hopefully that will be the way of the future.

00:11:42 PPA

But you need – whether you call them ambassadors or missionaries, or – you need to have someone who can speak to the community with credibility.

00:11:54 KS

Yes!

00:11:55 PPA

And your background is in geology, mine is in humanities and philology in particular. So you, I mean, I can't easily go to your department and speak to your folks. So do you have some reflections on that? How a university library ideally should be rigged to have a good communication –

00:12:18 KS

It's a really hard question. I think it's important for a university library to first of all appreciate the researchers' perspective, which is something I've been beating the drum for for a while and I think it's important to have people employed at the University Library who have that publication background and understand, so they can really effectively communicate with researchers so they can say: «Yeah, I know this could be a bit of a pain. Hear, this is important!» But you don't always have the luxury of having people at the library who have that experience, or significant experience. So I think it's important, as best you can to look at who is depositing data sets, for instance, and contact them and say: «Hey, you've used our services a few times, would you like to talk about it?» And it's hard sometimes because again, you're asking, you're asking more from the academics who again, don't have infinite, you know, unlimited amounts of time. So I think I think that is the struggle. If you don't have a well known Open Science Champion who is an academic, it's hard. It's hard to identify them. But I think they can be a very powerful tool if you can. So I know that I've had success personally talking to my own researchers in departments that I've sort of been responsible for because I do have an academic background. And I know what they find important and what they find cumbersome. And I sort of have a way to cut to the chase and say: «Look, this is what you have to do and you're going to benefit from it. And I have had feedback – I mean, I've had several difficult «customers», let's call them, from the academic environment who have really struggled with how much work it takes to do open science – but I've had feedback from several of them saying: «You know what? The editor of the journal where I've deposited, or I submitted my paper, they

commented how nice my data set was and now I understand why we want to do it like this». So to me that's showing progress.

00:14:02 PPA

Yeah, another anecdote from my side. I've heard of people trying to become professor. You need to submit all your work, and with these Research Assessment Reforms – the committee actually said this person deserves to become a professor because they had published data sets as well as traditional publications, and that was viewed as a positive thing and it showed that this person was, well, at the cutting edge of research development in their field. So this can actually work, can it?

00:14:35 KS

Yeah, it can, absolutely. And that's a really positive thing to hear. I think that something that, like, our colleagues here at the library would say is that the way that researchers are assessed also needs to change. Not just by their age factor, how many citations they have, or how many publications they have – I mean, that's all important too, of course, and most academics would say that's extremely important. But it's also important to assess how much this person is doing in open science practises, depositing data sets – which are publications, they have a DOI, they can be cited – so if they could be worked into the metrics that would be extremely beneficial, I think.

00:15:10 PPA

Another thing that many university libraries are engaged in, of course, is teaching and courses. And I know you've been involved in the *Data Stewardship Course*. Could you tell us something about that?

00:15:22 KS

Yeah. So I mean, you'll correct me when I'm wrong – because I never get the details right, but it's like sort of an open education course. It was developed here at UIT, in the library, by several of our colleagues and it's a three-part course. It's called Data Stewardship and it wants to build transferable skills so that even if you're getting a PhD in bioinformatics or social sciences or linguistics, you learn some common skills in research data management so that at the end of your PhD, OK, you're very hyper specialised in something very fancy. But you also have really good skills on how to handle research data. So there's three parts. So there's one which is an online module where you basically learn the practicalities of research data management. There‘s a second module where we do group work together in person for about a week, and that's what I've been heavily involved in. And the third part is – we send you out into the world, and you basically do a baby internship at a non academic partner to sort of become a data manager for a few days. So by the end of that course, you've learned some solid skills in research data management, but really importantly, research data management of not your own data, of data that's being used by someone else. So you – sort of, you gain some really good skills, so if you don't end up going into academia or working in a university environment, you could go and work for a company and be a their data manager.

00:16:40 PPA

Yeah, we're talking about the DocEnhance Project, which has developed a Data Stewardship Course. You said first, they take an online module, then they have about a week of workshops with, typically, librarians?

00:16:53 KS

Yeah, with us.

00:16:56 PPA

How has that panned out?

