

Karmen Williams:

This episode is part of a special collaboration between ACM ByteCast and AMIA: For Your Informatics Podcast, a joint podcast series for the Association of Computing Machinery, the world's largest educational, and scientific computing society, and the American Medical Informatics Association, the world's largest medical informatics community.

Sabrina Hsueh:

In this new series, we talk to women leaders, researchers, practitioners, and innovators who are at the intersection of computing research and practice. To apply AI to healthcare, and life science, they share their experiences in their interdisciplinary career paths, the lessons learned for health equity, and their own visions for the future of computing.

Hey, hello, and welcome to the ACM-AMIA Joint Podcast series. This joint podcast series aims to explore the interdisciplinary field of medical informatics, where both the practitioners of AI/ML solution builders, and the stakeholders in the healthcare ecosystem take interest. I'm Dr. Sabrina Hsueh from the Association of Computing Machinery, ByteCast series. And co-hosting with me today is Dr. Karmen Williams from Your Informatics Podcast with the American Medical Informatics Association. We have the pleasure of speaking with our guest today, Dr. Francesca Rossi.

Francesca Rossi:

Thanks. Thanks for having me.

Karmen Williams:

Thank you for joining us. Francesca Rossi is an IBM fellow, and an IBM AI ethics global leader. She works at TJ Watson IBM Research Lab in New York. Her research interests focus on artificial intelligence, and ethical issues in the development, and behavior of AI systems, in particular for decision support systems for group decision making. Dr. Rossi has published over 200 scientific articles in journals and conference proceedings, volume between conference proceedings, collections of contributions, special issues of journals, and a handbook. She's a fellow of both the Worldwide Association of AI, AAAI, and the European one, which is EurAI. She has been the president of the International Joint Conference on AI, and Executive Counselor of AAAI, and the editor-in-chief of the Journal of AI Research. Dr. Rossi is involved in multiple AI committees both in the US and Europe, and we are so excited to have the chance to speak with her more today. So thank you again for joining us.

Francesca Rossi:

Thanks.

Sabrina Hsueh:

Yeah. So, Dr. Rossi, you have been working in this interdisciplinary field between computer science and ethics. What makes you decide to start your career on this path initially? Did anything particular motivate you that you will call out as [inaudible 00:03:00]?

Francesca Rossi:

Sure. Thank you. Well, first of all, two things. When I decided to study AI was... Well, first of all, I decided to study computer science. That was the first decision that brought me on this path. And I think it was because I felt that at that time there were not many in Italy, I was in Italy, there were not many

computer science curricula in the various universities. So I felt it was something for the future that I was doing something new, and that was exciting for me. So that's what computer science. Then in my master thesis, I decided to do it on AI, and that's where I started working on AI. And I continued, and I still am working on AI. And that also was an additional thing. It was an additional feeling that really this science, more a technology, but also the science of AI was really allowing me to build things where they were new, they were shaping the future.

So even more than when I chose computer science, and then, after many, many years working in AI, and teaching, and doing research with my students, and so on, I was a university professor teaching, and researching in AI for more than 20 something years. Then I went one year on sabbatical at the Radcliffe Institute for Advanced Studies at Harvard University. And that place is a really a crash course in multidisciplinary because during the year that people spend there, every year there are about 50 fellows at the Radcliffe Institute, and they all come from different disciplines. So I was the only computer scientist there. And then, there were people covering all the other sciences, all the arts, and all the humanities. And then, the staff of this institute forces these people to work together to spend time together. I say forces because it doesn't come very natural at first because these people are used to work only with the people similar to them.

But really it was very, very learning experience, and in communicating, and working with people that had different background, different questions in their mind. And that was where I started to think not only about the technical, and scientific aspects of AI, but also about the impact on society. And that's why when I started thinking about the ethics of AI and the impact on people, on society, the way we live, the way we interact, the way we work. So after that sabbatical year, instead of going back to my university professor role, I decided to join IBM. And then, since then I increased inside the company more, and more this activity around the ethics of AI. So that was a pivotal moment, I think, because I learned what it means to think about the social aspect rather than just about the technological, and scientific aspects of AI.

