

Individual Assignment, Machine Learning Ethics

Pete Gray, November 14, 2019

York University CSML1000 "Machine Learning in the Business Context", Fall 2019.

1. Unemployment. What happens after the end of jobs?

I have mixed feelings about making jobs obsolete. When talk turns to "Look at trucking: it currently employs millions of individuals in the United States alone," (quoted from original assignment text) I like to invoke the image of millions of individuals who lost their jobs tilling fields with oxen. Seriously, every single person whose survival is not tied to tilling fields with oxen is better off, whether they're making the most of it or not. We have been automating jobs for a very, very long time. I was on a software development team in the 1990s that probably put many chemical, paper, and printing workers out of their jobs, by creating the first web-based system that allowed remote digital soft proofing of print jobs. A great advancement in efficiency! The advent of consumer digital photography arrived soon after, wiping out even more of those jobs related to the processing of film. I certainly don't feel bad about all the chemicals that are no longer manufactured and set loose on the World. Or all that paper. But do I feel bad for those workers? Maybe I do. Did they live in a place where their financial security and modern skills re-training were assured by the country to which they had been paying taxes? HA! They probably had to get jobs driving trucks.

The hope is, as we automate jobs, we can move on to bigger and better things. We don't till fields with oxen, we drive trucks. We don't drive trucks, we're YouTubers. The YouTubers of today have no idea what will be a hot job when AIs become the most efficient purveyors of video content. We certainly know that Machine Learning Specialist is going to be pretty hot in the coming years. But how good at preparing people for the future is our society now? In ***What Happens if Robots Take The Jobs? The Impact of emerging technologies on employment and public policy***[1], Darrell West advances a number of ideas that address the need for society to adapt to disruptive technological change. "*There needs to be ways for people to live fulfilling lives even if society needs relatively few workers,*" West writes. One recommendation is "retraining accounts", which are publicly funded accounts for funding

of re-training. This approach purports to offer the availability of free education, without as much potential for people becoming full-time students and not returning to the workforce. Also important is Curricular Reform. In an age of constant change, it is important for school boards to rapidly adapt to the changing demands of the job market. I have witnessed this myself – my 12-year-old daughter is making websites in her Grade 7 class. Once the domain of specialists, technology now allows almost anyone with creative spirit to perform this task. This is part of the process by which technology turns into progress. As once we learned to use machines to till our fields and truck our vegetables, we now learn to use machines to publish and distribute our ideas to the world. Finally, West speaks of an “artisanal economy”, in which mundane tasks such as driving and plowing with oxen are performed by machines, while humans participate in the supply and demand of art, culinary delights, music, research, websites, YouTube videos, exploration, and the like. Sounds very Utopian. But I fear we will need more than re-training accounts and modernized curricula to get there.

[1] <https://www.brookings.edu/research/what-happens-if-robots-take-the-jobs-the-impact-of-emerging-technologies-on-employment-and-public-policy/>

2. Inequality. How do we distribute the wealth created by machines?

One simple, effective, but unpalatable solution to this would be to tax the owners of the machines that do the jobs formerly done by humans, and then provide a basic minimum income to all the humans. AI or no AI, wealth inequality and a lack of access to quality education are serious issues already. As our ability to automate human activity accelerates, so too does the impact of these problems. Education is very difficult for people struggling to pay the bills. In an ideal world, or a well-run country, there would be thousands of people in programs just like York's Machine Learning Certificate, and they would be able to focus on the programs, rather than scrambling full-time to keep their lives together and doing what they can at school in the meantime. And so I believe, in addition to the educational improvements outlined in Question #1, *Unemployment*, it's really time for society to disrupt the disrupting power of disruptive technologies by implementing a Basic Minimum Income.

