



How AI Disrupts the Law

Professor Sandra Wachter

11 October 2023

There are three areas that I would like to discuss today, where I think we can see that AI is disrupting the law as well as society. Those three areas have to do with misinformation, discrimination, and workplace automation. I didn't realize it rhymes, but it does, and it's also true. So, those are the three areas that I would like to discuss and just show a little bit how I think that AI is causing trouble in a way and what we can do about it. Let's start with the first one, which is misinformation and AI.

I think everybody will be aware that algorithms can cause a lot of trouble when it comes to misinformation. Some people say that we are in a post-truth society and that we can't trust anything anymore that we see online. And I think one of the most important examples of that is the Cambridge Analytica case in 2018. As everybody will be aware, we saw that algorithms and platforms were complicit in trying to shift voting behaviour and had an impact on elections around the world, including the UK and the US, but this was not the peak of how technology started to impact public opinion and spread misinformation. During the COVID crisis, we saw another very troubling example where President Trump told people that they should drink bleach to prevent a COVID-19 infection. Not only is that type of treatment ineffective, but it's also very dangerous and some people may die. This information was spread on social media very quickly, people believed it to be true and followed that advice. Similar advice was also followed during the insurrection of the Capitol in January 2021, where social media platforms were used to incite violence. It is troubling just from a democratic perspective to see how technology can have a real-life impact, not just on the health of people, but also on our symbols of democracy.

But not just democracy and physical health. Also, the planet suffers because of misinformation. There is a lot of misinformation about climate change out there. I think most people will be aware that climate change is a fact, but if it wasn't for social media and other types of algorithmic dissemination methods, we wouldn't still discuss whether climate change is a fact or not, we would do something about it. So, you can see that social media has a role to play when it comes to misinformation and its impact on public opinion. But is it new? Is that something that has never happened before? Haven't we always had misinformation? This reminds me of a wonderful quote that Mark Twain once said: "History may not repeat itself, but it rhymes." A fun fact is that Mark Twain never said that. This is an example of 200-year-old fake news.

I'm going to give you three historical examples of the past that show a type of misinformation that is pretty harmful and harmful to people. So, the first one is Queen Elizabeth. People said that she was a man, that was a rumour that was spread about her. She was a very popular ruler. She kept peace in the countries. It was, she was a supporter of the arts and sciences. She advanced her education, she ensured economic prosperity. Another rumour that hasn't left our public sphere is that Marie Antoinette said "Let them eat cake" but she never said that. This is a 300-year-old rumour spread by Jean Chakra, an activist, politician, and philosopher who was interested in furthering the French Revolution. Maria Antoinette did not only have to deal with rumours, but she also had to deal with pictures or paintings or drawings that were spread about her, such as one that she had a romantic relationship with her, and that she was unfaithful to her husband.

Well, that gets us to the core question, which is, what is the truth then? How can we know that anything is true if we have always had trouble figuring out what is true? I would like to cite the example of the Flat Earth movement. People believe there's a conspiracy against them, that they have been lied to, and that in fact, the earth is flat. People laugh at this because we think, "How can you stand up against established dogma and science? That's quite ridiculous." This is also how Galileo must have felt in the 17th century when he also stood up against dogma and the established sciences and religion. He believed that the earth circles around the sun and not the other way around, and he paid for it. He was put on house arrest for eight years until his death.

Coming back to the most recent example of COVID-19 and people advising not to drink bleach, we should always listen to the experts who have something to say about it. But you could ask, was it *a/ways* smart to believe our experts? Was it always the right decision? Here, I would like to remind you of the very interesting practice of bloodletting, a medical method that was deployed for more than 3000 years, where you would cut a person and let blood out of them because you believed that you could cure certain illnesses with it, such as hysteria or depression. It was only banned in the late 19th century.

If we have always struggled with truth and always struggled with lies? Is there anything new under the sun? How is technology any different? I believe it's different this time. I would like to go back to something that Mark Twain said: "A lie can travel halfway around the world while the truth is still putting its shoes on." But of course, Mark Twain didn't say that either. That's another fake news about him.

I think that one of the three reasons why I think AI is different is because of the speed and scale that misinformation spreads. People in the 18th and 19th centuries didn't have the means to disseminate information as quickly as we can now. Nowadays, we are in a situation where everybody has a microphone, and they are given a stage. For example, Facebook has roughly 3 billion users at present. And people like Alex Jones or President Trump have millions of followers where they can just spread their information. The second different thing is how convincing they are when they lie when algorithms start lying to us. For example, with fake Rembrandt portraits, fake faces or fake Tom Cruises. So, if it's done to an actor, to a political leader, what does it mean for society? What does it mean for peace? And then, of course, ChatGPT came along, and everybody could experience what algorithms can do and how they can generate text that is quite convincing, such as academic writing. But that also comes at a cost, right? What does it mean for education if professors are no longer able to assess whether the homework or the essays are authentic or if they have been generated by artificial intelligence? And what does it mean for society if we cannot trust that the students learn the things that they were supposed to learn at university? But that's not only an issue for academia in terms of education, it's also a problem for academia in terms of science because ChatGPT is also able to write academic papers.