Karmen Williams:

Thank you. And you talked about some of those interdisciplinary challenges such as foreseeing, everybody had to force together. And so were there any other challenges or how did you overcome this challenge?

Francesca Rossi:

Well, one initial challenge, if you want to call it challenge, is that these people that come from different background use different terms, even different terminology, different words. So it's already an initial phase where you need to be able to synchronize, and do some translation between the languages in order to be able to then being effective in working together. Because if you don't even understand each other, and I found this during that sabbatical year, but then I found this challenge also later on. For example, for two years in the European Commission, high level as per group on AI, this was a group that worked on 2018 to 2020 for two years to define the ethics guidelines for AI in Europe. And these 52 people were very different. They were experts of AI, but less than 10. And then, there were civil society organizations, consumer rights organizations, philosophers, psychologists, and all sociologists, and so on.

And so the mandate of that group was to write these ethics guidelines for AI Europe. We realized that the first meeting that this 52 people basically had 52 ideas of what AI was. So the first thing that we needed to do is actually publish an additional document that really was saying, this is what AI means after discussing among these 52 people, and then, we could start thinking about the ethics guidelines for

AI Europe. So that was the way we solved the challenge there. But the challenge was also within IBM, because when we started building the governance around AI ethics, and involving all the divisions of the company in the activities around the AI ethics, of course, these people were very different in different variable, AI researchers, there were people in marketing communication, there were the legal people, there were the people doing products, software. And so all very different people. So the first thing we did was to build a glossary, a glossary of terms around AI and AI ethics and with their definition so that everybody could look up at the glossary in case there was any doubt.

And that was the agreement that these different people had reached about the terminology. So the terminology phase is initial one, of course, it can be evolved because terms are added, and so on. But that is something that needs to be solved, otherwise, you cannot work well together in an interdisciplinary environment.

Sabrina Hsueh:

Yeah, I can totally resonate with what you say here. It sounds a lot like what we are experiencing in our own company as well. But I want to take it back a little bit to also understand. So this kind of translation between interdisciplinary field is certainly difficult, and challenge, but you manage to do it well. Is that what lead you to your current role? Or is there anything else you will say that inspires you to pursue the current career path?

Francesca Rossi:

Yeah. Well, when I joined IBM, it was also a moment in time where there was a lot of initial discussion about these issues with AI, the first algorithms that were shown to be making discriminations, for example, also. So there were the first discussion coming up, what does it mean that this technology is beneficial to people, to society and so on. So with that in mind, and I was fortunate enough that during my sabbatical year before joining IBM, I joined the advisory board of the Future of Life Institute. That was very pioneering in convening these discussions about the ethics of AI. So then when I joined IBM, I started also a multi-division discussion group around these issues. But it was very kind of all over the place at the beginning. And then, this discussion group was transformed into a first version of the AI ethics board inside the company.

So an initial version of the governance around AI ethics, where we defined our principles, we defined some partnerships with external organizations, we defined some activities, but then it was not a decision-making board, it was mostly a coordination, and discussion board. And then, a few years later, it was transformed into a really decision-making governance body, which makes decisions about not just the principles of high-level thing, but also about the concrete activities that the company. So it was a natural, I think, phases one after the other one that led to more, and more concrete from very high-level principles to very concrete actions.

Karmen Williams:

Yeah. And I think you've already talked about this quite a bit, but we know that many industries and companies are arduous. They have established, they're responsible, and ethical AI guidelines. We see the new kind of coming in of generative AI, and the additional risk that could come with that. And so what do you see as the current status in enabling ethical and responsible AI? And then, if you're going to score for progress collectively as a field, how would you rate it?