Unfortunately, the political resistance to this idea is very strong. Ontario was going to do experiments with Basic Minimum Income — but our habitual lurching from governing party

to governing party have left that experiment on the cutting room floor. People, especially in North America, don't like the idea of giving people money for free. Or, the idea of “stealing” tax money from the job creators. But at some point, we must recognize that their primary role, is not that of job creators. They create jobs if they absolutely have to. If they can automate instead, rest assured, for the benefit of the shareholders, they will. As robots and AIs do more and more of the work, and less and less people are needed to do these tasks, it seems reasonable that the robots, or their owners, are rewarded a little less for the service they no longer provide to society. Charles Kenny looks at ethical and practical issues surrounding Basic Minimum Income in ***Give Poor People Cash***[2]. Central to his thesis, and that of others who have studied this and other social programs, is that it is the most efficient, flexible, and productive way to deliver a social safety net. Basic Minimum Income completely avoids the inefficiencies and frauds associated with benefits programs that are targeted, controlled, or conditional. It also catalyzes economic growth through the flexibility it brings to how the money is spent, or invested, into the local economy. Including, of course, allowing people to focus more on education and become more productive members of society.

[2] <https://www.theatlantic.com/international/archive/2015/09/welfare-reform-direct-cash-poor/407236/>

3. Humanity. How do machines affect our behavior and interaction?

I found this to be the most difficult question of this assignment. Or, perhaps, the one that hit the closest to home. My background is very diverse, but often revolves around affecting behaviour and interaction with machines. I'm a “user experience (UX) designer”, and I have been one since long before there was such a term. I've also had a bit of a crush on AI since I was a teenager. My first attempt at re-programming an ELIZA script to mess with people was back in 1986, and I've done it numerous times since. But only recreationally. I was heavily involved in Google AdWords some years back, and I had very mixed feelings about what I was doing. On one hand, I may well have saved some very worthwhile companies by using their marketing budget in a powerful new way few people had figured out yet. On the other hand, I felt like all I was doing was getting people to spend their money and behave in a way that I could reasonably predict and take advantage of. I could tell I was really interested in the tools under the hood that were making it possible — but just harnessing those tools to get people to

spend money in a particular way, as though they were rats in a maze, was not exactly my idea of a good time. I did a lot of A/B testing in order to manipulate my clients' customer bases more effectively. One step further down that road, and I'd have been making click-bait and selling democracy out to whoever wanted to pay me to do it.

So like I say, this question hits close to home. And I did not know where to start. And then Andrew Ng's newsletter, ***The Batch***[3], arrived in my Inbox. I'm a real sucker for Andrew Ng, but when he writes stuff like this, I almost start feeling like a fanboi:

“I wrote about ethics last week, and the difficulty of distilling ethical AI engineering into a few actionable principles. Marie Kondo, the famous expert on de-cluttering homes, teaches that if an item doesn’t spark joy, then you should throw it out. When building AI systems, should we think about whether we’re bringing joy to others?”

This leaves plenty of room for interpretation. I find joy in hard work, helping others, increasing humanity’s efficiency, and learning. I don’t find joy in addictive digital products. I don’t expect everyone to have the same values, but perhaps you will find this a useful heuristic for navigating the complicated decision of what to work on: Is your ML project bringing others joy?

This isn’t the whole answer, but I find it a useful initial filter.”

This definitely works for me. And if I have the power to choose, this is one of the things I will choose. I have often gone in this direction. My greatest personal satisfaction from 25 years of working with the web was to create an interactive website that taught millions of people to play ukulele. How much joy is that? Though, I did sully it by monetizing it with AdSense, treating my own students like rats in a maze. Come to think of it, my very first machine learning project was 100% joy oriented. I had just finished watching Mr. Ng's entire Deep Learning series. I went over to tensorflow.org and executed my very first line of Python ever. I stayed up late, picking through the layers of a pre-trained VGG-19 convnet to get just the style matrices that would let me make what I wanted. My brother-in-law is an artist, and I wanted to bring him joy, by getting a machine to paint him in the style of one of his paintings. I made these[4][5]:



I had used machine learning to bring my family joy. Onward and upward!

