The Authority File: Episode 430 Transcript

Transcript generated by Zoom's automated AI captioning.

Bill Mickey: So as we've talked about there's a lot of concern over responsible use of AI you know, but I'm wondering, and you've hinted at some of this, some of the stuff that's going on behind the scenes at Springer, but what are some of the advantages you see with AI offering, you know, in the publishing process, you know, particularly for research, discovery, and accessibility?

Anil: At a very sort of basic level AI will definitely, is enhancing accessibility and efficiency. For example, Springer Nature has implemented AI tools to generate sort of plain language summaries making research more approachable for non-specialist audiences are much more focused and faster in terms of dissemination. We also have faster language translations, for instance, which we have done for our German language books that were previously not available in English, and

Anil: these translation, as you know, academic translations in particular have been laborious and time consuming in the past. So while we still need human interventions to validate the quality, but the general basic translation services provided by the AI models are improving every day, and they are improving in quality. So broadly at Springer in AI I would characterize this into sort of 3 broad models.

Anil: One is called sort of text generation and text generation is, I'm not certain, Bill, if you remember, but we did publish an AI authored book in 2018, 2019 that was essentially a sort of research summary put together using an algorithm.

Bill Mickey: Yeah, I'm gonna ask you about that actually.

Anil: Yeah, the others are sort of text conversion, where, which is where we broadly place sort of research summaries now, or translation projects. And the last one is sort of text interrogation. And this text interrogation, what happens is you basically use human expertise. And here, when I say human expertise. It's usually our authors or our editors who have a background and knowledge of a particular subject in depth, and they work with specific AI models based on questions, etc.

Anil: We've we, in fact, had a project where we got our authors and our editorial teams together, and then we they worked with an LLM together to develop the framework for a book on accounting practices and principles. It was in German language, and the associated text, through a prompt and the prompt and or prompt ping pong method. So you would ask the AI a prompt, it would generate an answer. Then, using that, they developed the table of contents, as

well as the text that went into it, and that meant that the book actually got, was validated at each point, or the text was validated at each point.

Anil: And so, to a very great extent, it really led to cutting the time required for developing a manuscript quite substantially using such a project. So that was an example that we did as well.

Bill Mickey: Awesome. So yeah, I did want to ask you about that, the title you had AI write back in 2018, 2019 and I'm glad you brought up the accounting title now, too, because it sounds like perhaps the methods you use to create both of those titles were slightly different, I'm guessing.

Bill Mickey: But so I'm wondering, you know, with the title in 2019, which was machine generated and about lithium ion batteries, you know, what was the goal of that, like what was the goal of creating that prototype? And what have you learned from that exercise?

Anil: Yeah, so it was very much experimental at that time. But yeah, you know, the discussions around artificial intelligence and or, more directly, algorithms, etc. have been kind of circling for the last 5 to 6 years, right? There was certainly pre-pandemic when it actually started.

Anil: And so part of the reason why we wanted to experiment in that space is that as early as possible with algorithms and AI is to explore ways of how we can innovate with our content as much as possible, so meant in general to develop something that would contribute to supporting our communities and our readers in the short or medium term.

Anil: It also sort of underscores establishing some thought leadership which seeks to bring a future a bit earlier for the communities. You know, that's how we actually see it. So it was very experimental to see if we could actually develop something at an improved level of efficiency.

Anil: So the AI authored book project like you mentioned on lithium Ion batteries was aiming to explore the sort of nascent capabilities of AI and content generation. That was the 1st model that I was talking about. It demonstrated that while AI can compile information, so what it was is basically a compilation of information, but then it followed the broad structure of a book.

Anil: It sort of lacked the nuanced understanding and critical and meaningful analysis that human authors provide, and while developing that we had human intervention, but only in terms of just validating whatever not the content itself, but just that the AI didn't actually fall out of sync with what was the process of creating a book.

Bill Mickey: Okay.

Anil: So there was value in automation. That is something that we discovered, but only as a tool, and certainly not as a replacement. But we, we quickly sort of came to the conclusion that that is not the ideal, and we have certainly moved away from that content generation approach

Anil: because it and we recognized its limitations and moved. And that movement has been in the direction of the other approaches. I mentioned what my colleagues in our content innovation teams, for instance, called the human in the loop solutions that said, as they affirm, using generative AI quite reliably and correctly, within the framework I mentioned earlier. It's critical, and not a really a small task to actually do this. And that's where we got in

Anil: already existing experts and our others are experts. They create content also for other experts who are trying to develop expertise in allied areas. So we are quite fully committed in terms of how we interrogate and publish things, that the human validation is, has to be absolutely built into that particular workflow. So whether it's in the text sort of interrogation methods,

Anil: the fact that experts are always present makes sure that we have sort of fully committed to making sure that it aligns to our core guiding principles which I mentioned earlier, while also being a productive tool to our others, and for our readers. So there are some efficiencies gained, but nevertheless, within and clearly placed within the responsibility and accountability of the experts with whom we work very closely.

Bill Mickey: Yep, yep, excellent

Bill Mickey: So I wanted to go back to maybe talking about the text conversion method you discussed earlier and more specifically, language and translation. And very curious to see, you know, as a global publisher, what potential have you realized out of translation processes and also for peer review like, how does that work in that realm?

Anil: Yeah, so certainly in the in terms of breaking language barriers and translating, I mentioned German to English translations, you know several books in German which we publish mostly for a market in in Germany, Austria, and Switzerland that can be made available to a wider non-German speaking audience internationally.

Anil: And we have had a lot of success here. It's not just German that we translated. We have also experimented in Chinese and Japanese language and Portuguese as well, and with GPTs being released, LLMs in play, there is no real limit to what language you can actually choose, and we have had a lot of success there in terms of making content available in multiple languages for a global audience.

Anil: It's so in terms of breaking language barriers and streamlining, you mentioned peer review, it's crucial to maintain again, yeah, the key thing to say is human oversight. To ensure accuracy and upholding ethical standards has to be followed.

Anil: Translation has just shown us that there is a huge opportunity there, and global scholarship certainly profits from that. So that's an area where there is a lot of already translated content available, but we, of course, get it validated by our all those editors before it actually is disseminated.

Anil: The peer review topic is actually a bit more complex. Sure enough, those the potential in review or matching or detecting conflicts of interest and identifying recycled content, and also for initial screenings, and considering the number of articles and the books that require a peer review are constantly increasing

Anil: and demand certainly outstrips supply in finding peer reviewers consistently who are able to make a timely review, which is also a process at the core of academic publishing. Hence the timeline for these decisions. So that is a necessity. But the review process is one that is also, I would say, deeply human. AI could support it, but maybe not replace it.

Anil: Just a few weeks ago Nature did a survey which found that researchers surveyed overwhelmingly expected AI to be used in language, correction, or preparatory work, but when it came to offering a decision for quality, then AI was not sort of something that they thought they could support as a tool.

Anil: So AI for a peer review, although we can in theory say that they could be used for screenings and other, there are several possibilities, obviously, but researchers had a specific view on that. But that could be just a survey at a given point in time. We have to see how those, how that evolves as we go forward. The other part of the peer review process, where AI or there was some concern around using it for peer review is privacy. For privacy reasons automated peer reviews are seen as sort of less adoptable. But, like I said, this is an important area of ongoing learning. And we have to see how this will develop going forward.