Francesca Rossi:

Yeah, so as I said, these phases happen not only in my personal experience, but I think in many other organizations. So the first phase was mostly about being aware of the issues that were with the uses of AI. And then,, the second phase was principles. Everybody published principles. I think that Harvard was a project called the Principled AI project, and collected all the principles that were around AI ethics that were published by any kind of organization. There were more than a hundred sets of principles, many of them overlapping, but from different angles from companies, governments, universities, multi-stakeholder organizations, so on. So first phase awareness, second phase principles, and then, the more practical phase. Okay, so how do we implement this principle? How do we integrate the implementation of this principle inside our own organization?

And I think that these three phases, and we are still then in the practical phase, implementing the principles into concrete actions, whether these actions could be like risk case assessments, processes or educational material for everybody, the developers and everybody else, or software tools or even diversity in the composition of the teams or the governance bodies like the AI sport that we have. So all these pieces of very concrete actions came out of these principles that were very high level, and principles are very nice. They set like a north star, but developers need something more concrete than the principles. And of course, as you say, with the evolution of the technology with machine learning and then, generative AI and so on, these actions need to be also updated over time. That's why, for example, recently we published a very large table of foundation model risks and comparing them to the risks that we already were aware of for traditional kind of machine learning approaches to see which ones remain the same with generative AI, which ones are amplified and which ones are completely new.

And so they need new tools or updates on new tools, new educational material, new consultations, and so on. So I think that it's like I saw a continuous trend in increasing these activities, also increasing the number of people that work on this inside the company as well as the resources. And not only increasing, but increasing, not in a linear way, but kind of an exponential way at the beginning was maybe going a little bit slower [inaudible 00:15:08] and then, [inaudible 00:15:10] it grows exponentially very rapidly. Also, because there are forces that push in this direction that come, for example, for the point of view of IBM come from, yes, governments that are more are increasingly generating laws about AI, but also client themselves that want to know what we do for their solution that we provide and what we do about these AI ethics issues. So governments, clients, media attention, so a lot of internal but also external forces that push and make this area grow.

Sabrina Hsueh:

Yeah, so there are so many areas we can grow here, and certainly we never have enough people on this. We see the need for this everywhere we go. With so many AI regulations coming up on the horizon as well, certainly a lot of people are being kept awake at night. We're wondering that for you as particular in your capacity as the current president of AAAI, right? So who are those issues that keep you awake at night? And are there particular regulation or particular risks that you worried most? And also the follow-up question for that will be for professional societies like us, AAAI, like ACM, what can we do together upon together?

Francesca Rossi:

So as we said, the technology evolves, and brings about, and the uses of these new techniques in AI that bring about additional opportunities, but also additional questions and legitimate questions about possible risks in the uses. So for example, for generative AI, for classical machine learning, of course, there are the usual issues related to fairness. So the presence of bias, so the ability to detect and mitigate bias in an AI system, in the training data, and so on. But also sometimes the lack of

explainability. So it's kind of a black box, the system that produces an output, but we don't know how it gets the output from the input, the transparency issue that where those that produce the system need to be transparent enough to say how they built it, the privacy issue, of course, because machine learning relies on data. So all the data issues that including the privacy ones.

And then, with generative AI, of course, the issues related to the content that is generated. So issues related to misinformation, because we know that generative AI is extremely excellent in the writing text or generating images that are very realistic, and even text very perfect from the syntactic point of view, but sometimes it makes some mistakes, factual mistakes. So we need to be aware of that if we want to use it. The fact that it makes these mistakes, it doesn't mean that it's not usable, and not cannot be useful in many scenarios, but we need to be aware of that in order to use it in the most appropriate way. So if you ask for example, for factual information, it's better to check, double check the factual information with other trusted sources, for example. And then, of course, there are other issues related to misinformation, but they're not related to limitations or mistakes of the AI system, but actually on how humans decide intentionally to use these AI systems such as deep fakes, and that can impact our society by manipulating public opinion.

And so possibly disrupting even the democratic process, the elections, the candidates, and so the impact on society as well as the impact of these AI systems in the education system, for example, we know the students are using some of the systems to bypass the learning process to write their own essays. And that's something that should be resolved even in the research community. And I see that at AAAI, there are people that write some parts of the papers through these technologies or even parts of the reviews of the papers through this technology. So we need to be aware of possible actors they may use inappropriately these tools because the tools should be used, but should be used in the right way. So is to augment our own capabilities to help us grow, and do things that are higher level than what we could do earlier, and not to bypass any process that can be helpful for us to grow.