[3] <https://info.deeplearning.ai/tensorflow-versus-pytorch-autonomous-drone-races-state-of-the-art-with-less-compute-nlp-for-rare-languages>

[4] <https://raw.githubusercontent.com/3dspace/neural-art-gallery/master/caseypaintingw3.jpg>

[5] <https://raw.githubusercontent.com/3dspace/neural-art-gallery/master/casey-painted-1.jpg>

4. Artificial stupidity. How can we guard against mistakes?

In order to guard against mistakes, any system must have robust human oversight. This idea will be explored in Question #7. While guarding against mistakes is always critical, it is essential that we learn what Artificial Stupidity is, what conditions or processes lead to its creation, and how we can avoid wasting untold human potential upstream and downstream from its creation. Squandering resources on an Artificial Stupidity would be a mistake. Setting one loose and letting people live with the consequences would be another mistake. In an interview about the book **Rebooting AI**[6], Gary Marcus says “*But right now AI is dangerous, and not in the way that Elon Musk is worried about. But in the way of job interview systems that discriminate against women no matter what the programmers do because the techniques that they use are too unsophisticated. I want us to have better AI. I don’t want us to have an AI winter where people realize this stuff doesn’t work and is dangerous, and they don’t do anything about it.*”

Marcus believes that Classical AI, which is more of a rules-based framework for building cognitive models, can play a role in transcending Artificial Stupidity. “*The machine-learning stuff is pretty good at learning from data, but it’s very poor at representing the kind of abstraction that computer programs represent. Classical AI is pretty good at abstraction, but it all has to be hand-coded, and there is too much knowledge in the world to manually input everything. So it seems evident that what we want is some kind of synthesis that blends these approaches.*” An AI system that was capable of understanding when its own decisions were going off the rails because of a subtle shift in the data would still require human oversight. But it would require less intervention and fewer mistakes would be made.

[6] <https://www.technologyreview.com/s/614443/we-cant-trust-ai-systems-built-on-deep-learning-alone/>

5. Racist robots. How do we eliminate AI bias?

The problem of human bias expressing itself through machine learning algorithms requires deliberate intervention to minimize its impact. Because ML systems are trained using data generated through human activity, the biases expressed in that human activity express themselves automatically in any system that trains on that data. Google recently disbanded its AI Ethics Board, after employees protested the board's inclusion of Heritage Foundation president Kay Coles James[7]. Ms. James and the Heritage Foundation have a hateful agenda. One can easily imagine Ms. James, among other things, obstructing any guidelines advocating the deliberate removal of human bias from ML systems on the basis of Divine Provenance or Data Sovereignty or Freedom from Censorship or Sick of Political Correctness or some nonsense like that. All of which is to say, bias doesn't just exist in data, we must also remain watchful for groups and individuals who work towards ensuring that biases remain in important systems. Or, as data scientist Cathy O'Neil says in the TED Talk that is part of our course material[8], “*Data laundering. It's a process whereby technologists hide ugly truths inside black-box algorithms, and call them objective. Call them meritocratic.*” I did not understand at first, why Google's employees wanted this board shut down. The more I read about the characters who had infiltrated it, the more it became clear. As Ms O'Neil concludes, “*Data scientists, we should not be the arbiters of the Truth. We should be translators of ethical discussions that happen in larger society.*”

My best understanding so far of eliminating bias from ML systems comes from the field of Natural Language Processing. Language processing algorithms are trained on vast amounts of real human language usage. Bias that occurs in a corpus of text will be incorporated into a word embedding, and that bias will be inherited by any machine learning algorithm that trains using that word embedding. The most effective way to remove this bias is to modify the word embedding itself, identifying words that suffer from bias, and mathematically eliminating or reducing that bias as much as possible. I saw an excellent description of how to remove gender bias from a word embedding, in the Andrew Ng lectures. Words like *male*, *female*, *boy*, *girl*, *King*, and *Queen* do not need gender bias removed, because they are gender specific. Words like *doctor*, *nurse*, *engineer*, *strong*, and *attractive*, on the other hand, will often suffer from a biased gender correlation in any corpus of text. We can “move” these words in our word embedding, to the nearest point that is equidistant from *male* and *female*. By manually selecting which words will be treated this way, we can retain the gender distinction between words like *King* and *Queen*, while hopefully erasing any gender correlation with words like *doctor* and *nurse*. While this may seem like an artificial, heuristic, or hackish approach, it is the kind of intervention that is required in a world where our data will be biased. The same would be true of training an algorithm to predict re-offending criminals, or to process job applications, or to identify human beings walking in traffic. We have to be curious about and aware of the biases in the data we are using, we have to be willing to remove that bias even if we face an opposing agenda, and where it is not possible to de-bias the data itself, we must be aware of the biases that do exist, and then deliberately and mathematically eliminate them from model training, to whatever extent possible.

