Exploring AI Ethics & Publishing

Recorded at STM 50th Anniversary Frankfurt Conference
15 October 2019
with

• Attorney Carlo Scollo Lavizzari
• Niels Peter Thomas, Springer Nature

For podcast release
October 16, 2023

KENNEALLY: Public policy and legislation across the globe regulate and restrict technology in countless ways. Yet the laws of technology itself are few. And even among those, most are generally not well known.

Welcome to CCC's podcast series. I'm Christopher Kenneally for Velocity of Content.

Melvin Kranzberg, an American academic whose field was the history of technology, created just six laws of technology nearly 35 years ago. The most essential of Kranzberg's laws is the first one – technology is neither good nor bad, nor is it neutral.

The emergence in our time of enormously powerful computing technology has moved the concept or artificial intelligence out of science fiction and into commercial and cultural fact. AI is building a strong presence in the scholarly publishing industry, where there are now many new AI-related initiatives, products, and services.

These opportunities challenge us to consider the area of ethics as applied to AI uses within publishing.

On Thursday at the Frankfurt Book Fair this week, I will moderate a Frankfurt Studio panel discussion, **AI Solutions – Trained With Your Content?** Court decisions, regulation, and legislation will ultimately create the legal guardrails for protecting copyrighted content from unchecked infringement. In the meantime, debate over any limits to be placed on training LLMs must address concerns over equity, quality, transparency, and authenticity.

In 2019 at the STM Association's Frankfurt Conference, I spoke Carlo Scollo Lavizarri and Niels Peter Thomas about publishing industry-led initiatives under development even then to govern AI and use the technology to the benefit of scientific research.

… Carlo, I'd like to open with you. STM has been on this story for some time. And indeed, looking back to Future Lab, I believe in 2018, you identified that the future reader – or the reader of the future, rather – is a machine. That was the starting point for some of the work you are doing. So tell us about that.

LAVIZZARI: Thank you very much for this introduction. Indeed, I can recognize the Future Lab of STM and AFK who has given a great panel this morning for recognizing this early on the horizon. I think STM actually has a history of identifying topics, including – I can only recommend the Future Lab to all of you to attend. And indeed, AI, I think by now – 2019-2020 – you cannot open any website or any newspaper or anything not AI is really the topic there. So already starting back in 2017-2018 at STM, we found that it is necessary to break this down a little bit and unpack and look, what does it mean for publishers?

We have to realize that for the foreseeable future, it's about data and quality data and whether publishers can be perceived as partners to make this data available, because the entire AI world is really a development now on big data and big computing power The algorithms have been around for 40-50 years. That's not new. What's really new is the capacity to harness data. And a further avalanche of data is coming towards us, in any case, that will swamp all discussions. At least that's my prediction. Whether subscription or open access, it'll be all about data then.

We've broken this down to four areas. How does AI help publishers to do a better job in house, so to speak – back office – even though back office doesn't really exist anymore in a digital company? But second, how can AI help to make the content that publishers sell more valuable – or offer? And thirdly, are there specific skills in organizing content, in curating content, that publishers can actually extend? And AI may, in fact become the main thing the publishers of the future will do – and indeed, maybe the reader of the future is perhaps a machine.

KENNEALLY: Well, expand on that point, because this is an area that publishers already have some expertise in, right – that is, curating, organizing data, information. So this is a natural progression of that work.

LAVIZZARI: Correct. So data – in German, we have this word called data salad. It's basically an assortment of, I suppose, a haystack that is unusable of data. So what you really need to do a successful AI project is organize the data, prepare it – I'm sure we will hear more

about this – so that it can be used for machine learning, for reinforced learning, directed learning, for testing, for very fine – for debiasing all the things that need to happen so that an artificially intelligent entity that perceives the world effectively using data as their map of the world – so that they are not ending up with misguided results.

KENNEALLY: Talk about the role that publishers can play in helping to address those concerns around trustworthy data and the danger if that's not done.

