AI and Ethics

Richard Wheeler

Reykjavik University

Today

Today we break up into two groups acting as ethical committees, and review a real use case from my own work. It's a hard one, but … typical. You have 45 minutes to answer some questions and make decisions *as a group*, then five minutes for one member of your group to present your decisions and recommendations. After 20 minutes of your meeting, there will be a short break in which you can ask for clarifications on the system and situation.

Try to observe how you discuss ethical issues in a group. Do you appoint a leader, have a formal process, or just have a discussion and agree on a result?

Group dynamics are … unusual.

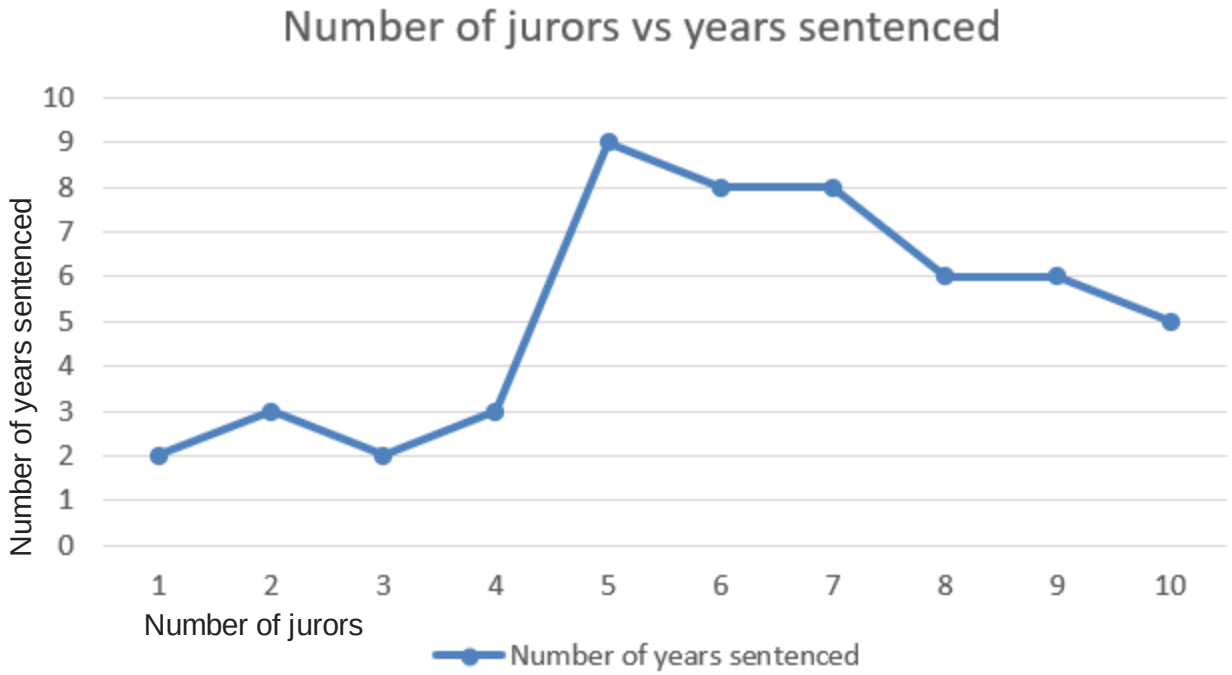
On Group Thinking

How humans make decisions in groups is seriously *weird.* We make decisions in groups we'd never make on our own.

Let's say you committed a crime, and were already found guilty. Now, a panel of regular people, *jurors*, will decide how many years you will get in prison in the range of 2 – 9 years.

How many jurors between 1 – 10 would you choose?

On Group Thinking



When asked before the jury meeting for their recommendation, most jurors will say around 3 years. But together, they decide as high as 9 years. Why?

These effects mostly, or nearly only, occur when the jurors are in group meetings interacting. These same effects are used by cults and extremist groups for recruiting.

The Case Study

When I worked for the World Health Organisation Headquarters in Geneva, Switzerland, I designed a system for tracking and assessing efficacy of tuberculosis treatments in (a large Asian country). Tuberculosis treatment was then by diagnosis and treatment with tablets taken regularly. This was in the days before data analysis was common, and I was using data mining and modelling to figure out what the likely effectiveness of the WHO programme was. To my shock, I discovered the WHO were sending to this large country *three times* the amount of drugs needed to treat all of the recorded cases. Of course the WHO knew, they just wanted to see data and proof.