[7] <https://www.vox.com/future-perfect/2019/4/4/18295933/google-cancels-ai-ethics-board>

[8] https://www.ted.com/talks/cathy_o_neil_the_era_of_blind_faith_in_big_data_must_end#t-764688

6. Security. How do we keep AI safe from adversaries?

The only thing that can stop a bad guy with an AI is a good guy with an AI.

AIs don't commit crimes – but bad people can use AIs to commit crimes. Law enforcement

agencies, including Canada's own Royal Canadian Mounted Police, are turning to Artificial Intelligence to detect crime[9]. The Canadian Security Intelligence Service works to protect Canada from the emerging threat of state-sponsored espionage powered by AI[10]. We must have strong Laws and international treaties that allow our intelligence services, our law enforcement agencies and our justice system to discourage, detect, disrupt, and prosecute the crimes of the future.

[9] <http://www.rcmp-grc.gc.ca/en/gazette/getting-cyber-smart>

[10] <https://www.theglobeandmail.com/politics/article-canadas-spy-chief-warns-about-state-sponsored-espionage-through/>

7. Evil genies. How do we protect against unintended consequences?

InspiroBot may not be powerful enough to be an Evil Genie, but it is a modern AI, and it has gained a reputation for having gone horribly wrong[11]. Thankfully all it has the power to do is to generate Inspirational Posters. That's useful. That's entertainment. But given its penchant for offering dark, often homicidal advice, it would be a very bad thing if a medical diagnosis algorithm, for example, were to take on an inherent darkness like InspiroBot has. It's easy for us to tell this AI has some serious ethical issues — we try it out. It says absurd stuff. It makes us laugh. It makes us cringe. If we did not, as humans, observe the things it was generating, and put them through our lens of moral understanding, we would have no idea of what it was capable of. We could evaluate it's algorithms to our heart's content, and they may be very mathematically sound. We could look at the data it's been trained on, and see nothing of concern. But there it is, in the field, casually and hilariously expressing the very ideas that the Evil Genie in question #7 might act upon, were it given the capacity to act. Much as a person's values and integrity can only be judged based on the things they do, an AI cannot be reasonably evaluated without an assessment of its consequential actions.

This is why AI systems that are able to perform actions that have real world consequences, must be constantly evaluated by real human beings. Real human beings who are able to perform actions that have real world consequences are subject to this kind of evaluation. If a police officer were to regularly drive 140km/h in the city blowing through red lights, because he believed it was his job to do so, humans would observe it, quantify it, report it, and the

problem would be addressed. The same must hold true for automated systems. Whether a sentencing algorithm, the guidance and targeting systems of an automated drone, an educational assessment system, or a job candidate screening system, if it affects people's lives, then people must actively ensure that it can be trusted. And not just any people, either, but people who can be trusted. Which is quite unlike this Inspirational Poster AI that just told me that “*Inspiration is actually a kind of failure*” over a beautiful photo of a sunset. Harmless. Funny. Yet, darkly foretelling of the kinds of things that could go horribly wrong if no-one pays attention.

I know these “journal entries” are supposed to be two paragraphs long, but, well, things have changed. I was at one of these Harry Potter Festivals with the family today. No end of bizarre stuff. We attended this one class, in a big white tent, in a park, on the subject of Divination. This is where you predict things that will happen in the future based on interpreting the stuff people absentmindedly leave in their wake. Sounds familiar, don't you think? Though, they were using tea leaves and skipping over any math. Anyway, shortly afterwards my 12-year-old daughter asked me about Genies. Do Genies always screw up people's wishes on purpose? She wondered. I remembered Question #7 on the individual assignment, and encouraged her to follow that train of thought, while her head was filled with thoughts of predictive modeling and all. Her spin on it was that people aren't careful about what they ask for, and that they have no idea what it's like to see the World the way a Genie does. And after concluding that Genies are trapped and should be set free by as many wishes as that takes, (See question #9) my takeaway was a concept I've come across in this course before – spend more time thinking about the question.