LAVIZZARI: For people to trust AI advances and AI tools to use, there needs to be transparency, needs to be clarity as to the provenance of the data used. If I'm being accepted or rejected as a patient in a hospital in a decision tool situation, I need to be able to trust this decision. Similarly, for a hiring situation, if the person making the triage is in fact the machine, I want to be sure there are no unfair biases based on historical data that this machine is going to use.

So I think publishers will have a great role to assure that the data used to train machines before they are let loose, so to speak, in the wild is correct and also to almost become escrow agents for the log entry, effectively, of the data that the machine generates while it is running.

KENNEALLY: We've been using almost interchangeably ethics and law and legislation, and for many, they are more or less the same. But they are different, and important distinctions are there. Draw those out. Because as I understand it, ethics are guides and principles, and the legislation that comes are the sort of commandments of the society. Why is it important to develop, perhaps, some kind of ethical principles first, before that legislation should come?

LAVIZZARI: I think it's not necessarily a question of first and before and after, but human activity and also technology is simply so multifaceted that laws alone will never be able to capture entirely the attitude, ultimately, that people have to professionalism, people have to put into developing high-quality artificial intelligence tools. It's almost like the financial markets. You can have laws galore, but at the end of the day, the ethics of the financial advisors are secured through professional standards and ethical rules of those financial advisors. I think it's the same here.

KENNEALLY: Well, Niels Peter Thomas at Springer Nature, you confronted these issues, particularly the ethical issues, head-on with this project that was announced last spring, a book published – an e-book published – I think I have the title right – Lithium-Ion Batteries, a Machine-generated Summary of Current Research. It was research that – or at least the project, I should say, was conducted in partnership with Frankfurt Goethe University, so one that's very local as much as it is a global interest. And it was a project

that took, as I understand it, something like 18 months from start to publication. Were you thinking about these ethical issues? When did you begin thinking about them? And what were some of the questions that came up in your meetings?

THOMAS: Actually, Carlo was referring to the machine as the reader of the book. And we tried to look at – from the other side – the machine as a creator of knowledge, the machine as an author, as a writer of books. So we actually did it – so half a year ago, we published this book, and it was actually an idea that was born exactly two years ago at Frankfurt Book Fair when I was standing together with colleagues thinking, is that actually possible? We thought it through, and we thought it takes at least one and a half years, and in the end, it was exactly the time we needed.

But for us, it was – from the beginning, it was very clear that this is only half a very important challenge in terms of technology. How do we master to really, as you said, bring the data together, let the machine learn, produce something really meaningful? Because our ambition was that the chemistry community should say this is useful to us. We need it. So we want to read it. It must be a human-readable machine-written book. So that was our ambition.

And we thought from the beginning, we want to be first, because we believe it should be a publisher who does it first and not a tech giant, because – exactly, we said we want to make very transparent from the beginning and say this is exactly what the machine can do right now. This is the state of technology as of today. We need to start a discussion on the ethical aspects of it and the responsibilities. So we made it very transparent in the book what the limitations of the technology are. I mean, this is our first attempt, and the community gave us very feedback to it. But it's far from being perfect.

And now, we need to really think about this other half of the challenge, which is really a legal challenge – what happens if somebody infringes the copyright of this book? Whose rights actually are violated? These are very new questions that we have to ask. But then, also, who takes the responsibility? What if the machine summarizes something where we say this is inappropriate? How do we do it? But also very practical issues of what we know is in our industry the standard, like peer review. We can't go back to the machine and say please redo chapter three, we don't like it right now. It will be exactly the same outcome.

So we need to really negotiate new standards, and I believe we need to do this not as a single publisher, but need a new idea in the whole industry. So this is why I very much welcome the initiative here to bring this together, because I think these ethical and these responsibility standards that we need to develop are at least as big a challenge as the technological challenge is and was.

KENNEALLY: Right. It raises so many questions. One, though, is fundamental, which is what does it mean to be a publisher? The idea or the notion that one can be a publisher by pressing a button seems to be somehow contrary to what we all feel.