There was an interesting experiment done in January 2023, where two batches of abstracts were sent out to expert peer reviewers. One batch of the abstract was done by experts, professors, by scientists. The other batch was completely fake and generated. The human experts were only able to detect the fake ones 68% of the time, which is not a lot. That's not good for the scientific community, not just because science needs to be true, but also because we use science to make policy. We want policymaking that is based on evidence, and that is based on scientific methods. However, if scientists can be fooled, how can we make sure that we write good laws that deal with climate change or pandemics? The scope and the reach are different.

And then, there is silence. The silence around the algorithms. We do not hear, see, feel, or smell algorithms, but they are always here. They are around us. They put us in filter bubbles. Every time you interact with digital technology, it learns about you. It learns very quickly who you are, what you like, what ethnicity you

have, what sexual orientation you have, who you voted for, and whether you're religious or not. And based on that information, everything is tailored for you. We think we have a shared reality online, but we don't. We have a fragmented reality. So, everything I see is different from what you see, but people don't know that. 80 to 90% of people believe that what they see on Google is neutral, that everybody sees the same search result, and that everybody sees an objective reality. We have known that since the Facebook leaks on October 21, when we started to see behind the curtains of big tech companies and to understand a little bit better how their business model works. This model works in a way where they capitalize on the attention economy. The idea is to keep you on the platform for as long as possible because that means advertisement revenue. However, research shows that what keeps people engaged is toxic content.

Another thing that algorithms are doing is they only show us part of the voices on the internet. So, algorithms think that content created by people of colour or people from the LGBTQ community is more toxic than the content of cis people. This is why it makes the spread of misinformation, and tailored misinformation, just very, very different. This leads us also to the second part where I think that AI is different and it's disrupting the legal framework that we have.

If you just ask the question, *how does an algorithm work?* It is looking at the past, and trying to predict the future. That's all it does. For example, let's just say I want to hire a law professor at Oxford, and I want to use an algorithm to help me with that. The thing that I would do is I would feed into the algorithm to all the historical information that I have about past hires. I would ask the algorithm, what do all those people have in common? And the algorithm would create a profile for me because those people were good law professors. Therefore, what will happen is that the algorithm will come up with that profile of that past information, then a new person will apply for that job, and the algorithm will say, "Well, do you look similar to that profile? If you do, you get invited to a job interview. If you don't, you're going to be rejected." That is the problem, right? When you take data from the past, you will transport the past inequalities into the future. And that is a problem. For example, here in the UK, we had a problem with the Quam algorithm in 2020. During the pandemic, it was not possible to have the A levels in person, so we had to move them to an online platform. What did happen, quite unsurprisingly, is that students of colour and students who were not at independent schools were disadvantaged. The class system and the unequal access to education in the UK were picked up by the algorithm and it reflected in the grades that were given to the students.

However, discrimination always happens. For example, if you live in China, there is a social credit scoring system. That means that the Chinese government is deploying a scheme where they take privately and publicly available data and use an algorithm to decide who is a "Good citizen." If you get a high social credit score, that means very good things for you, such as receiving better offers in supermarkets. If you are a bad citizen, it might mean that you're no longer allowed to leave the country. More specifically, being a video gamer in China makes you a social score drop. Another example would be the use of an Apple machine; for ten years we have known that if you are buying stuff online and you use an Apple machine to do so, you will pay higher prices. So, you're being discriminated based on the type of machine that you use. Then, in the Netherlands, we have that if you apply for insurance, you will pay higher rates if your address has only a number in it, as opposed to people who have an address with a number and a letter, who pay lower rates. Another type of discrimination is Face Recognition software. It determines how fast your retina moves, how much sweat you have on your forehead, the pitch of your voice, and how you gesture, which has an impact on whether you get an interview or not.

These would be groups that have been recently discriminated against: Video gamers, and people who might scroll fast on their screens, the letters they use on a job application, and how fast their retinas move on a computer screen. None of these groups are protected under the law, though. Why would they? They never

needed our protection at all. Non-discrimination law focuses on things like ethnicity and gender and sexual orientation, ability, and age, but Apple users, fast retina movers, slow scrollers, those people don't find any protection, yet they are the basis for decision-making for very important decisions in our society. Sadly, the law is not prepared for it because it never had to be. AI is disrupting the law and society because misinformation is different, and because discrimination is. With all the things that we have discussed, you could rightly ask, what does it mean for our job market? If you have an algorithm that is twice as fast at half the cost, is your job still safe?