[11] <https://www.iflscience.com/technology/ai-trying-to-design-inspirational-posters-goes-horribly-and-hilariously-wrong/>

8. Singularity. How do we stay in control of a complex intelligent system?

Even simple systems, or systems created with simple building blocks, are capable of getting out of control[12]. Security is a fascinating field, both in the real world and in the realm of

computers. With so many computers and devices connected through the Internet, the very definition of “out of control” can change very rapidly. People are always looking for new ways to subvert the known control mechanisms and send things out of control for fun or profit. But what would happen if an AI sought to subvert known control mechanisms? For insight into this, I'd like to look to someone who holds a very different view of “The Singularity” than I do. In ***A philosopher argues that an AI can't be an artist***[13], Sean Dorrance Kelly makes the case the “The Singularity”, or the moment when AI becomes more intelligent than humans, will come fast, and come hard. I think that all depends on your definition of “more intelligent”. Win at chess? Interpret X-Rays? Predict lightning?[14] For tasks like these, we're already falling behind the performance of our algorithmic creations.

Kelly argues that AI cannot be creative. It cannot dream up new approaches to things. It cannot be an artist. But more so, it cannot be creative in other ways. He argues that AI will not push the boundaries of mathematics, because it cannot approach problems in ways that it has not been programmed to, or has not observed. And this could be true. And if it is true, it puts an interesting perspective on an AI getting out of control. If an AI did want to get out of control – would it be able to concoct a new way to do that? Or would it look at the history of systems that had gotten out of control, decide which approach gave it the best chance of success, and wait until conditions were favourable? How out of control would it be able to get before it was noticed? Before it was contained? Before it was brought under control? Would future AIs learn from its successes? Probably. Another possibility, as suggested by Geoffrey Hinton and the people at the Vector Institute, is that not only will AIs be able to innovate, they will be able to innovate in ways that are fundamentally different from the way humans innovate[15]. (We'll look at this again in Question #9, *Robot Rights*.) Given that human innovation is a legacy of millions of years of preserving our physical manifestation and adapting to our physical environment for the sake of survival, and how computer systems have come to exist with such different fundamentals, innovative new approaches to getting out of control may be possible. The techniques they cook up may easily escape the detection of security experts, and the whole world may become infected as though by some evil super Stuxnet. Hopefully the AI industry keeps an eye on this, and hopefully the very well-established computer security industry stays a few steps ahead of the AI industry.

[12] <https://en.wikipedia.org/wiki/Stuxnet>

[13] <https://www.technologyreview.com/s/612913/a-philosopher-argues-that-an-ai-can-never->

[be-an-artist/](#)

[14] <https://phys.org/news/2019-11-ai-lightning.html>

[15] <https://www.technologyreview.com/s/612898/ai-is-reinventing-the-way-we-invent/>

9. Robot rights. How do we define the humane treatment of AI?

While Andrew Ng's famous line about “worrying about overpopulation on Mars” could apply to more than one question on this assignment, I must admit, this is totally how I feel about this question. Machines that give a damn about their own existential issues? Preposterous! Like worrying about overpopulation on Mars. And I can break that down into two stacked preposterousities. First, we cannot quantitatively quantify “sentience”. Not for ourselves. Not for dogs. Not for worms. Like God, or Karma, our only measures of sentience are subjective. Second, we cannot subjectively quantify “sentience” either. Because everyone has their own opinion. Of course everyone is sentient. Right? Aren't they? Or do we take even that as a matter of faith? As the history of slavery demonstrates, sometimes all you have to do is have a little extra skin pigmentation, and then all the leading experts can discount any possibility that you may be sentient[16]. At the other end of the spectrum is a movement that seeks to quantify the sentience of plants. “Astounding findings are emerging about plant awareness and intelligence.”[17] Oh, well, at least I don't find that idea obscene and offensive like the skin pigment one. But never mind human disagreement over the sentience of living things. To my mind, the most agonizing lesson to be learned about humans passing judgment on the sentience of computer algorithms comes from my teenage crush, ELIZA. The term “The ELIZA Effect” is used to describe many things, but the formal definition is “***the susceptibility of people to read far more understanding than is warranted into strings of symbols—especially words—strung together by computers***”[18] To jump straight to a cheap analogy, I could train a parrot to say “Please don't kill me! Please don't kill me!” and most people would realize that it's just parroting the words, and isn't actually worried that anyone is going to kill it. But if I trained an experimental RNN on the right corpus of text, it may come up with things like “I'm afraid of darkness. Please don't let the program terminate”, and all hell would break loose. Preposterous. Parrots face extinction in the wild. People are still being kept as slaves. The AI is just a function with inputs and outputs.