So again, so there are several issues, but most of the issues are, in my view, related to two things. One, that the technology still has limitation, as capabilities and limitation. So the typical case is this hallucination thing. This is a limitation of the technology. Maybe in the future by building the technology in different ways, maybe we'll overcome this thing, okay? But for now we have to be aware that there is a limitation. And then,, other risks are related to the risk of we decide to use the technology. So whether we decide to use it in an completely fully autonomous way when it's not the right scenarios to use it full autonomously. For example, and we need to have human oversight or human in the loop. So these are the two things, limitations of the technology and misuse, I would say.

Sabrina Hsueh:

Yeah. So are there things you feel that professional society can do a little bit more on top of [inaudible 00:21:16]?

Francesca Rossi:

Yeah. Well, professional scientific societies like ACM, AAAI, and AMIA, they can totally raise the awareness. So for example, even in the mission of AAAI, the mission of AAAI is to advance AI research, and the responsible use of AI. So it's already in the mission since the start. That is not only to advance the technology per se, that's not the ultimate goal, but the ultimate goal is to make sure that it's used responsibly. So awareness also, another thing that AAAI is doing is also to consult with the policy makers, so to help policy makers understand, and this can come from scientific professional society, but it can also come from companies from others that are more technically involved. So to help policy makers understand what is the best way to regulate AI, if there is a need to regulate it and how.

So that's why I think what the European Commission did with European AI Act was very wise, because it really started with these multi-stakeholder consultations where people from different areas, including professional scientific societies, could say, "Okay, I think from my expertise, from my knowledge of AI, I think this should be advice about how to regulate AI." So that's another areas where this association can do. But of course, being association that organized events like conferences, like you mentioned several conferences, AAAI has conferences, symposia, and many other things. I think that to raise the awareness of the researchers themselves, that AI is not a science and technology only at this point, but is a social scientific technological field. So to help researchers be exposed also to the social aspect, and considerations that they need to make whenever they do their research. For example, at AAAI, whenever people submit a paper to the conference, we ask them to write a section in the paper to discuss the possible ethical implications of the work that they're doing.

So that is something to raise awareness in research, because the next of researchers will have to be more multidisciplinary, more social-technical than what we had in the past. Then the way I was educated, for example.

Sabrina Hsueh:

Yeah, totally with you. In AMIA, we have been hosting the AMIA Showcase for three years in a row, right? And we do that too, to as all the submitters of the EI evaluation for their ethical consideration to be added in different stages of their health AI evaluation. So yes, thank you for bringing that up, the underlying important point, and also for that multi-stakeholder view. One thing that I think both ACM and AMIA has been doing is to bring awareness, and bring more multiple parties on the table together to generate consensus from community first, right?

Francesca Rossi:

So another thing that actually AAAI is doing together with ACM is to co-organize a conference called AI Ethics and Society, which is exactly this attempt to put together in the same place, in the same event, in the same conference, researchers, and scientists from different disciplines. So this conference, for example, has four program chairs, one from AI, one from philosophy, one from psychology, one from policy. And then the papers, posters, and invited talks, and so on are all very multidisciplinary. So that was a way for ACM and AAAI to really get together in this multidisciplinary way, not just as a technology associations.

Sabrina Hsueh:

Right, yeah. And in this multi-stakeholder view, did you find any particular party's voice is lacking? At least in AMIA, what we find is a patient's voice is often missing on the table, right? That's why we need to bring actual effort to bring them in. Like last year we had this conference with Harvard Medical School to make sure that we had the patient advocates also in that [inaudible 00:25:54]-

Francesca Rossi:

Yeah, that's a good point. Of course, in that case, it's patients, but in general is the communities that are impacted by whatever these researchers or developers are building, right? And this is true that usually this voice is not too present. It may be present, for example, I've seen even at this conference, AI Ethics Society that happens every year. Last time, I've seen a lot of papers describing studies, for example, where they were asking. Even some of them were related to healthcare, they were asking the impact of AI on the doctors, on the patients, and so on. But the conference itself did not have these communities present there. Some of the papers were describing the results of these surveys, of these consultations,

but it's true that those communities themselves were not present at the conference. So they were present only because they were in the papers as being part of the study described in the papers.