00:16:59 KS

Planned out? Yeah, so we basically set up a week and we – sort of, we have sessions, about five sessions and each session has a theme like, for instance, in one session we might talk about Data Management Plans – how to write them, how to evaluate them, what makes a good one, which makes a bad one. We made up a session on how to archive your data – like, what information do you need to archive your data in a FAIR manner, so, you know, including things like a readme file, correct metadata, thinking about licences. So we have these targeted sessions and we have, sort of, problems that they have to solve or they actually might, might actually try to archive data and see what data they actually have to physically enter into a database before they can proceed to depositing anything. So it's really hands on. But importantly, it's, like, we all handle data as researchers. But now when you have to handle data for somebody else and you have to really think about clarity and reusability, and that's what the course is really great at. We've just been marking the exams actually this morning and it's really nice to see the reflections from the students who say, you know: «Handling my own messy data and bad file structure is one thing, but handling for someone else really makes me appreciate how important it is».

00:18:09 PPA

So that's after the third module?

00:18:11 KS

Yes. When they've been with a company for a few days.

00:18:14 PPA

Right. Who are these typical course participants?

00:18:20 KS

Generally, PhD students, although we do have postdocs now and again. They come from everywhere – biological Sciences, geological sciences, linguistics, informatics, you name it, we've had them. So we get everybody, people who may not have much quantitative data – maybe they deal with pictures or videos or audio recordings – to people who are hard scientists with Excel spreadsheets full of numbers. We get everybody. It's really interesting to hear the different perspectives.

00:18:52 PPA

You said age is a factor. So these are, usually, what they call junior researchers or early career researchers. How would you say, for a university library doing this outreach thing in open science, is there anything particular to think about, depending on what age, or I mean, academic age group you're targeting?

00:19:16 KS

Yeah. OK, this is really important, actually, because you have the, as you say, early career scientists. These are the ones who are PhD's and maybe postdocs who are relatively younger academically and they're more open to new ideas and they understand – they're more malleable, they can understand the importance of open science much more easily, and they know that they need to prove open science practises to, you know, get a EU funding granted, for instance. And also they're always in for the free lunch at these events. So they're there, they are there for it. And so we often speak a lot to PhD students, for sure. The harder part with, sort of, trying to show academics the importance of open science, is talking to the older academics – and I will include myself in this group because I've been around for a little while now. These are people who have grown up academically publishing with traditional publishing houses – you know, reviewing for publishing houses, only paying a paid charge if the journal demands it, maybe paying for Open Access, but those are the people who have not, or have not been exposed as much to this open science system and are potentially a little bit suspicious of it, like: «Share my data openly? My God, someone's going to take it!» or: «You want me to give a preprint? Why should I do that? Someone will take my idea and steal it!» They're a little bit more – and this is myself, I include myself in this – they're a little bit more reluctant. I mean, they – obviously, no one's gonna say: «Ohh! Publishing open is bad.» Of course we want to share your science as wide as you can – you know, to be generous, but also maybe just to get get a bigger readership. Of course! But to convince them to spend their time, to commit to these open science practises, that's harder. But one would hope that if their students and their postdocs are, sort of, bringing this knowledge back to them, it maybe become more of a happy practise, they may be happier to do it. But certainly we really want to target, sort of, the established, let's call them established academics, to sort of come over to the open science side or at least consider it.

00:21:20 PPA

I would like to thank you now, towards the end of our podcast, for the efforts that you made for three years here, where you certainly put a mark on our local university library. Are there any further thoughts? Some Take Home Message that you want to give us before you move on to the next stage in your academic career?

00:21:43 KS

I don't – I just know that, I think – now I must think, and this is difficult. I think that we've come a little way in the time I've been here and I don't want to say: «Oh, I did so well, I helped us do all these things!» but I think it's important to realise – and the university library is essentially a service provider, we're providing services to academics – and I think, I hope – what I've tried to do is to show us that: «That's great that we're providing the services, but you have to make the other people want the services». So you have to – it's basically marketing. You have to market and appreciate the customer, who is the academic here, and appreciate that, you know, their time is limited, I mean everybody's time is limited but their time, their capacity for improving their open science practises is limited. So we want to make it – we want to show them the benefit of it, which in general will benefit open science overall.

00:22:34 PPA

Thank you, Katie Smart, for coming to the podcast.

00:22:37 KS

You're welcome, Per. Thanks for having me.

00:22:41 PPA

Open Science Talk is produced by the University Library of UiT the Arctic University of Norway. Thanks for listening.