

Food for thought: Is AI moving fast and breaking things?

point.zastone.ba/food-for-thought-is-ai-moving-fast-and-breaking-things

June 17, 2023

After the memorable Point Breakfast Buffet and very productive panel about innovations in fighting disinformation, moderator Antonella Napolitano opened the “Is AI Moving Fast And Breaking Things?” panel. In the opening remarks, Antonella expressed her joy of being back at the POINT Conference and stated the fact that this panel was just before lunchtime so she joked that this talk about AI is going to be something that actually gives the participants food for thought before their actual food.



Photo: Vanja Čerimagić

According to Antonella, AI is definitely something we’ve been hearing a lot about in so many ways, especially in the last six months. There has been some sort of media frenzy, especially about generative AI such as ChatGPT. So, it was a good thing this panel featured such great speakers who were able to handle this topic efficiently.

The conversation started with Eva, who is an engagement manager at the Mozilla Foundation’s insights team, on the topic of what AI not only can achieve but is achieving today. Eva stated that they at the Mozilla Foundation use what is called a trustworthy AI. She explained that trustworthy AI in practice stands for a couple of things. Firstly, it means that they at Mozilla Foundation believe in transparency and accountability in taking bias and AI seriously, and secondly, in looking at data governance, meaning they track the data that informs algorithmic decision-making. She described that it also means that the Mozilla Foundation funds researchers and start-ups as well, as an act of support of the efforts towards battling biased AI.



Photo: Vanja Čerimagić

“Technology is not entirely good nor entirely bad. It is always a challenge with AI”.

As an example, she mentioned the project called Melalogic which was founded in the United States by an engineer whose wife passed away from melanoma. At the time, AI detection had come far enough to identify skin cancer only on white skin and his wife was a black woman. So, the AI data that was used to identify skin cancer was simply trained on white skin and wouldn't work and even turn out deadly for non-white people.

Therefore, the above-mentioned engineer started working on the platform to collect data sets regarding any number of skin conditions for non-white skin. She stated that it was an example of addressing existing disparities in healthcare and how AI can be biased. On the other hand, she admitted AI has had a good impact, noting that ChatGPT, for example, certainly eased the challenges that come with written communication for people who have dyslexia or are facing language barriers.

“I think perhaps the best thing to come out of Chad GPT is that in the United States, perhaps, the administration will take AI regulation seriously”.

Speaking of transparency and accountability, which is also the foundation of the work of investigative journalists that have been investigating the arms of algorithms and AI, Antonella gave the word to Justin-Kasimir Braun, who's a data journalist at Lighthouse Report to tell a bit more about the investigation. They've recently published a report on the arms of some of the systems as well as the way they've been interrogating it and trying to bring transparency and accountability to this process.

In the last two years, they have embarked on this big project to investigate the application of welfare fraud detection algorithms across the European Union. This included sending hundreds of FOIA requests to welfare agencies across the EU to try to get access to as much technical documentation of AI systems that are being deployed in the context of welfare systems. Eventually, they got access to the entire life cycle of the algorithm of a welfare fraud detection tool in the city of Rotterdam in the Netherlands. Justin-Kasimir explained that this tool is essentially the file that you can feed new data into that will then give you an output score which tells you how risky the algorithm thinks a certain person is to commit welfare fraud. The investigation showed that the tool showed a discriminatory pattern towards young mothers and especially young mothers of migration background. Justin emphasized that is really an example of how an AI system can be used in a context where very very vulnerable people are being subjected to harrowing investigations. Lastly, he noted that he hopes the academics, journalists, and accountability workers can use the disclosures that they have made to further this investigative approach.



Photo: Vanja Čerimagić

“I think this provides a very stark example of well, not only how you can use AI but how AI and algorithms can actually be designed by flaw design, can discriminate against people and can further harm them”.

Kris Shrishak, from the Irish Council for Civil Liberties, talked about how the European Parliament voted on the AI Act this week. The AI Act is the European Union's approach to regulating artificial intelligence systems to be clear. It does not regulate the technology but only the use of this technology in specific use cases. Even though the first draft was finished in 2021, Kris stated that it is realistic to expect the act to be implemented by 2026 at the earliest. He briefly explained that the Act will only regulate certain applications and use cases that are considered to be highly risky. In addition, there are specific prohibitions on certain use cases that the EU considers to be unacceptable, such as biometric identification using artificial intelligence systems, social scoring, and deep fake, for example. Those are categorized as AI systems that require specific transparency, but Kris explained that transparency could mean so many different things to so many different people.



Photo: Vanja Čerimagić

“Does it mean to just inform you that you're interacting with an AI system or is it to inform regulators that specific systems have been deployed in the Union, for example?”

On the other hand, Kris stated that a very important improvement was made in this version of the Act and that is the definition of what an AI system is. He agreed with Eva, that documenting what data sets are being used to develop systems is a very important part of using AI. Lastly, he said that the AI Act would certainly have an effect and impact outside the EU as well.

The next speaker was Nasir Muftić from the University of Sarajevo. Nasir looked back on how to regulate AI adequately since there is a race between big players on the market which some authors call digital empires.



Photo: Vanja Čerimagić

“The thing that is very important to have in mind is that the big jurisdictions, the big digital empires, are the ones who attract and who are attractive to the smaller ones as their rules, their world view”.

He noted that in the case of Bosnia and Herzegovina and some smaller states, the EU has this sometimes formal authority and sometimes only some sort of soft power by which it influences the actions of market players as well as regulators in those states. So, if the actions of the EU will be replicated, how are those states able to implement them in a sufficient or sound manner? He emphasized that it is possible that any government can literally copy or use the piece of legislation as a model and adopt a similar instrument in its own national legal system. However, when it comes to implementation, a government may lack some of the resources that are necessary for a specific instrument to be applicable in its national legal system.

Nasir also mentioned that the Council of Europe is currently developing their own legally binding instrument in the AI regulations field. He said that there are also many other legally binding instruments, such as the Liability Directive, Revised Version of Product Liability Directive, Digital Services Act and Digital Markets Act, that are a part of the AI legal regulations.

Nancy Yu from Huridocs spoke last on the panel. She described that, first and foremost, we need to get to the root of what responsible tech means.



Photo: Vanja Čerimagić

“That means to have an obligation to do something, simply an obligation to do something”.

She emphasized that the most important thing in investigating AI systems is data labelling and the second is about retraining or relearning the system. In machine learning, data labelling is essentially the process of identifying the raw data images, text, files, and videos, and adding one or more meaningful and informative labels to provide the context so that the machine model has something to learn from. She noted that if there are problems at the root of original data sets then the whole system can be inherently problematic. And to correct it, we need humans and we need responsible humans to do the investigation mentioned above.

“So I’ll just conclude here by saying that to be responsible, it means that as system designers and technologists, we need to grapple with these questions and we need to be answerable and accountable for the outcomes produced”.

There were a few questions from the audience, which the speakers gladly answered and concluded this very informative and important panel. Just as Antonella said at the beginning, this discussion was something that truly gave the participants food for thought before lunchtime.

Author: Lamija Haračić