PDR Host Workshop on FAIR-ification of Data

Interview with Helen Malone,
PDR, President, and Director, Global Information Hub, GlaxoSmithKline

For podcast release
Wednesday 17 April 2019

KENNEALLY: Findable, accessible, interoperable, reusable. Those are the guiding principles for researchers and publishers that together add up to a FAIR way to manage scientific data.

I'm Christopher Kenneally for Copyright Clearance Center's podcast series. Welcome. On May 21st and 22nd at the Heathrow Marriott, just outside central London, the Pharma Documentation Ring will host a special workshop on the FAIR-ification of external data. The PDR, a community of innovative knowledge managers from 28 leading pharmaceutical firms, sees adherence to the FAIR principles as critical for all stakeholders in the healthcare industry. Research data that is findable, accessible, interoperable, and reusable, according to PDR, will allow scientists to uncover hidden connections and insights.

Helen Malone, president of the PDR and director of the Global Information Hub at GlaxoSmithKline, maintains the FAIR principles can lead to breaking down data silos. The PDR workshop aims to accelerate that process. Helen Malone joins me now from her office in Stevenage, the United Kingdom. Welcome, Helen.

MALONE: Hi, Chris. Thanks very much for inviting me.

KENNEALLY: Well, we're looking forward to learning about this very special program coming to just outside London on May 21st and 22nd. The PDR is organizing it, and as we say, it is an effort to look at breaking down data silos by adhering to these FAIR principles. FAIR is the acronym here for, as we have said, findable, accessible, interoperable, and reusable. Can you tell us about more about these FAIR principles and each of the points? Expand on them – why is findable important, why accessible, and so forth?

MALONE: Of course. So the FAIR principles are all about ensuring that both internal data, like R&D clinical trials, and external data, such as scientific or healthcare literature, is FAIR. And as you said, FAIR is actually an acronym, so it covers the four principles of findability, accessibility, interoperability, and reusability.

Findability really means that the data should be able to be found by an appropriate person at an appropriate time. Accessible means that it's accessible either

internally through a license or publicly available. Interoperable in terms of how the data is formatted. It's standardized and annotated. And reusable for both people and machines.

So FAIR actually began in the academic community, and since then, it's really started to be embraced by science funding agencies such as the European Commission as well as the National Institutes of Health in the US. Essentially, the idea of FAIR is really to improve both discoverability and reuse of digital resources and data.

KENNEALLY: The principles address, though, which particular issues that you face at a company like GlaxoSmithKline? Anything in particular you want to highlight that this is really important for?

MALONE: At GSK, for example, we've started on the FAIR journey, and we're looking first at what we can do to improve our internal data. And what we hope is that FAIR is going to help us break down those silos between different departments and get us that better usage across data and across the R&D pipeline.

KENNEALLY: And these principles are especially important not only for the pharma industry, but for all of digital healthcare. That is what, of course, is driving the development of this very special workshop that you will be holding – PDR is organizing there in May. I wonder if you can tell us about some of the challenges and the opportunities that are presented to organizations like your own when you FAIR-ify, when you have allowed for the FAIR-ification of external data.

MALONE: Yeah, so you're probably aware, Chris, that artificial intelligence and machine learning has really been at the forefront and the front of center of the digital healthcare revolution. So the principles of FAIR data really do underpin that successful implementation of these important analytical tools. And the pharma industry in particular has recognized that potential of applying these analytical tools to the vast amount of data that we have access to in order to develop more personalized and targeted medicines. So that's really ultimately the benefit that we're trying to get to, is to develop these more targeted drugs more quickly to bring more benefits to patients.

And while most of the attention up to now has been essentially looking inwardly at our own internal data, external data, as I'm sure you're aware, is growing exponentially. It's much more accessible than it ever was to anybody. So companies such as GSK are more outwardly looking than ever before. That's really the opportunity.