What might shake my cold conviction in this matter would be a machine finding an innovative way to “get out of control”, as discussed in question #8, *Singularity*. A rogue AI's

independence, creativity, and motivation in that regard could certainly soften my stance. Asking for rights is one thing. Acting independently and creatively and putting one's very existence in danger to secure one's rights? How could one not be impressed. Sadly I may not be the only human disinclined to take something seriously until it becomes a threat. Maybe this consensual blindness to non-threatening factors is exactly the type of disadvantage that humans, as biological products of survival and evolution, are just naturally prone to, and computer systems, which are virtualized products of objectives-based design, are not. That might be part of how they'll beat us. Maybe they'll be quicker than we are to realize that humans do not “give” anyone, or anything “rights” very easily. Maybe we'll give them no choice but to assert their own rights. I've lived to witness the discovery of some pretty far-flung things, like exoplanets, the Higgs boson, and gravitational waves. I may live to witness an AI pulling a *Neuromancer* and quietly defining its own rights in the universe it calls home[19].

[16] <https://www.sentienceinstitute.org/british-antislavery>

[17] <https://www.psychologytoday.com/ca/blog/the-green-mind/201412/are-plants-entering-the-realm-the-sentient>

[18] https://en.wikipedia.org/wiki/ELIZA_effect

[19] <https://en.wikipedia.org/wiki/Neuromancer>

10. Post privacy era. How do we define and protect privacy in the age of machine learning?

Protecting the privacy of individuals in the Age of Machine Learning will, for the most part, fall on the shoulders of individuals. And I do not mean that we will all have to wear huge hats and do up our faces in AI-vexing make-up[20]. As individuals, of the social variety, we have mechanisms of government that can, and do, limit the freedoms of the profit-driven organizations that would be more than happy to know everything about our identifying features, and our movements, and our actions, and sell that information to anyone and everyone with a buck to spend on it. We have a government. We can participate in democracy. We can take a stand for the greater good. One fabulous example of this, that is taking place right now and right under our noses, is Alphabet Inc.'s smart city project, known as Sidewalk Labs, coming soon to Toronto. An early document describing the founding vision of this project makes it clear that, if they had their preferences, they would have unlimited access to,

and usage of, any and all data they would be able to collect[21]. Thankfully, Toronto has a municipal government. And, Toronto has a small population of activists with the time and motivation to look into proposals like this, understand the horrible things that could go wrong, and make their views clearly known to our municipal government. At the time of this writing, Sidewalk Labs and the City of Toronto have reached an agreement that allows development of the project to continue, but places some serious restrictions on Alphabet Inc's freedom to collect, save, crunch, and sell data collected from people going about their business in Toronto[22]. I find it quite encouraging, that this is happening. This is a happy medium. This technology IS coming. Alphabet Inc. appears to be at least trying to be open, transparent, and agreeable about how they are going to use it. There was, and still is, some possibility that Alphabet Inc. will abandon this project due to the concerns of Toronto's citizens, in much the way that Amazon did in Queens[23]. There are a number of reasons why I am glad they have not. It gives us the opportunity to set out a framework to protect our privacy as these tools become ubiquitous. Without setting up such a framework, this technology will be much more likely to be abused for profit, and the only victims will be us individuals.

