# Ethics in Data

## A definition of Ethics

* The basic concepts and fundamental principles of decent human conduct. It includes study of universal values such as the essential equality of all men and women, human or natural rights, obedience to the law of land, concern for health and safety and, increasingly, also for the natural environment. See also, morality.

## What do we mean by ethics?

* The application of general ethical principles to the actions and decisions of businesses and the conduct of their personnel.
* Not materially different from ethical principles in general because business actions must be judged in the context of society’s standards of right and wrong.



### Ethical universalism

* A common understanding among multiple cultures and countries about what constitutes right and wrong. This gives rise to ethical standards for all firms and all people.
* In business this translates to: and understanding whether a business related action is right or wrong is judged by universal standards. You will take the same action no matter where you are doing business in the world

### Ethical Relativism

* Different beliefs, customs, and behavioral norms, across countries and cultures give rise to multiple sets of standards of what is ethically right or wrong. As companies increasingly operate globally. Whether business actions are right or wrong are based on local ethical standards. In short while bribes may not be acceptable in Canada it might be in other places.

## Candidate Ethical Principles

**Golden Rule**

* Do unto others as your would have them unto you

**Immanuel Kant’s categorical imperative**

* If an action is not right for everyone to take it is not right for anyone

**Descartes’ rule of change**

* If an action cannot be taken repeatedly, it is not right to take at all.
* This is saying you cant change your action based on different seasons of the year

**Utilitarian Principle**

* Take the action that achieves the higher or greater value (for the greatest number)

**Risk aversion principle**

* Take the action that produces the least harm or potential cost

**Ethical no free lunch rule**

* Assume that virtually all tangible and intangible objects are owned by someone unless there is a specific declaration otherwise.

## Moral vs Business Case

**The moral case for ethics**

* A strategy that is unethical is morally wrong and reflects badly on the other character of the firm’s personnel

**The Business Case for Ethics**

* An ethical strategy can be both good business and serve the self-interest of shareholders

## Ethics in Information Systems

* Ethics refers to the principles of right and wrong that individuals, acting as free moral agents, use to make choices to guide their behaviors. Information systems raise new ethical questions for both individuals and societies because they create opportunities for intense social change and, thus, threaten existing distributions of power, money, rights and obligations. Like other technologies, such as steam engines, electricity, the telephone, and the radio, information technology can be used to achieve social progress, but it can also be used to commit crimes and threaten cherished social values. The development of information technology will produce benefits for many and costs for others.

## How to think about Ethical, Social, & Political Issues

* Society as a calm pond
  + Clear expectations of right and wrong
* Technology as a rock dropped in the pond, creating ripples of new situations not covered by old rules
* It make take years to develop etiquette, expectations, laws
  + Social and political issues cannot respond overnight to these ripples

## Key Technology Trends and Ethical Issues

* Moore’s Law
  + 1970s Gordon Moor was the cofounder of intel and made a prediction that computer power would double every 18 month to 2 years. This has remained true to today. Computer continue to get faster and more powerful. More capable software etc. Data storage costs are decreasing.
* Data storage costs declining
* Data analysis capability advancements
  + Machine learning
* Networking capabilities increasing