Why was there such a big mismatch between the amount of drugs being used and the number of registered patients? Were there unrecorded patients? Or were they oversubscribing by accident? Were some of the drugs spoiling somehow or being damaged? Or were they stockpiling drugs somewhere for future use? Were the drugs even reaching the patients? My system was meant to prescribe individual drug regimes to individual patients, patients who seemed to be receiving three times the prescribed amounts of drugs. Should my system cut the shipments to the country? What's going on here?

The Case Study

What an expert system in 1995 produced as output…

For one year's TB treatment in one county ->

There were 31960 violations of best practice in treatment ->
And 7 warnings so serious they *have* to be investigated ->
Followed-up cases not equal to patients ->

Number of relapsed patients is going up not down ->

People seem to be getting sicker after treatment, not better ->

Number of people treated and diagnosed are not equal ->
Slides used are more than 3 times the diagnostics performed ->

Cases are not being followed up correctly; patients disappearing ->

Gender distribution is 4 times off statistical norm ->

****

*WARNING: DBNAME: composite CHART: composite PROVINCE: "[redacted]"

*YEAR: "1992" QUARTER composite COUNTY: "[redacted]" COUNTYCODE:

*"130826" POPULATION: 319600 VIOLATIONS: * 7 WARNINGS * Total

*cases quarter 2 and follow up count unequal: (A39 not equal to E5) *

*Relapse patients followed up in 1993 and previous year cases *unequal: (B61 not equal to A11+A12) *

Chart 5: Other initial cases *sputum conversion at 2 months *greater than at 3 months: (E7>E17) *

*Smear positive cases and clinical visits unequal: (F1 not equal to *D7+D18) * Slide usage outside of allowable range: (E58>F8*3) or

*(E58<F8*2) * Total cases quarter 3 and follow up count unequal:

*(A39 not equal to E5) * Unequal gender distribution in new smear-*positive cases: *([males * 4] < females) * **********

****

The Case Study

After much analysis, we discovered that the drugs were being carried and delivered by the Army, and literally every guard post, driver, carrier, and hospital worker had to be paid off to do their job. So about 65% of all of the drugs were being stolen or misappropriated, and the AI system I built showed clearly at which junctures things were disappearing. It was millions of Euros of drugs per year. In this large Asian country, the national health authorities had their own tracking system that faked all of the numbers and usage, and assigned anything missing as prescribed, lost or spoiled, or not received. The misappropriated drugs were then being sold and exiting the country on the black market to be bought in countries that had refused the WHO's help, such as Cambodia, etc., many of the sales going directly to foreign governments, including those with western boycotts. What to do now? I assumed we should meet with the country's government and expose the corruption, and stop supplying them the treatment drugs, or reduce them to only the amounts actually needed. Or maybe we should let the AI system I built specify the amount of drugs to be shipped to them?

Let’s Talk Ethical Processes

The questions:

1. Was it ethical to design this system to make recommendations for treatment? Was it likely to be accurate or predictable? What could be done to make it more ethical? Should there be humans in the loop?
2. What should the WHO do about the 65% of disappearing treatment drugs? What would it mean for the AI system?
3. What process for discussion and decision-making by myself or the WHO might have been most appropriate?
4. Is there a policy recommendation here for making further similar decisions?
5. As a group, can you decide what to do, and what is your decision?

The Case Study

When I worked for the World Health Organisation Headquarters in Geneva, Switzerland, I designed a system for tracking and assessing efficacy of tuberculosis treatments in (a large Asian country). Tuberculosis treatment was then by diagnosis and treatment with tablets taken regularly. This was in the days before data analysis was common, and I was using data mining and modelling to figure out what the likely effectiveness of the WHO programme was. To my shock, I discovered the WHO were sending to this large country *three times* the amount of drugs needed to treat all of the recorded cases. Of course the WHO knew, they just wanted to see data and proof. **Why was there such a big mismatch between the amount of drugs being used and the number of registered patients?** Were there unrecorded patients? Or were they oversubscribing by accident? Were some of the drugs spoiling somehow or being damaged? Or were they stockpiling drugs somewhere for future use? Were the drugs even reaching the patients? My system was meant to prescribe individual drug regimes to individual patients, patients who seemed to be receiving three times the prescribed amounts of drugs. Should my system cut the shipments to the country? What's going on here?