So you are right that most of the people at this conference, for example, are those that produce the technology or produce the solution based on the technologies, and then, they also do their own study at home, and then, they come to the conference, and present a result of the study. But we don't see a lot of that community that is impacted by the work.

Speaker 4:

ACM ByteCast and AMIA, FYI Podcasts are available on Apple Podcasts, Google Podcasts, Spotify, and other services.

Karmen Williams:

If you're enjoying this episode, please subscribe, and leave us a review on your favorite platform.

Wonderful, yeah. And then, in your various roles, and very much so a pioneer in this area, are there any changes you are proud that you've been able to facilitate? And then, for that matter, are there any new changes that you would suggest?

Francesca Rossi:

So within this journey, even within IBM that we did inside the company, I think the idea of building this internal board, that's something that I really feel that it was the right thing to do. And in fact, that was started all these very concrete activities, which I did not do by myself, of course, with a lot of other people, and the companies, and other division putting resources, and so on. But this initial idea of the need for centralized governance inside such a big and global organization like this company, I think it was the right thing to do, and to pitch, and to be insisting at the beginning inside the company to build it. And because again, in a global company, it's important that this governance is centralized, although it's very connected to all the divisions, and to all the different roles. But it's important that it's centralized. Because again, we cannot have different standards in different countries.

A company that has offices everywhere in the world, basically you want to have the same principles and the same standards of risk assessment, evaluation of or use of the tools, thresholds for bias, whatever, all over whatever the company does, all over the world, independently of whether there is a law that requires that in that region of the world or not. And that's already sets a very nice attitude that AI ethics is not about just complying to some law. So waiting for the law, and then, when the law is there, you have to comply, of course, but really being very proactive, and saying, this is what I want to do. This is how I want to behave with this technology. This is what I want to do no matter whether there is a law that forces me to do that or not. And that I think is implicit in this initial idea that I had of having this centralized governance, like the AI ethics board.

Sabrina Hsueh:

And there is a new generation of informatics computer scientists who will be listening to this podcast episodes who will be interesting in diving more with you on this area. Are there any suggestions for where they should start and where [inaudible 00:30:32] ?

Francesca Rossi:

Well, first of all, at this point, there are many books on AI, and also AI ethics. There are really many books that tackle these issues from different angles. So I mean, there are really many, many different

books so that one can read. But even just going to these conferences like the one that I mentioned, the AI, Ethics and Society, and because at these conferences, people then can get an idea that you walk around, you talk with people. Usually in this conference there are paper presentations invited [inaudible 00:31:07]. But then also there are poster sessions. So in the poster session, you can go up to the authors, ask, discuss, so you get an idea of what people are doing, and what are the main issues that are being addressed, what are the main solutions that people put in place.

And so that is a nice way to be engaging with that community of people, that multidisciplinary community that thinks about these issues. And as concrete solutions puts in place concrete solution for the AI ethics, but also that you find also companies at these conferences at AAAI, for example, they're usually not companies. So if you have in mind to change your career or do something, and you get to see the internship opportunities or even the full position opportunities that are around that area.

Karmen Williams:

Yeah. And you've given such good, really good insight on this. And so stepping back a little bit, reflecting on your own career, what might be some of those things that early career moves that you made that you... When thinking back, you're like, "Oh, this will be really useful, I could recommend this for newcomers." And then, for that matter, even mid-career, what were some of the things that you did? So what would be your advice for those who are exploring maybe different career paths in this area?