THOMAS: Yeah, so I think as the managing director for books for Springer Nature, it's a very compelling idea. If we don't need any authors anymore, maybe we don't need any editors anymore if we can do it all by the machine. But that, of course, will not happen.

So this is – as we very transparently say in the subtitle, it's a machine-generated summary of existing – of current research. So we are producing a new perspective, but we are not producing the – in a technical term, we are not producing new knowledge. But for readers who have no time to read all the 5,000 to 10,000 articles that are somehow summarized in such a book, if you don't have that much time – and most of us haven't – then it creates a new perspective. So it is kind of a new view and a bias-free summary of existing research, which is then almost, by the push of a button, to be created.

And that gives us a completely new role. That's true. And we have to think about do we want this role? Can we take this role? Should we? Or should we involve more experts, more authors to check it, to correct it? Should it be our responsibility? Is it the responsibility of a review of the community? I think these are very, very important questions, but there is no final answer to them yet.

KENNEALLY: Right. And player pianos and phonograph records didn't do away with concert pianists, either, so there is an example for us of where they add to – they are supplemental. And to the point about it, perhaps we need to tell people a bit more about the book. It's a collection of work. It has gathered information on, I believe, thousands of articles. Is that right? Tell us more.

THOMAS: Well, I cannot go too deep into the technology, because in the end, it's indeed quite complicated. But the basic idea is that we show to a pipeline of algorithms everything that we have published in a certain discipline in the last, let's say, five years or so. In this case, it was everything about lithium-ion batteries, everything about electrochemistry. So there are between 5,000 and 10,000 articles, book chapters, database entries, and so on that we showed to the algorithm. Then the algorithm clusters them and says there are certain areas where there is more published knowledge about, certain areas where we don't know much about it.

And then after this clustering, there is one human interaction. This is basically a figure that we give to the algorithm. We want to see this book in 250 pages, and we want four chapters. That's what humans need to say. And then the algorithm creates a table of

content by selecting the appropriate clusters and bringing them together to superclusters in order to say, what do we want to bring here together? So there is one chapter about anodes, one about cathodes, one about data models, and so on. So this is all done by the algorithm.

And then the algorithm looks at what is published in this cluster, looks at certain measures – what is important, what is less important, where do we have more knowledge according to specific parameters, what seems to be more important to the community and whatnot, and then extracts the most important facts and text parts of the original works – of course, properly cited, so my legal colleagues assured me that we are not infringing copyright here. But then it is all summarized by the machine.

And then there is some semantic parsing and some – and we then changed the size of the text. So we bring it together so that in the end, you have a text with some direct quotes, but mostly text or sentences that are not in the original works but are combined knowledge of the different sources of the content that we showed to the algorithm in the first place.

KENNEALLY: But I believe a librarian came to you and raised some alarm bells – or raised some red flags, I should say – regarding the potential here for abuse. Just tell us about that.

THOMAS: When we published it, it was really our intention to start a discussion on the consequences and to start a discussion on how do we control this technology in the end? Since then, we are discussing it with lots of librarians. And indeed, one librarian came to me and said it's really very fascinating. It's very interesting. I support the idea that this is a good idea to experiment with. But you can abuse this technology and somehow mask a plagiarized text, so you could indeed – you could machine-generate a text and then hand it in as your PhD thesis or something like that with only minor modifications. So there are possibilities. There are always possibilities to somehow misuse a technology.

So I think the question for me is here that even by using the technology in the intended way, we have to be very careful. But there is also this even more – potentially more dangerous aspect that the technology itself can be used for something that we don't want it to be used. But I'm an optimist. As long as we have sessions like this, I think – as long as we are transparent about it, as long as we care about these issues – I think we can solve them, so this is why it makes me an optimist.

KENNEALLY: Well, thank you. I really appreciate the discussion we've had with Carlo Scollo Lavizzari and Niels Peter Thomas. Thank you very much.

(applause)

END OF FILE