

Ethics of Artificial (Narrow) Intelligence - Summarized Presenter Notes (v3)

Nicholas Kalogirou | July 16 2020

PART 1 - Professions, Ethics, and ANI

How do we design technology for human thriving?

Professional engineering background. APEGA's first rule of conduct for engineers in Alberta.

["In their area of practice, engineers shall hold paramount the health, safety, and welfare of the public, and have regard for the environment."]. Example: effect of boiler safety regulation for the Alberta Boiler Safety Association in 1897 saving lives, after 25K+ died / injured in last decades of the 19th century.

Professions throughout history have developed ethical codes, which help practitioners REASON through the impact of their practice on people and environments. Like engineering in the 19th century, **AI doesn't have a professional body** or an ethical code. There is a growing recognition that with data in every part of our lives, that we need to develop our reasoning of both harms and benefits of AI algorithms

Ethics is a **critical, disciplined questioning** of what is right and wrong. **[How would you know whether you are doing more good than harm?]** is one of the central questions of ethics.

Today's talk will focus on Artificial Narrow intelligence (ANI), which just means **intelligence in a really specific area**. ANI has big ethical implications and impacts in this world right now. Not covering machine ethics / Artificial General Intelligence in this talk.

PART 2 - Nested Systems Awareness

Here's a model to explore ethics on different levels, each nested within one another.

Models are a toy version of reality, which help us understand something about reality. Models are given its context and data by the **modeller** (eg data scientist). Modellers typically get context and information from the **organization** (eg corporation, institution).

Last summer I encountered a moral dilemma. I followed a lead for a customer analytics engagement for a casino. The question was, what information can we collect on casino goers, in order to market to them, so they spend more time and money at the casino. [see benefits / harms on slides]

The question at this level: how do our models connect to **not just economic systems** but other systems like population health, justice, defense, and education? This casino engagement is a good example of how we can look beyond our immediate circumstance and **reason at different levels**.

Extending our reasoning - our societies survival depends on the planet earth's [natural and life systems]. If natural and life systems breakdown, our society will breakdown as well.

The impacts at the of data science at these larger levels often have to do with **accelerating existing problem areas**. [refining AI example - AI that could both worsen, or help mitigate climate change].

We're in a constant process of deciding, **how much information from the higher levels** do we let inform the lower levels? The better we can reason at larger levels, the better chance we have to care at larger levels. The marker of our **ethical progress is how we are able to care** at progressively larger levels.

Ethics of Artificial (Narrow) Intelligence - Summarized Presenter Notes (v3)

Looking at the modeller

When we include, or exclude data from higher levels, this tells us about our goals, our views, and our values, and our ability to care at different levels. **Ethics doesn't just make us question the systems we're a part of, but makes us question ourselves.** Ethics holds the mirror in front of us and asks what do I value, what's my level of caring, why do I care, what responsibility do I accept?, courage to question?, how do I verify my own knowledge?

If as a data scientist you want your models to be more fair, and accountable - then we have to develop those traits in ourselves. **How are we adopting the identity, attitude, behaviours, of someone who is more fair, more accountable, more truthful, more compassionate?** If you develop attention to care, compassion, and truth, you will find opportunities to practice care, compassion, and truth.

As the modeller, in ethics we have to do both the work of looking in, but also looking out.

PART 3: Who and What of AI Ethics

Individuals like **data practitioners, technologists, academics** work in various **organizations** like corporations, non profits, academic institutions, governments, working on ethical **research, guidelines, technical tools, laws and regulation, and deployment.**

For example the area technical tools for fairness, bias, and explainability is the area with the most development going on. There's also work going in governments, the most famous example being the EU's General Data Protection Regulation, or non profit groups like the **Center for Humane Technology (highly recommend "Your Undivided Attention" podcast)** or **AINow (ex-Google/FB employees)**, who call for urgent, equitable, policy work to happen.

The top 3 themes in a recent **literature review** of AI ethics guidelines in 2018 are **accountability, fairness, and privacy**, with over 80% of guidelines referencing this. An interesting finding is that these top 3 categories are aspects for which technical fixes can be developed.

What are the gaps?

Seeing ethical AI as a technical problem; no consequences for violating ethics standards and codes; looking to vague guidelines and not ourselves for ethical development; skipping questioning in the name of profit.

Concerning societal consequences of AI

The biggest problems with AI is not the AI itself **but the acceleration of existing social challenges.**

AI is accelerating trends in **inequality**. Low wage earners are seeing threats to their job security and wages. Algorithms in our lives solidifying the gap between poor and rich. In our **online social platforms** we are being pushed into social and political tribes by algorithms, with troubling repercussions for our politics, civil discourse, and democracies.

A second major concern is the **loss of liberty** being accelerated by AI. We are using AI to **consume masses of personal data**. In the digital age, as digital citizens, privacy IS liberty. We're also starting to see AI as a tool of government and corporate control through increased surveillance of so many aspects of our lives.

Lastly, **the truth is under attack** with the aid of algorithms and social technology platforms. We are pushed into social and political tribes, flooded with misinformation and divisive messaging. We are open to manipulation without oversight of algorithms, especially in massively scaled social platforms. Truth is a pillar of civil discourse.

Ethics of Artificial (Narrow) Intelligence - Summarized Presenter Notes (v3)

PART 4: What Can We Do as Individuals

Develop your self. Find your own **reasons why you care**, because your caring, will affect your motivation.

Continually develop character as a human, and **not just technically**

Discuss values. encourage open, honest discussions with about your own and your organization's gap between value and action. Without open discussion, we have very little chance of action.

Question broadly. Use and develop **ethical checklists** and techniques that **help you ask** broader system questions (deon). Try to not just ask about cost efficiency but who and what is being **empowered? Keep on asking** the questions through deployment.

Public Pressure. Lastly, we can **encourage public discourse and pressure** on corporations and governments. We need to **hold harmful AI algorithms accountable - to do this we can support informed public policy, public action, professional ethics, self organization.**

deon: A Data Science Ethics Checklist

Deon, and its a data science ethics checklist with questions and examples along the data analytic lifecycle.

Proxy discrimination example - geographic area (postal / zip code) is highly **correlated with socioeconomic** status and **race**. So we have to ask if its appropriate to use that variable, as it may be possibly discriminatory

Fairness across groups example - facial recognition is significantly worse at identifying people with darker skin.

Nick's Process Explorations

We have some very useful checklists and technical approaches, but we also want to develop ourselves, our character, and our virtues. The first version of my practice, is just a **simple combination** of look at oneself, one's org, one's motivations, then moving on to the ethical checklist like deon.

First and foremost, we have to find out **why we care** - I think we develop this through the trait of **compassion**. Second, we have to look outwards and be **aware of impacts at broader systems**. Third, we have to **accept some level of responsibility**. After this step, we could use an ethical checklist like deon to proceed. And around the whole process, is a loop of ethics - both caring and checklists have to be **continually revised over time**.

Conclusion

What I've taken away from exploring all this, is that **IT IS possible** to develop our ethics. There ARE other people on this frontier to help. It will take **compassion, responsibility, and aware use** of AI to build a resilient world that supports our well being.

"We make our world significant by the courage of our questions, and the depth of our answers - Carl Sagan"

And **it all starts with CARING and COURAGEOUS questioning** - it's never going to be easy. To build a more ethical world, we have to encourage and help people to question, try, think, and act; to do the **hard work of acting on higher ethical standards for the public, the environments we live in, and for future generations.**

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>.