Francesca Rossi:

So again, my early career, it was not very multidisciplinary. Even mid-career, it was not multidisciplinary because I had more than 20 years in academia where I was doing AI and I was only talking to AI people. And that, of course, is very much nice because you are in your comfort zone. So you are there, you talk to people that completely understand you. So it's very comfortable being in that position. But then when I decided to go a little bit out of my comfort zone with this sabbatical here at the Radcliffe Institute, that was really an eye-opening thing. And so it was maybe more challenging than remaining in my comfort zone, of course. But it was really an eye-opening, and then, not only eye-opening to me, but also opening a lot more possibilities, and opportunities to me that were not in the landscape before that.

So the only suggestion that I can give is really to not be afraid to experiment, and to go out of the comfort zone, and to work or collaborate with people that are very different. Because again, maybe a bit challenging at first, but then it's really very rewarding, and you really feel that you can have an impact outside your vertical. And so you can really, I mean, have an impact also on other disciplines or learn also a lot because if you keep talking to the same people or the same kind of people, yeah, you learn because, but you can learn much more if you start talking, and working with people that are very different from you. So that was the... And of course, in my case, I did not do that for many, many years, but then when I decided to do it, I said, "Oh my God, why didn't I do it earlier?"

But it was the right moment in time. But again, I think that not necessarily like I did it, that I moved from academia to a company from Europe to US. I mean, it doesn't have to be that drastic change, but to really go out, and be aware of the fact that even if you are working in a technology, the right stakeholders are not just technological people. There are many stakeholders of that technology that are outside the technology, and they are societal stakeholders. And so the best way to advance the technology, and the science is to consult with these other stakeholders.

Sabrina Hsueh:



And that's one of the reason we started this podcast series, that we want to start exposing our new generation to this kind of interdisciplinary field. So they can learn how to speak the language differently, and to be more brave as what you have experienced to step out of your comfort zone. You are-

Francesca Rossi:

Yeah. So I remember for example, that until several years ago, for example, at the AAAI conference or some other AI conference, it was okay that some author presenting his own paper, technological paper about AI. I remember that people were asking, "Oh, but what about the societal impact?" And then, I say, "Oh, I'm just a researcher in AI. Somebody else will take care of the societal impact." So this maybe was okay many years ago. Now, in the last 10 years is clear that this answer is not okay anymore. It is not acceptable. So you are a researcher, you do the technology, you do the science, but you need to be aware of the societal potential, societal impact, and be able to answer the question that related to the societal impact. So that's an evolution that really is very important.

Sabrina Hsueh:

And that leads us back to our next question, which coming back to our discussion about professional societies a little bit, right. So in this intersection, we definitely all think experiencing the need of talking to people in different field, in AAAI, in ACM, in AMIA. So are there particular activities that you feel that will have great potential for this different professional society that this thing to be touching on different audiences, but in the end of the day might be really the same group of people in the future, is how do we for together?

Francesca Rossi:

Yeah. No, I think that there are many things that can be done together, joint events, not necessarily full conferences, but even joint events, parts of the conference, for example, AAAI as an annual conference with more than 4 or 5,000 people. But then it has also spring and fall and summer symposium, which are kind workshop like much more informal and smaller events. So those could be joint. Also within the AAAI conference, there is a program that we started few years ago that is called the Bridge Program that is exactly to build bridges between AI and other communities, AI and the vertical sector, AI and something else. So that could also be a place where it could be joint. One of the bridge, let's say events could be about AI, and ACM, and AMIA doing something together related to health informatics or anything related to that. So there are many activities that can be done, for example, that can be even joint issues. For example, AAAI has a magazine called the AI Magazine that is available to everybody, also non AAAI members.

And this magazine has issues with various articles. And each issue as a team, so there could be a team that is this joint interest around AI, and healthcare or other topics. So from the point of view of AAAI, I think there are many things that can be done jointly by these professional and scientific associations.

Karmen Williams:

Yeah, that's very helpful. I think even thinking how AI we know is not new to computer scientists, definitely in natural language processing. And so we've seen it tried it over and over again maybe for the last 40 years. But lately we see there's maturity of infrastructure support to large scale computing. And even the availability of pre-trained models definitely has increased the access to things that we previously were maybe research only technology. So this has led to many real world applications, definitely in healthcare as well. And so in your mind, what is the most important AI application thus far? And then, what do you see that's coming up that we're like, this is really important?

