



## **Interview with Joris van Rossum**

**For podcast release  
Monday, April 19, 2021**

KENNEALLY: Technology forever changes our world, usually starting with science. Galileo's invention of the telescope sealed the fate of an anthropocentric universe. After the Janssens developed the microscope, van Leeuwenhoek improved the device enough to reveal microorganisms. And of course, computers like those Bill Hewlett and David Packard first developed in a Palo Alto garage are now found in every lab.

Welcome to Copyright Clearance Center's podcast series. I'm Christopher Kenneally for Velocity of Content. In our own time, robots and algorithms using artificial intelligence are becoming commonplace tools in research. AI is especially relevant where large volumes of data and information are processed, leading directly to the scholarly and scientific publishers that digest the data and produce even more of it.

Following 18 months' work, the STM Association will release a white paper, "Best Practice Principles for Ethical, Trustworthy, and Human-Centered AI" as part of the upcoming STM spring conference to be held online April 27th through the 29th. Joris van Rossum is STM's director of research integrity. He joins me now from Amsterdam with a special preview of the report. Welcome to Velocity of Content, Joris.

VAN ROSSUM: Thank you, Chris.

KENNEALLY: It's a pleasure to have an opportunity to speak with you about an issue that is of interest to researchers, to publishers, and indeed to the general public, as we see robots and artificial intelligence and algorithms so much a part of our own lives. When it comes to AI, I think what STM has put its finger on is that in order for it to work in the way we all hope it will in research in science and technology, it has to be grounded in values and principles. Tell us about those.

VAN ROSSUM: Yeah, indeed. Let me first start by saying that indeed, the potential for AI in science and research is huge. As you mentioned, technology always has played a big role in science. It's changed science from being observational to more experimental. As you said, the computer – we can't think of research anymore without a computer. And the enormous production of data in the last decades, in combination with AI technology, really, really, promises to make science way more efficient. Some speak of so-called smart science, meaning that AI is not just testing hypotheses against vast amounts of data, but also creates new



hypotheses, developing new theories, exploring new connections, and determining unknown causes. But for that to happen, having sound principles and making sure that AI is applied in an ethical, trustworthy, and human-centric way is really crucial, as of course, those are principles central to science itself.

KENNEALLY: Indeed, they are. And the difference, as you point out, with AI is that it is more than just gathering the data, because the algorithms themselves can be made to make predictions, recommendations, even decisions about research or about other activities. So AI is data with a difference, we could say.

VAN ROSSUM: Absolutely. And I think the advantage, indeed – or the promise – is that it can do a lot with data. The disadvantage is that it's quite obscure. The algorithms are quite obscure. Which means that we as publishers have to be very transparent – have to be transparent when it's used, how it's used, etc. So with that challenging nature of the technology also comes our responsibility to be transparent about how it's used, etc., and to make sure it's not misused or doesn't lead to consequences that are unintended and unwanted for.

KENNEALLY: STM publishers – scientific, technical, medical publishers – they are both users and producers of data. For our listeners, Joris, tell us a bit about how those activities happen. How are publishers using AI today in their work?

VAN ROSSUM: Yeah, I think when we developed this report, we realized what a unique position we as publishers have. First of all, we are key providers of information and data and articles on which AI is run as input data and as training data. Of course, we are really well designed to do so, having so much experience in curating data, selecting data, reviewing data, etc. So we are in a unique position to really provide high-quality data. And as we all know, garbage in, garbage out – having the right data, having high-quality data is really crucial for an efficient and a trustworthy application of AI.

But second, publishers are also increasingly using AI, either developed in house or supplied by third parties, to support internal workflows and services for authors, editors, and reviewers. That has actually been going on for quite a while. AI is being used in recommending journals to authors – uploading your abstract, and you get the right journals you should submit it to, recommending reviewers to editors, recommending content to readers, streamlining submissions by carrying out technical or language checks, and last but not least, helping authors to improve their English by means of artificial intelligence support tools. Another important element is plagiarism detection, of course, now being used ubiquitously by almost all publishers. Also, we're investigating how we can use AI to prevent fraud and data manipulation, for example.



