Welcome to episode seven of SSP's Early Career Publishing podcast. This is Meredith coming to you from Boston, and I'm joined by my co-host Sara from her home in London. For those of you who are new to the podcast, this series aims to serve all facets of the scholarly publishing industry by bringing together insight, information, and guidance from leaders and experts in the field. Today is the first episode of a two-part series that focuses on open access in scholarly publishing. To prepare for this series, we gathered a list of questions from some of SSP's early career professionals. And we'll cover a few of those questions in each episode. In part one, Sara and I will be talking through the basics of open access. And for part two, we'll welcome back Ann Michael of Delta Think to go a bit more in-depth on OA topics. Sara, first of all, welcome back to the podcast, and congrats on adding a new member to your family.

Great. Shall we get started with our basic OA questions and answers?

Absolutely. Let's do it. So the first question that we were sent is what is open access? Whilst this may seem obvious to many of you, for any of our new entrants to the industry, this is a very, very good question to have answered because it's critical to what we do. So in the strictest sense, open access is defined as the distribution of content online without any cost or other barriers to access. It's usually licensed by a specific open access license, where the content tends to be free of most reuse restrictions, but requires attribution. So for example, I'm a scientist, I take some work from another scientist. And I have to say - it's open access, but I
fewer barriers to entry, they wanted faster publication times, and global free access. So, of that open access was a key part of author choice, they demanded publication outlets that had fewer barriers to entry, they wanted faster publication times, and global free access. So, of

need to say - this work that I've used was done by this scientist. So, attribution. So papers, books, data sets, or any other piece of content that's published open access is there for free to read and use by anyone globally. Some people do argue that for an article or piece of content to be pure open access, the content needs to have zero barriers. So that includes for reuse. So anyone could take an article and reuse figures, text, data, whatever it is, even without attribution. Other people would say that a paper only needs to be published in front of a paywall, so no one needs to have to pay for it to be open access, and there are many, many shades in between. Strictly in its definition, open access has many definitions. I have no intention of taking a side here, but just want to make clear to anyone listening, that there are differences in how open access is defined. Our next question is what is the difference between open access, hybrid, and subscription models? This is an excellent question. So a subscription journal is exactly what it says it is. A reader has to subscribe to get access or be part of an institute or department, lab, faculty, whatever the case may be, that subscribes. This tends to be what we would call traditional publication. So when someone is publishing in a subscription journal, authors might choose to put the version that they originally submitted online in a repository, such as the Archive or on their own personal website. And that's called green open access. So even within a subscription journal, you can still have a form of open access. Most publishers do have a website dedicated to open access that explains the difference between green open access and gold. And I strongly encourage you to have a quick Google and find some of these resources. A hybrid journal is a journal in which some of the articles are open access and some are accessed via the subscription model. This kind of journal does require that the author pays a publication charge to the publisher if they want to publish open access. But it also means that content that isn't published open access can be accessed via a regular subscription. So hybrid journals are really designed to give authors choice as to whether or not they publish open access or whether they go via the more traditional subscription route. Finally, open access is a journal that's funded exclusively via the payment of an adequate process charge, which we would call an APC. All authors except for those that may be given a fee waiver for any given reason will pay an APC to the publisher. And that covers everything. It covers editorial, publication, web hosting, and the many other costs that are actually involved in the publication process that are typically covered by subscriptions. Where an author can't pay due to hardship or where they're based or funding constraints, they can actually request a fee waiver so they can still publish in an open access journal, but they might not have to pay the APC. We will cover this more in the second part of our podcast when we're joined by Ann Michael from Delta Think. Our next question is why has open access gained so much interest and attention in our industry? So open access has been an absolute game changer in how publishing works. I'm not going to attempt to name the first OA journal because if I did, I would be Googling for hours and still potentially get it wrong. But by the mid 2000s, the launch and success of journals like PLOS ONE showed that open access could offer authors something that they hadn't had before. Because open access doesn't require a subscription, journals were free to publish as many papers as authors demanded, and as peer review allowed. Journals like PLOS ONE were able to publish thousands of articles per year. And it was a pretty straightforward route to publication for authors who might have previously struggled in more niche publications. Authors paid a flat fee, they were published online relatively quickly, and they didn't have to wait for things like print dates and print publication. In addition to this, the publication of articles without a paywall gave access to research globally. So researchers lacking access to subscription journals had barriers to access eliminated and could read anything that was accompanied by an OA license. So this was of course key in furthering the goals of the scientific community to advance science and research. And of course, once a critical mass of successful journals emerged, interest and attention followed. Publishers saw that open access was a key part of author choice, they demanded publication outlets that had fewer barriers to entry, they wanted faster publication times, and global free access. So, of
course, any business model that successful generates attention and open access was no different. What we now see is funder mandates that request that authors publish their content open access. So, for example, the Gates Foundation, they require all authors to publish their work open access. And where a funder mandates OA, publishers need to provide avenues to enable these mandates to be delivered. Open access is here to stay, it’s become a key part of the publishing infrastructure. I think we could do a podcast on its own really just of this question. One final thing I will say though, is while we tend to think of open access for journals in the hard sciences, it now increasingly covers ALL types of research output, and that includes books and the social sciences and humanities. Our next question is how does open access work for the author? So what are the mechanics of open access? This is one of my favorite questions because I like to dispel the myth that there is any difference at all. Open access works pretty much identically to non-open access or subscription content. There is very little difference from submission through the peer review to final publication. It is almost identical. The key difference is in fact an author clicking a button that says I want to publish this open access in that early stage of the process. What editors want is sound, well done science that fits the journal's aims and scope, or the book - publishing program's aims and scope, regardless of whether it's open access or subscription. So authors will submit a paper to their journal of choice, they'll make sure all of the usual things -- formatting, approval from all authors, adding suggested reviewers, etc. -- all those things will be there, as usual. Authors might be asked to confirm that they know the title's open access and has a fee. That's the key difference until the paper is actually through peer review and ready to be accepted. At that point, the authors might need to choose a license to publish their work under. It needs to be paid for by the authors themselves, the institute, the lab, the funder, whoever it is. And at that point, it's placed in front of the paywall when it's published, instead of behind a subscription paywall, and it can be accessed globally. Otherwise, it's pretty much the same. Meredith, over to you for the next question.