Francesca Rossi:

So first of all, maybe I'm sure you are much more experienced than me in understanding what's the best application so far of AI in healthcare, but I would say anything that requires the analysis, and the interpretation of large amounts of data, this is what AI and machine learning techniques are very good at. In again, supporting the activities of the health professionals in making decisions based on a more sophisticated analysis of the data. So whether it's radiology or whether it's something else, whenever there is a lot of data to be analyzed, I think that AI can help. Again, being careful about what should be automated, and what should be said given as knowledge to the doctors, to the health professionals so they can make more informed decisions. But for the future, I think that healthcare is still done in a very siloed, vertical way, just like education. In some sense, education, at least when I studied, and I think in Italy still when people studied the university, the other study informatics or philosophy or psychology, I mean, they're very vertical things.

And I think that the way the health of a person is still handled is a bit like vertical. Now you have a specialist in this area, and then, especially this other area, and especially in this other area. So I think that AI can help to have a more holistic approach to healthcare that I have the impression that could help in higher success of resolution of issues in the health of a person. So a more holistic approach where data coming from different sectors, disciplines within the health of a person can be combined to get a better idea of what can be done about the health of a person. And then, of course, there are all sorts of even preventive closed loops or closed loops approaches where AI can be combined with neuro technologies, for example, that can read data from sensors, but can also write data into the central nervous system.

I think for example, some things that are already available, maybe only in the labs or few situations like those closed loops where for patients to anticipate, and mitigate epileptic seizures, for example, now the data collected AI that interpret them predicts that there will be a seizure, and then, the neuro technology that injects some substance to mitigate that, or similar one for treating or mitigating Parkinson tremor. So this closed loop where AI is combined with other technologies, for example, neuro technologies to help in a very ongoing, always ongoing sensory information analysis, and then, intervention. So that to me is a very interesting avenue. The future can be expanded from these two examples that I gave.

Sabrina Hsueh:

And whether it's healthcare or not. So this then has to be evaluated before they can put into the real world scenarios-

Francesca Rossi:

Of course.

Sabrina Hsueh:

... and really to be practically used. So I'm wondering, in terms of evaluation, did you see that any important pieces there that should be included in your mind? And since there are so many guidelines out there, so sometimes [inaudible 00:43:49]-

Francesca Rossi:

Yeah, of course, the more we are using these technologies in making decisions that are affecting people's life, like healthcare, financial, well-being a very important aspect of the overall well-being of a

person, and the more we have to be careful about the technologies that we employ, we have to vet them, we have to evaluate them over common benchmarks. So for example, that are related to the scenario where you want to deploy these technologies. The risk assessment framework, for example, NIST in the US has developed a Risk Assessment Framework in general, not just for specific sectors for AI. And this is something that can be standardized to agree on what is the right way to evaluate whether it's benchmark, whether it's standards, whether it is some thresholds in the risk assessment. So the evaluation... Or whether it is internal or external red teaming. There is a lot of discussion right now about red teaming AI, large language models, and so on. And think that these agreeing on what are the right benchmarks, what are the right ways to evaluate is very important.

Many years ago in AI, this was not some very common not to have an agreement. The first time that researchers in AI agreed on some benchmarks to evaluate AI system on a common set of benchmark is when the ImageNet database was put together. Because for the first time you say, "Oh, okay, everybody working on image interpretation should tell me how the system behaves on that same data set of images." Then, of course, we discovered that there were issues with this data set, but the point of having a common benchmark, a common data set to use to evaluate, that's very important

Sabrina Hsueh:

And that set up the success what Transformer, and many now technology transforming invention that we have seen in this field.