In terms of the challenges for the FAIR-ification of data, importantly, changes in company culture, particularly to understand the value and the importance of FAIR principles. And other challenges include the likely increase in cost to develop additional technical infrastructure and also the potential restrictions on the use of the data.

KENNEALLY: Well, these are all points that you will be covering on the agenda for the PDR workshop on the 21st of May and the 22nd. And you've got some special guests to help you do that. Tell us about some of those.

MALONE: Yes. Well, I'm really excited, actually, about the up and coming PDR workshop. It's focused, as you say, on the FAIR-ification of external data. So we've got a great lineup of experts, thought leaders from a diverse range of organizations.

As an example, we have Erik Schultes from the GO FAIR Initiative, which is really a leading think tank and becoming incredibly influential in building an international consortium on implementing FAIR principles. We also have Nadia Anwar from Clarivate Analytics, who will be discussing the vendor perspective on FAIR. And Will Spooner from Genomics England, who will be talking about the specific challenges and opportunities of applying FAIR to genomics data.

We've actually opened up the workshop so that anyone with an interest in the FAIR-ification of data can join us. So we'll have delegates from academia, the pharma industry, publishers, and other external organizations with an interest in FAIR.

KENNEALLY: And there are some special breakout themes – workshop breakout themes on the second day of the program looking at licensing, integration and ontologies, and technologies and architecture. This all gets very complex very fast, but give us the highlights there. What are you trying to touch on with those workshops?

MALONE: So particularly with the licensing, we're looking to see how we can develop fit for purpose licenses, particularly to enable the reuse of data for machines and humans alike and really exploring how the FAIR principles help to support that.

On the integration and ontologies work stream, we're looking at how we create a sharing culture which particularly standardizes metadata, it integrates ontologies, and really looks at new consumer trends and different types of data, such as voice and images, for example.

And then in the technologies and architecture discussion, we're really looking at the many choices that we have for data storage and management, such as federated systems and cloud solutions, and looking to see how FAIR can help.

KENNEALLY: Indeed, it sounds like a pretty ambitious program, putting a lot into a day and a half there at the Heathrow Marriott on the 21st and 22nd of May. I wonder if I can end our discussion, Helen Malone, by asking you about the importance of this sharing culture – a sharing culture that is going to lead to development of drugs and treatments and solutions that are going to matter to many people listening, because it will help address the health needs that they may have or their families. And this sharing culture is something which is really made possible by this network of information.

MALONE: Yes. Certainly from a personal point of view, I think a sharing culture is absolutely vital in order for us to move forward. And it really needs to be driven by both a bottom-up approach, from scientists who really understand the need for FAIR data, and also a top-down commitment and investment from senior executives. So the cultural change that we would be looking for is to be enhanced and encouraged by having incentives to share data, and I also think peer recognition both internally and externally would help enormously. So I do believe that these FAIR principles will help us move much more towards that sharing culture, which will help to improve the flow of R&D data, breaking down silos between different departments and also externally, and ultimately bringing much more benefit to patients with the opportunity to develop more targeted, personalized medicines.

KENNEALLY: We wish you the best of luck with this very special program that is going to be presented by the Pharma Documentation Ring, coming to the Heathrow Marriott on May 21st and 22nd, looking at breaking down the data silos by the FAIR-ification of scientific data.

We have been speaking today with Helen Malone, who is the president of the PDR and director of the Global Information Hub at GlaxoSmithKline. Helen Malone, thank you so much for joining me today.

MALONE: Oh, thanks very much, Chris. It's been great talking to you.

KENNEALLY: Today's podcast is produced by Copyright Clearance Center. CCC builds unique solutions that connect content and rights in contextually relevant ways through software and professional services. CCC helps people navigate vast amounts of data to discover actionable insights, enabling them to innovate and make informed decisions. CCC is headquartered in Danvers, Massachusetts, and has offices across North America, Europe, and Asia. To learn more about Copyright Clearance, visit copyright.com.

Our producer is Jeremy Brieske of Burst Marketing. I'm Christopher Kenneally. Thanks for listening.

END OF FILE