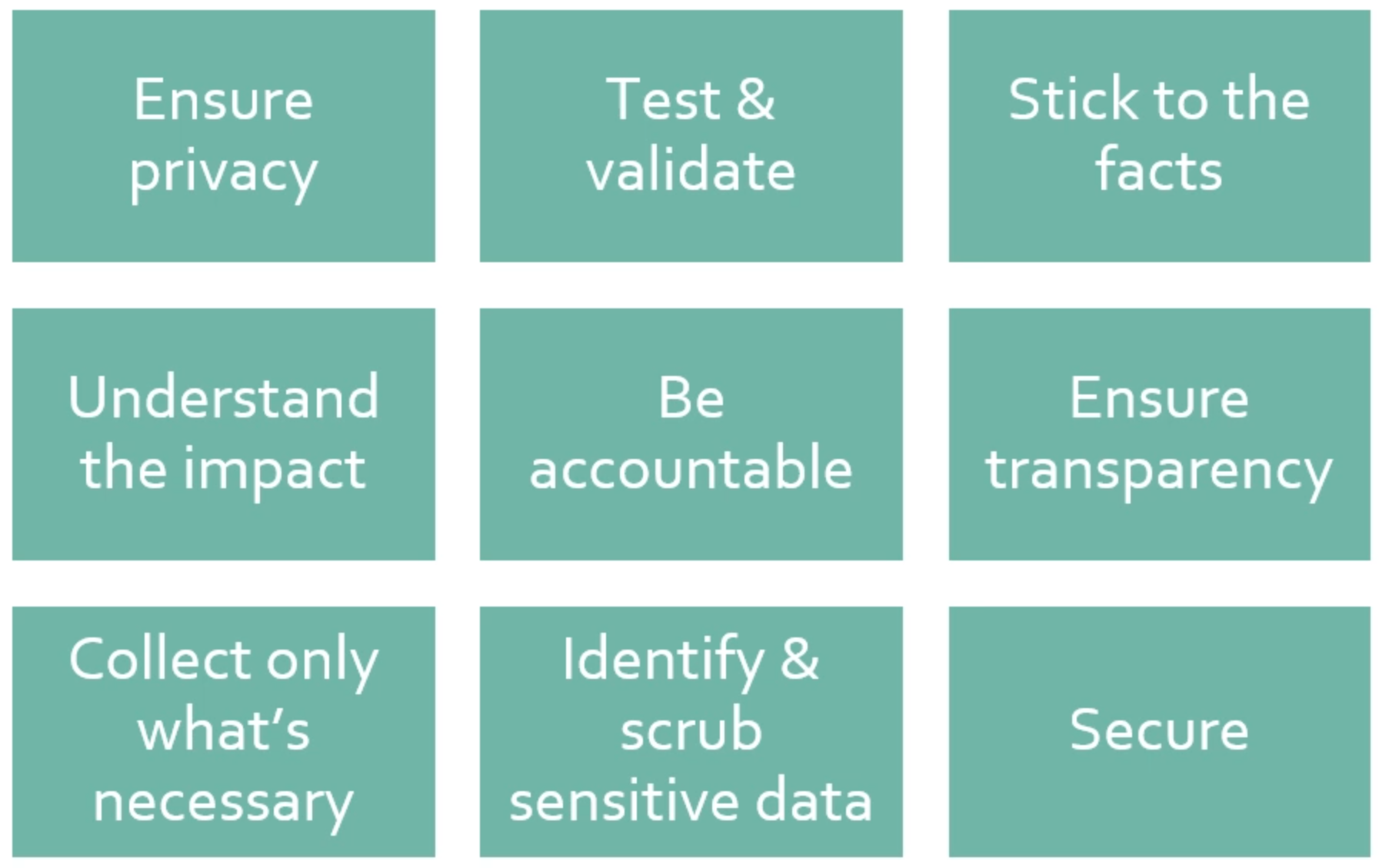
**Question:**

Just because we can should we?

## Data Ethics Considerations

* Intentionality
  + What do you intend to use the data for?
  + Is it ethical?
  + Once you decided what and how it will be used for and who will be accountable for the congoing use?
  + You can publish and share the insight you gained. However, other people can take that and make their own decisions.
* Accountability
  + Who is going to be accountable for the decisions?
* Consent & Choice
  + Do you have consent for the info you have?
  + Does the person who is sharing the information with you know what you will be using it for?
* Transparency
  + Is there transparency in the data so you can get into the data and understand how it is being summarized? insight being provided?
* Privacy & Protection
  + Have you done everything you needed to do in order to protect the data you are working with?
  + Is access permission based?
  + Is data encrypted?
* Respect
  + Need to have a healthy respect for the data impact of our analysis
* Legislation
  + Do we comply with all necessary legislation for Canadian Privacy and is data residency required?

## Your responsibility



* Don’t always assume that the data you are accessing is always correct
* Be critical of data that your accessing and make sure you can test and validate and verify it.

**Quote:**

If you torture you data you can make it say whatever you want.

* Keep your biases out of your analysis

## Why does this matter to you?

In context of the course?