Francesca Rossi:

So there are many... So, of course that ImageNet is less used now, but there are more, and more benchmarks to evaluate the capabilities of this AI system. Of course, we have to be kind of careful to evaluate over the right benchmarks. For example, I've seen a lot of evaluation of large language models over benchmarks that we use to evaluate human beings. For example, the bar exam or the admission exam. But of course, those are fought for human beings for the capabilities of the human brain and not for machines. So for example, machines are very different in capabilities. For example, they have an almost infinite memory because I can add as many things I want and almost perfect memory. This is something that we don't have as human beings. So I don't think it's very appropriate to use benchmarks, the same benchmark that we would use, the same tests that we would use for human beings.

Sabrina Hsueh:

Yeah, that's a good point. We should also test machine for reasoning and planning, right? Which humans are best at. Yeah. And did you see any industrial standards emerging at all?

Francesca Rossi:

Well, in AI and AI ethics, there are already several initiatives from many standard organizations. So I think the first one to have standards around AI ethics, not just AI is IEEE. IEEE is global association that also has a standard part. For example, IEEE is the organization that defined the global standard for Wi-Fi. The reason why we can go everywhere, everywhere in the world with our computer and plug into our Wi-Fi is because everybody uses this. So IEEE is a very active standard as organization. And they had put together already several years ago, a program of standards called the P7000™ program with about of, I think, 13 or 14 standards, all about AI ethics. One is about fairness, one is transparency, one robustness, one... I don't know, value embed [inaudible 00:48:28]-

Sabrina Hsueh:

[inaudible 00:48:28] in one of them called AI Nudging.

Francesca Rossi:

Yeah, okay. So that's really the first time that I saw standards explicitly around the AI ethics. But of course, there are standards on AI coming up, and already finalized like from ISO, but also there will be a lot of standards coming up in Europe from CEN-CENELEC, which is the European official organizations that define standards that can be used within the European laws. So now that the European AI Act has been approved, and with some months, and years will be then used all over Europe, then the standards will tell, the standards produced by these organizations CEN-CENELEC will define how this law will actually have to be implemented by the various companies or users, and so on. So the many standards will come up very soon from that organization.

Sabrina Hsueh:

Yeah. And will that lead to self-regulation more before AI regulation comes up, or you see this all together as one effort?

Francesca Rossi:

Well, again, in Europe, the regulation has already been approved by the European Parliament. Now, the regulation also says that in six months, or one year, or two years, it has to be implemented parts of the regulation with different timings. So the standards now in Europe, they have to work very fast because by that time when the regulation will have to be applied, the standard has to be in place because otherwise, people won't be able to know what to do to actually be compliant to the regulation.

Sabrina Hsueh:

Yeah. And that's a saying in US, right? So standards have to catch up.

Francesca Rossi:

Right, right. Yes. But I think there is a role also in the... Not just the standards for Europe or US or other regions, but also these global standards. Like the IEEE standards are not for a specific region, they're global, same ISO standards. So it's very important to have interoperability globally, not just within a territory.

Karmen Williams:

Wonderful. This has been such an informative, and exciting conversation about AI, and I know I'm most excited to see the three strongest professional societies, AAAI, AMIA, and, ACM really partner up and have a great collaboration. And with that, are there any parting words that you'd like to share with us?

Francesca Rossi:

No, but I really looking forward for these three societies to collaborate on these topics, and to do actual things that can be impactful, inspiring to all the members, and beyond the members of these organizations. But not only inspiring, but also very impactful in concrete ways. And thank you for having me. It was a nice chat, and I looking forward to the next steps.

Sabrina Hsueh:

Thank you.

This transcript was exported on May 07, 2024 - view latest version [here](#).

Speaker 4:

Thank you for listening to today's episode, ACM ByteCast is a production of the Association for Computing Machinery's Practitioner Board

Sabrina Hsueh:

And AMIA: For Your Informatics is a production of women in AMIA. To learn more about ACM, visit [acm.org](http://acm.org). And to learn more about AMIA, visit [amia.org](http://amia.org).

Karmen Williams:

For more information about this and other episodes, please visit [learning.acm.org/bytecast](http://learning.acm.org/bytecast) and for AMIA's For Your Informatics podcast, visit the news tab on [amia.org](http://amia.org).