There's actually a third way where we use AI, and that's in external-facing tools and services – using it to classify content, to recommend data to external users, and bringing together related information from disparate sources. Fueled by AI, publishers are also increasing serving as providers of analytics and insights – for example, insight into research trends or as input for R&D and identifying targets for drug development.

So again, it is quite a unique position we have dealing with AI. And I should maybe add a fourth element – of course, AI as an area of research. We support that research by, of course, publications – articles and books around the subject. So quite a close relation we have with that field.

KENNEALLY: It's a close relationship and a complex one, and you've pointed out the various ways that it is. So in order to arrive at some best practice principles, it has to gather up a lot of important considerations. That was part of the work that you and the group working on this STM Association white paper engaged in. You've come up with five categories that you identify as being best practice principles for ethical and trustworthy AI. Tell us about those five categories.

VAN ROSSUM: Yeah, it's interesting that more and more of these principles are developed. For example, the European Union has provided an overall view of best practice principles, the OECD as well. And the themes, I would say, overlap. For example, transparency and accountability is an important area where we define principles, quality and integrity, privacy and security, fairness and sustainable development. Everybody is thinking about these important elements of trustworthy and ethical AI. But as I mentioned before, we as publishers I think bring specific perspectives to this, and hence also our work on this white paper.

KENNEALLY: So following from transparency and accountability, the second principle is quality and integrity.

VAN ROSSUM: Indeed, yeah. That's the core of what we do, of course. One important element I already talked about – that is, the data. So how do you make sure that the data that is used for AI is the right selected content? That is very important. Having wrong training data, having wrong input data can lead to a lot of unwanted consequences. We all know about racial biases, etc., that can arise because of AI. Here, the publishers, of course, have an important role to play through curation, through ontologies, through peer review, but also through the creation of databases with selected content. We have the unique position to really ensure that people can select the right content for input.

But of course, quality relates to more of what we do. I think about the peer review process, for example – typically, a process where we ensure that the quality of the published material is ensured. But also we can apply this peer review process – or, let's say, the quality



cycle – to the use of AI itself as well. So I would say there are many levels where we can benefit on the work that academic publishers have done over the years.

KENNEALLY: Right. And we talk about data, Joris van Rossum, and it has this neutral quality to it. Yet it is, as you point out, far from always neutral. Data can be marshaled to all sorts of purposes, and that is because human beings are involved. There are human beings collecting the data. There is the data that the human beings themselves are producing. So questions of privacy, even security arise.

VAN ROSSUM: Yes, absolutely. Privacy and security – of course, it's new. It's actually becoming the subject of new legislation around the world. Again, hence also very important for AI. So what we have developed is 10 best practice principles and operational steps to ensure respect for privacy and data protection when designing, developing, or using AI systems for data used and generated by the AI system throughout the lifecycle. In this white paper, we developed these 10 principles that, indeed, ensure that privacy is secured, a crucial element of a trustworthy AI as well.

KENNEALLY: Fairness comes up as one of the principles here, one of the five best practice principles. What do we mean by fairness? What kind of fairness are we thinking about – fairness in terms of equity? What other concerns?

VAN ROSSUM: Yeah, this is a very interesting aspect. I think to understand fairness in relation to AI, we have to, again, remember what AI does. AI basically strengthens existing patterns. It looks at data, it looks at the past, and based on that, it makes recommendations or it makes predictions. But that means that existing patterns can be strengthened. That entails a risk that we have to be aware of – and hence also the principles to try to counter those risks.

Let's say if I am an author from a well-known institution, and I submit a paper to a journal. Let's assume that a journal uses AI to predict whether my manuscript is of high quality. What the AI then does is it looks on aspects like where do I come from, from what institution, from what country? If my colleagues from the same institution or my same country have been successful, AI can make the conclusion, oh, that's probably an author that will be successful. That is, of course, very, very risky, because then it strengthens the patterns, which means that people don't get a fair chance. If I come from a country which has traditionally not been very successful, it's going to be more difficult for me to get published. That is something we absolutely have to prevent as publishers. That's why we really have to think about what data we use, but also in what processes we apply AI.