After much analysis, we discovered that the drugs were being carried and delivered by the Army, and literally every guard post, driver, carrier, and hospital worker had to be paid off to do their job. So about 65% of all of the drugs were being stolen or misappropriated, and the AI system I built showed clearly at which junctures things were disappearing. It was millions of Euros of drugs per year. In this large Asian country, the national health authorities had their own tracking system that faked all of the numbers and usage, and assigned anything missing as prescribed, lost or spoiled, or not received. The misappropriated drugs were then being sold and exiting the country on the black market to be bought in countries that had refused the WHO's help, such as Cambodia, etc., many of the sales going directly to foreign governments, including those with western boycotts. What to do now? I assumed we should meet with the country's government and expose the corruption, and stop supplying them the treatment drugs, or reduce them to only the amount actually needed. Or maybe we should let the AI system I built specify the amount of drugs to be shipped to them?

The Questions

You have 45 minutes to answer these questions and make decisions as a group, then five minutes for one member of your group to present your decisions and result. After 20 minutes of discussion, there will be a five minute break in which you can ask for clarifications on the system and situation.

The questions:

- Was it ethical to design this system to make recommendations for treatment? Was it likely to be accurate or predictable? What could be done to make it more ethical? Should there be humans in the loop?
- What should the WHO do about the 2/3 of disappearing treatment drugs? What would it mean for the AI system?
- What process for discussion and decision-making by myself or the WHO might have been most appropriate?
- Is there a policy recommendation here for making further similar decisions?
- As a group, can you decide what to do, and what is your decision?

The Outcome

What did the WHO do?

The actual outcome was very enlightening. I was at a high-level meeting to discuss it, and the big bosses said that legally, they had an obligation to report everything found. But ethically, this would only lead to people dying of tuberculosis being denied treatment. They decided that in this case, their ethical obligation was more important than their legal obligation, which was actually fairly complex, and that their ethical obligation was to ensure the maximum number of people were treated successfully, whether in that big country or elsewhere, which is how it was working out. The drugs were not being wasted, just stolen and misappropriated to other countries where the WHO could not act. It was also decided that ethically, they had an obligation to help ill people in countries even under legal boycott, and this method of routing the treatment drugs through the black market informally was actually a good way to get around restrictions, as long as they were not the ones doing it themselves. They later impressed upon me the importance of not letting a bias for western legal formalisms confuse or obstruct our obligations to humanity, and that in the real world of global health care, few processes were ever fully legal or ethical, and that they typically had to choose among ethically and legally challenging alternatives, or do nothing, which they found to be unacceptable.

The Outcome

What did I do? On the AI system side, I decided that I would apply a fuzzy modifier of 300% to all treatments to account for the losses, and track the *actual* drug use per patient across that large Asian country. The same system was then optimised to guess how many external undocumented persons were being treated and cured as a likely side-effect. Was this ethical? What if the people using that system now do not know about the modifier and it somehow leads to triple doses of treatments?

The Outcome

Unintended consequences of this work:

- I had to embargo my Master's Degree thesis for ten years. I was not allowed to publish our results for ten years. Story: a map of military supply routes.
- The system was used for nearly ten years before a new version was created by their in-house team.
- This was the first real AI system designed for the WHO (1993/1994 or so), and led to the creation of an in-house data analytics team focused on AI, but this would take another eight years.
- This work won the PARC XEROX Prize for best AI post-graduate work in 1994. The prize was $50. :>)
- After this, all anyone wanted me to do was AI & Medicine.
- I worked for the WHO another four years, and then volunteered (mostly unpaid) for another 15 years. In that time I developed AI systems for tuberculosis control, measles diagnosis, mosquito identification, malaria treatment, developing world traffic safety, and drug treatment regime optimisation systems.
- Working for the WHO was very … *enlightening.* Who you *save* is very political. Meeting Bill Gates, humour in unlikely places: children's cancer unit.

The Outcome

What Have We Learned

What have we learned in this exercise?

For Next Class

For Monday in four days, write and submit around one page of analysis of the ethical consensus process. What role did you feel comfortable playing? Was it easy to make decisions together, and did you think it was fair? Were there things you disagreed with? Discuss your own feelings for the ethical issues involved. Were you more interested in policy, process, technology, philosophical ethics? What that was unexpected did you learn?

The assignment is to show knowledge and insight into what constitutes ethical issues in AI systems and the processes for decision making. Explore what parts of the discussion you found most interesting given your background and scientific interests.

Ask Me Anything,

…and I mean *anything.*