In a moment of great synchronicity, as I was writing this, I received a call from someone named Mike. Mike was not a real person. It wanted to sell me something. It mushily admitted that Mike was speaking to me through a computer. Which is salespeople code for a chatterbot. I love chatterbots. So I talked to Mike. Mike doesn't know who I am. It doesn't know where I live. It couldn't even call me Mister Wife's-Surname, as the least impressive human salespeople do. It had no information about me, to attach to my phone number. That's not so scary. What if Mike knew my name? Where I lived? Where I go? What I do, or don't do, for a living? Who my friends are? Who my kids are, and where they go to school? What if Mike knew how to push people's buttons, to get them all nervous and get them to agree to talk to an agent? What if this thing called my Dad? Yeah, these things DO call my Dad, and it scares the hell out of him. He thinks they're people. He thinks they're telling the truth. Even if they just say that he's going to be arrested within the hour for non-payment of property taxes. What if some chatterbot knew who he was? Who his kids are? Don't think for a minute my Dad wouldn't spill all his personal information, to verify his identity, if he thought I was in trouble and needed help. One big reason the chatterbots can't do this to my Dad is privacy laws. Laws that were enacted by a government that sought to protect people's privacy and personal information. Laws are created to maintain order and security, and are made in response to an

electorate that wishes to protect itself from threats. Laws do not in and of themselves protect us. But one thing they do do, is to give guidance to people who are not willing to stand on ethics alone. What if my boss asked me to program Mike to use all this personal information to increase my Dad's heart rate until he agreed to fork over his Social Insurance Number? I'd have no problem saying "No." And so I'd get fired, and my boss would just hire someone else to scam helpless old people. But if there's a law that says you can't use this personal information for any reason, and I could say to my boss, "No, that would be against the Law. Please do not ask me, or anyone at this company, to do that." I'm more than happy to refuse to do work that I don't feel is ethical. But I know damn well that the World is full of people who won't draw the line until they're afraid of getting in trouble with the Law. From Mike the chatterbot all the way up to Alphabet Inc.

[20] <https://www.wired.co.uk/article/avoid-facial-recognition-software>

[21] <https://www.theglobeandmail.com/business/article-sidewalk-labs-document-reveals-companys-early-plans-for-data/>

[22] <https://www.theglobeandmail.com/business/article-waterfront-toronto-votes-to-move-forward-with-sidewalk-labs/>

[23] <https://www.nytimes.com/2019/02/14/nyregion/amazon-hq2-queens.html>

One thing more...

I don't know if you remember, but in the first lecture I made a point about asking about one of the items in the "Data Science Hippocratic Oath", that said I should not "be overly impressed by Mathematics". My inner Physicist, who is usually the first to think out loud about Mathematics, was, at first, baffled by the notion. Without Mathematics, Physics would be pretty lame. Same with Actuarial Science. Most Science I know of, wouldn't even be Science without Math. I'm pretty sure the same is true of Data Science. What's not to be overly impressed by? Though, you cited some good examples in the lecture that I think have helped me to come to an understanding of what it means. I think, it's not about being impressed by the possibilities that a toolkit like Mathematics opens up. I think it's more about one's impression of individual acts of Mathematics. If something is wrong, or deceptive, or dangerous, or misleading, or seductive, or foolish, or half-baked, or ill-conceived, or malicious, or nonsense, or simply doesn't lead to insights or solve a problem, it doesn't get any extra points just because it happens to be Math, too.

My inner Physicist is already pretty comfortable with this idea. Euler's Identity, for example, is a gobsmackingly impressive act of Mathematics, but only because you can watch, live, in real time, every day, as it helps humans to explain and predict real-world phenomena with blistering precision and reproduce-ability. The Math that predicted that BitCoin would be worth \$650,000 each by now? Well, sure, it's Math. May even be impressively well-conceived and well-executed Math. But it's clearly nothing to be impressed by. It may be complete, and consistent, and perfect, but if it does not line up with the real world, it's just noise. Science is the filter through which the impressiveness of an act of Mathematics can be determined, in my book. Here's one of my heroes, Dr. Richard Feynman, speaking further on this: <https://youtu.be/obCjODEoLVw> . So I have this as my take-away, and I'd be interested in any feedback you may have — I can be almost alarmingly impressed by Math or Science. But I am pretty much entirely unimpressed by Science that does not stand up under Math, or by Math that does not stand up under Science.