But also, it has, I would say, a deeper risk. That is that if you look at it more broadly, AI tends to consolidate historical structures, which includes established scientific ideas and theories. But as the philosopher Thomas Kuhn has argued, scientific breakthroughs are characterized by



replacing paradigms with new ones. You used Galileo yourself in the introduction. I think that's a very good example. So the risk of using a technology that looks at existing patterns to make predictions, recommendations, or decisions means that it can suppress the opportunity for new ideas to emerge, thereby stifling innovations and scientific breakthroughs. Of course, as partners for science and for scientists, this is something we as publishers have to prevent. Hence we have to be really careful, which is described in the principles, when and how we apply this technology in a publishing process.

KENNEALLY: AI, artificial intelligence, it's characteristic of our century, of our 21st century. As well as characteristic of this time is concerns around sustainability and the proper way to develop economically and the use of resources. How does that factor in? How does the final consideration of best practice principles, sustainable development, factor into AI considerations?

VAN ROSSUM: Yeah, I would say, indeed, as you say, AI has a crucial role to play in maximizing – we have a role to play to maximize the benefit of AI to human health. Of course, that's the core of what publishers do. We improve science, and thereby we hope we improve the society at large. Again, this fits, I would say, in the general goal of science, and hence of publishers.

Very important, however, is that policymakers and research funders should create incentives for providers of key input data. Having high-quality data, having curated data – again, such a crucial component of trustworthy and ethical AI – it really depends on quality data, and that means we need incentives. We need incentives for people to produce the data, to curate the data, to store the data. That's also part of the recommendations we put down in the white paper.

Next to that, of course, we have to make sure that the AI is used efficiently, preferably with renewable energy, and it should be sustainable. But again, for that to happen, we need incentives to ensure that we keep making the investments crucial for correct application of artificial intelligence.

KENNEALLY: Well, I have to compliment you, Joris van Rossum, because your explanation of this white paper, “Best Practice Principles for Ethical, Trustworthy, and Human-Centric AI,” has been exceedingly clear to me, and I am hardly an expert in this field. I think that's the trouble, isn't it, finally – that the subject of AI, we all hear about it, we all recognize it, but there is a great deal of opacity to it all. It is such a complex field. Understanding of it is really challenging, but it's critical moving forward.

VAN ROSSUM: Absolutely. Again, let's indeed end where we started with. The potential is huge. If you think about what machines can do and could do plowing through enormous, vast amounts of data, coming up with new hypotheses, doing so much more than human beings can



do themselves – we like the saying, creative individuals and smart computers. It should support researchers. It shouldn't, of course, take over the entire process.

But as you say, it's opaque. It's difficult. That's why we need to think carefully how and where we apply it. But also, we need feedback loops. We need to work with the users. We need to work with all the various actors in the ecosystem to make sure that once things happen that we don't intend or that, again, doesn't contribute positively to science in general, we need feedback loops where that can be addressed and corrected. So it's going to be a continuous development.

AI, of course, is also in development. It's not a settled technology, and it's expected to evolve significantly. Hence, also, our principles are not intended to be exhaustive, but rather to contribute to the ongoing discussion and to make sure that we move forward, but move forward in a responsible way, taking into account all the benefits, all the promises, but also the risks, and doing that, I would say, with all the participants in our ecosystem.

KENNEALLY: Joris van Rossum, director of research integrity for the STM Association, thank you for joining me today on Velocity of Content.

VAN ROSSUM: Thank you.

KENNEALLY: Our co-producer and recording engineer is Jeremy Brieske of Burst Marketing. You can subscribe to this program wherever you go for podcasts and follow us on Twitter and Facebook. I'm Christopher Kenneally. Thanks for listening. Join us again soon for another Velocity of Content podcast from CCC.

END OF FILE