AI can do a wide range of things. It can write poetry, it can write screenplays, can produce paintings, and, in the case of ChatGPT, it can write a press release to summarize a police report or do a legal draft. Then, what does that mean for artists, doctors, journalists, lawyers, office workers, scientists, or even coders if part of their jobs can be automated? Unfortunately, this is not a new thing. Technology has always displaced and replaced some of the work, but it's different because we never had a technology that was disrupting so many sectors at the same time. I think we need to worry about and the legal of framework needs to adjust to it.

Recently, there has been some disheartening news on this matter. For example, IBM issued a public statement where they said that they're not going to hire specific back-office workers anymore because they will be replaced with AI; this means that roughly 8,000 people will be unemployed. Twitter, Facebook and Amazon have let a lot of their staff go, as well. Therefore, the question of whether AI will generate new jobs is unclear at this point.

But I also want to talk about solutions and proposed solutions. I would say that, at the moment, the bad news is that we are focusing on the wrong type of solution. In the case of ChatGPT, there were many open letters after its release that were calling to stop that type of technology, so we could catch up with our legal and ethical thinking around it. It sounded like a good idea, but, if we examine the wording of the letter, it says things like "We have to make sure that AI doesn't rival human intelligence." We have to make sure that it doesn't outnumber us, outsmart us, make us obsolete and replace us. We have to make sure that we don't lose control of it. And from the wording, you can say that the letter is very concerned with the idea of AI somehow becoming sentient and having its agenda. Then, another letter was published in June and it said, "We have to the risk of AI waking up and enslaving us just as seriously as we have to take the risk of pandemic and nuclear wars." However, there is no scientific evidence that we are on a path to sentient AI.

Algorithms do need data, which means that they are grabbing everything that they can find on the internet. But data belongs to somebody, it doesn't grow on trees. People do have intellectual property rights; the problem is that people's data is being taken away and repurposed without them being asked. Data collection needs a lot of resources: It needs a lot of water, space, and electricity. Just to give you an idea: One training session of ChatGPT requires the same amount of energy of 126 Danish households in a year, but for one training session to cool down (because you need to cool the servers in a data centre), you need 360,000 gallons of water every day. An interaction with ChatGPT, 20 questions asked, costs us half a litre of water. Those are resources that nobody talks about. These are real issues.

Unfortunately, around the world and in the UK, policymakers started buying into the idea that this is not the real battle. The real battle is that speculative fiction one where we worry about the killer robots at some point taking over. For example, in the AI Safety Summit, which will be held on November 2023; the whole focus of that summit is none of the issues that we have discussed. It is all about AI safety in the sense of making sure we don't lose control of AI. Losing control of AI is a very powerful narrative, but it is also a very wrong narrative because AI isn't an alien that lands on our planet and now we have to negotiate with it. We built it. Every design decision is an ethical decision. We made it that way. So, saying that we now have to plead with it and

make it listen to us is taking away the responsibility of the people who built those systems in the first place.

The European Union is proposing an “AI Act,” which will probably come into force either at the end of this year or next year. The framework has really good points in it, like a risk-based approach where you think about where technology is being deployed and depending on how risky it is, there are certain rules that you have to follow. The most interesting part is that at a higher risk application, the legislator said, eight different categories will believe are higher risk. None of that is banned, but you have to do certain risk assessment and documentation. However, it doesn't deal with any of the things that we have discussed. It does not talk about misinformation, employment, population displacement, or the workforce. As it stands, you will have the “AI Act” that will come into force by the end of the year, and then there will be another two years or so where standards will be developed... But who writes those standards? In the case of the European Union, this is done by private entities, meaning that they are not democratically legitimized. As a result, you will have centennial industries writing those standards for themselves. Other entities like NGOs or civil societies only have so-called observer status, which means that some of them get a seat on the table, but they do not get voting rights. This begs the question: How can you be sure that they're following those rules? Of course, it could work, but it could also go wrong.

What is the good news? The European Parliament has proposed amends to the Act, and I hope this will go through. For example, they have now suggested that the recommended system that can contribute to misinformation should be deemed a high-risk system. It has also been proposed that you should have a right to complain and that you should have a right to an explanation. The last good news is that I would like to dismantle the fake news genie out of the bottle and fight the idea that “There is nothing we can do about it.” Yes, AI is here to stay, and as such, you can either jump on the train or be left behind. We must question *who is giving us this information*. Because it is usually people who have an interest in the genie is out of the bottle.

Every time there is a new technology, we have a hype cycle. Every time something like that happens, people say, “Oh, this is never going to be the same...” But that is not always true. We are the ones that get to decide because we are the ones that shape technology if we want to. We are the ones that can decide if we want to adopt it. And we have a say in how to regulate it so we can make sure that technology serves us and not the other way around. Thank you very much.

© Professor Sandra Wachter 2023