Meredith Adinolfi 08:38

Great. Thanks, Sara. Our next question is what do internal operations workflows look like for OA content? And where do pain points exist that need technological or other solutions? The first part of this question overlaps a bit with what Sara was just talking about in the previous answer, but it's still, I think, some interesting nuances. So as Sara talked about, at the core, publishing open access articles or journals is not very different operationally from publishing subscription or any other type of content. In general, many of the workflows and protocols apply across different models. One important piece to consider for OA content is the information that publishers need to capture about funding to ensure that processing, invoicing, and ultimately, publication happen correctly and in line with whatever the author is expecting. This is particularly important for hybrid journals, where it's critical that information is captured properly in the system to ensure an accurate result in terms of the publishing model that the author has selected. So there's an increased need for publisher systems to support collection of key information at various stages of the process. It might happen within a submission or production system, depending on what that information is. So it's really important for publishers to make communication as clear and concise as possible, which will hopefully reduce the burden on authors as much as possible and ensure that they know what they need to provide when and how to fill out their paperwork for whatever they want to happen. This is particularly important because we know that submission and publishing requirements can be a recurring pain point for authors. So anything publishers have to add to that process can only increase that feeling. So there again, underlining the need for clear communication throughout. Also, because of the variety of types of OA licenses and funding situations that now exist,
publishers have likely seen an increase in queries and troubleshooting requests in these areas. So I think we can expect publishers to continue optimizing their workflows and systems so that they can ensure they can keep scaling up without compromising on quality of author service. Of course, a big piece of all of this is just simply listening to authors about their experiences, and based on that feedback, being open to experimenting within the systems that are in place.

Our next question we're going to cover is what pieces of metadata are key in open access publishing to make the outputs more discoverable? And is this different than in subscription-based publishing? Let's start here with just a quick definition of metadata to make sure that's clear. In the context of this conversation, metadata is basically information that provides details about an article's contents or components. And metadata is a significant part of the discoverability conversation, especially because of the rise in artificial intelligence tools, which require information to be structured and as machine-readable as possible to work properly.

Of course, a big piece of all of this is just simply listening to authors about their experiences, and based on that feedback, being open to experimenting within the systems that are in place. Of you might remember that a few years ago, the industry saw a good amount of buzz around an initiative called Metadata 20/20, which was a collaboration between a few organizations with the intention of providing more structure and consistency guidelines for the metadata that's used in scholarly articles. The ultimate goal there was to optimize the use of metadata to make articles discoverable, and as discoverable as possible, and also easy to process by both systems and people. Though we're not going to talk in any detail about Metadata 20/20, I mention it because I think the four core tenants provide a good framework to think about metadata, which should be compatible (meaning that it's parsable and readable by both human and machine), it should be complete (meaning it reflects the contents and components fully), it should be credible (meaning it's accurate and reliable), and it should be curated (meaning it must be up-to-date and consistently maintained in order to be useful).

Some examples of core metadata elements for articles are things like the DOI number, ORCIDs, license information, and funding or grant data. And that's just to name a few, there are many others that we could add to that list. So publishers have an important role here in staying up-to-date on metadata standards and practices and then applying those to the best of their abilities because this is really the place where a framework can be enforced that will ultimately lead to the best result in terms of discoverability. Now, as for the second part of this question, the importance of metadata is not necessarily specific to open access publishing. But because one of the advantages of OA publishing models is that articles are widely and freely available, to get the most value out of that advantage, discoverability becomes a really important piece. And now for our final question, a very broad one: why are publishers willing to embrace open access? I think this answer overlaps a bit with another answer that Sara discussed earlier. It's important to remember that researchers, publishers, and funders all share a common goal of wanting to advance science and therefore improve life. So a critical piece of that goal is ensuring that research reaches a wide audience and can be read and used and built on by others. So if you start to think about it that way, it makes sense then that publishers want to evolve their publishing models and options as the industry evolves and changes in terms of what the needs are and the expectations. Over the past few years, the demand for open access publishing by researchers in the scientific community is increased significantly, which means that publishers have had to continuously evaluate what they're offering, and also spend time listening to the community about what they want and what best serves their research. Well, as Sara said earlier, there are also funding body mandates around open access that make it important for publishers to have choices to offer all researchers regardless of their institution or funding body. You might often hear the phrase "author choice" in the current environment, and this is exactly what that means. With new funding models and options developing over recent years, it has become more palatable for publishers to offer a range of OA options and different types of journals. We're going to talk a lot more about those developments in part two where we're welcoming back Ann Michael as our guest. So for now, we're going to leave this here as a teaser for the next part of the conversation.
Sara Grimme  15:06

We do hope this has been informative and that we haven't thrown too much information at you. We hope you'll join us in part two of our open access series where we will be joined by Ann. And in the second part, we'll delve further into open access addressing such issues as funder mandates, Plan S, transformative agreements, and the future of open access. Thank you so much for listening.