



An interview with... Merve HICKOK – Alethicist.org / “100 Brilliant Women in AI Ethics 2021”

Description

Interview with... Merve Hickok

Founder of Alethicist.org and voted as one of the “100 Brilliant Women in AI Ethics 2021”

1) What is your personal definition of ethics?

Following your moral compass and values to do the right thing even when you are under pressure, or even when no one is looking, while respecting others' lived experiences and actively expanding your knowledge.

Merve Hickok

2) You have founded [Alethicist.org](https://alethicist.org). Tell us a bit about it...

It grew out of my frustration years ago when I was deep into researching (for my own benefit and understanding) all the problematic issues with AI systems.

I was looking at issues like different aspects of bias, opaqueness of AI systems, some of the principles and how they translated into frameworks etc.. I was trying to get the bigger picture, understand different perspectives and distill them into practical work. However, I kept going down the rabbit hole with a lot of works.

The field is relatively new so you can find so many new publications about these subjects, but it was hard to distinguish the objective ones from promotional works. There are also a lot of technical papers (naturally) that intimidate people who are interested in the field but do not know where to begin.

So I thought I could put together a curated list of works on these subjects to help others who are either trying to understand these concepts better or actually doing relevant academic research.

[Alethicist.org](https://alethicist.org) is a living website. I update the content regularly every month. I also try to share my own work (webinars, podcasts, articles etc) as a lot of efforts is geared towards awareness raising and advocacy.

3) You have chosen “ethicist” as opposed to “ethics” in your company name.

What does this difference mean to you?

I wanted to highlight the human and human decisions behind AI & other disruptive technologies.

AI systems do not have their own values, decisions, moral compass. They are artefacts of human decisions and interactions.

I sincerely believe that while we are in this phase of understanding the impact of these systems on individuals, on society, and on our institutions and while we are trying to improve our practices and products, we need dedicated roles in organizations. This is not to say that ethics is a single person's responsibility in the organization – definitely the opposite. It is everyone's responsibility, but a dedicated role during a transitional period who can organize and mobilize teams, can help drive change quicker, better and with more accountability. I actually wrote about what that role looks (or should look) like in an organization <https://medium.com/@MerveHickok/what-does-an-ai-ethicist-do-a-guide-for-the-why-the-what-and-the-how-643e1bfab2e9> .

My moonshot dream is a day when my role becomes obsolete.

A time where we have made such changes in our organizations, in our team compositions and our interactions and what we want to optimize that the ethical questioning and prioritization on humans and human rights are just second nature to everyone involved...that it is embedded in the essence of our projects and products

4) You are invested in AI and “bias”. Do you believe that AI will be able to solve this human tendency/feature?

One good thing about all the debate about AI bias has been the understanding that our society is biased; our decisions about data and models and optimization are biased and that this all relates to power. Some of the scandalous results have actually helped us materialize what we have been discussing all this years about the structural issues in our societies. We, as developers, decisionmakers, policy and lawmakers, etc now have to make conscious decisions. Do we want to use these systems to lock in these structural injustices and possibly deepen and magnify them? Or do we want to change and transform our societies and institutions for the better? We need to take responsibility about imagining a better future, working towards it and holding ourselves accountable for what we create and use (AI systems in this case).

5) Where do you believe AI bias to be most dangerous?

Any area where it has the potential to impact a person's life or quality of life, access to resources and opportunities. The issue with AI systems is not only the bias issues but how do we treat the outcomes of these systems (which is our own biases whether intentional or not). We take data as objective ground truth, and the predictions that the system makes as facts, as causal relations. So when you start with those problematic fundamentals and use them in impact areas I mentioned before, application areas like justice system, military, law enforcement, finance/credit, recruitment, health, education, etc all become risky areas. You can make life/death decisions in some, or you can lock a person / marginalized groups in deeper perpetual loops that give them no possibility of breaking it (for example credit scoring, employment, predictive policing, etc).

6) What do you particularly feel committed to?

Social justice and a future where technology is not used to assert asymmetric power. A future where everyone benefits from technology as they define the technologies for themselves.

A lot of my work is raising awareness about the implications and consequences of current practices with AI systems and data practices, and then how to improve our work.

These systems have impact on our lives whether you define yourself a consumer, a citizen or an employee. We interact with them 24/7 but do not necessarily understand the consequences. With my ethics, governance and policy work, I am trying to move the needle as much as I can where these AI systems and practices are not accepted as inevitable but that we question the framing that is imposed on us.

7) The EU's newly proposed Regulation on AI, in which way do you believe it essential?

It is the first comprehensive regulation that actually has a chance to become law at the end. We have seen EU's ability to change the landscape with GDPR, and I think the impact of this proposed law will be even more substantial. Ethics is a voluntary governance so there are limits to what the companies and even governments will commit to if left to themselves.

The market cannot regulate itself fully in a way that protects human rights, and address social justice and societal concerns. Business models and priorities sometimes clash with these concepts.

On the flip side, only relying on law and compliance is not enough either. We need ethics and that moral compass to challenge the technologies as well as law while our societies and realities change. So the proposed regulation on AI systems which is coming after years of high-level discussion about ethics and principles etc is very welcome. It also provides some certainty and guidelines to businesses too. It will still take a few years for the proposed legislation to become law and hopefully some of the controversial and/or missing points in the proposal are addressed by that time.

8) You have been voted as one of the "100 Brilliant Women in AI Ethics 2021". What does that mean to you?

It was an absolute honor to be voted into the list and be in the company of some amazing women working towards ethical and responsible technology in their own professions and activism. Women in AI Ethics' mission is to increase recognition, representation, and empowerment of women in AI Ethics. These are women who imagine a better world for everyone and are passionate about social justice, equity and inclusivity.

9) You will start a new online course on June 1st ! What will you be focusing on and how can it practically be useful to anybody in AI?

This is actually my second online course with RMDS Lab. The first one is now available as self-paced course on the same platform and it was a deep-dive course covering the fundamentals of AI & data ethics, bias in product lifecycle. The course starting on June 1st covers the fundamental concepts in AI ethics from a principles/human rights/regulatory perspective; how the facial, object and language recognition and analysis systems manifest bias, how they are used in different settings and what the consequences are. The second half of the program focuses on surveillance methods and impact on our lives with discussions of real-world application in workplace, education, public spaces, consumer spaces etc. Last session is focused on similar subjects but in smart home and smart city context. Throughout this program, learners will gain a strong understanding of these systems, and the consequences of these systems in their lives as well as social justice applications.

The 2 courses are independent of each other and do not require any technical background. Link to the registration <https://grmds.org/deep-dive/AI-Bias-Surveillance-Recognition-Analysis-Prediction>

10) Who is your favorite person in history /science /literature / AI...

Hard question ? I am a very curious and inquisitive person so having read his biography and his approach to world and his surroundings, I would first say Leonardo Da Vinci. But Ada Lovelace, Grace Hopper and Ruth Bader Ginsburg are my trio of heroes

You can register to Merve's AI Bias & Surveillance Course

No technical prerequisites needed!

Copyright House of Ethics

Interactive training course

AI Bias & Surveillance: Recognition, Analysis & Prediction

6 Sessions

Copyright House of Ethics

Overview

Detection and analysis systems are ubiquitously used across domains of life (education, work, leisure) to access resources and opportunities. This six-week program covers the fundamental concepts in AI ethics, how the recognition and analysis systems manifest bias, how they are used in different settings and what consequences are. Throughout this program, learners will gain a strong understanding of these systems, and the consequences of what they develop, use and implement in their lives.

Register now

Date Created

May 30, 2021

Author

katja-rausch