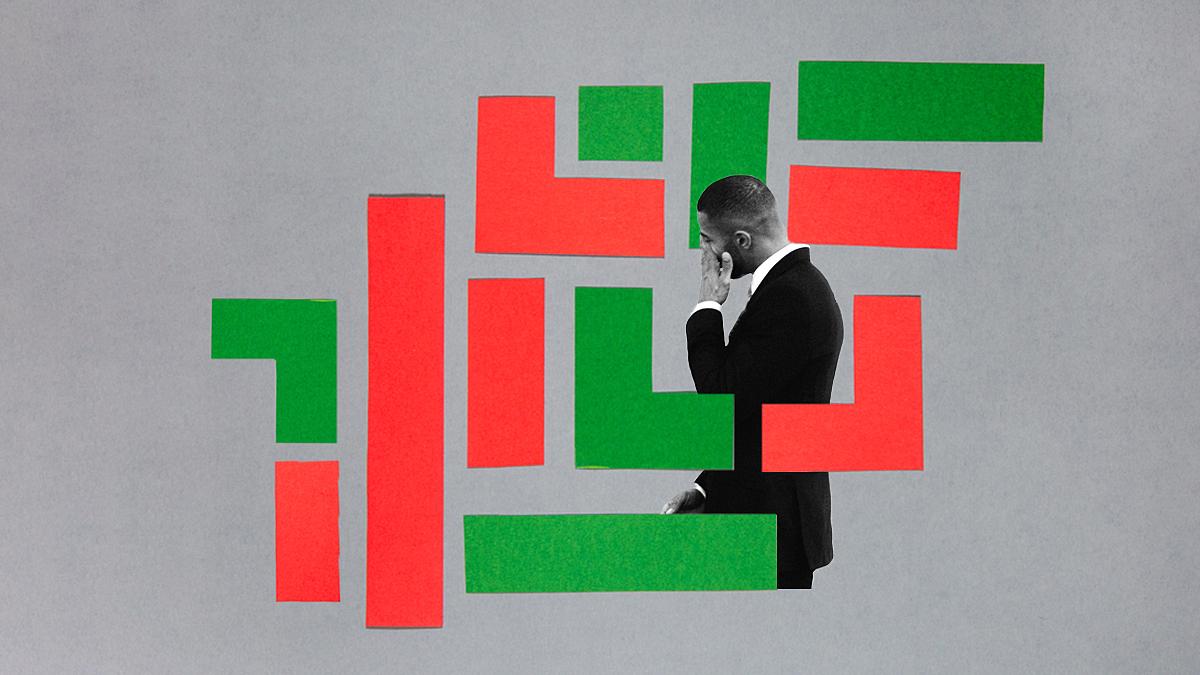
In context of your life and career?

# When Does Predictive Technology Become Unethical?

by

* [Eric Siegel](https://hbr.org/search?term=eric%20siegel)

October 23, 2020



HBR Staff/KKGAS/Sidney Morgan/Stocksy

**Summary.**What happens when algorithms can predict sensitive things about you, such as your sexual orientation, whether you’re pregnant, whether you’ll quit your job, and whether you’re likely to die soon? We’re not talking about mishandling, leaking, or stealing data. Rather,...more

Machine learning can ascertain a lot about you — including some of your most sensitive information. For instance, it can predict your sexual orientation, whether you’re pregnant, whether you’ll quit your job, and whether you’re likely to die soon. Researchers can [predict race based on Facebook likes](https://www.cam.ac.uk/research/news/digital-records-could-expose-intimate-details-and-personality-traits-of-millions), and officials in China use facial recognition [to identify and track the Uighurs](https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html), a minority ethnic group.

Now, do the machines actually “know” these things about you, or are they only making informed guesses? And, if they’re making an inference about you, just the same as any human you know might do, is there really anything wrong with them being so astute?

Let’s look at a few cases:

In the U.S., the story of [Target predicting who’s pregnant](https://www.predictiveanalyticsworld.com/machinelearningtimes/how-target-gets-the-most-out-of-its-guest-data-to-improve-marketing-roi/6815/) is probably the most famous example of an algorithm making sensitive inferences about people. In 2012, a New York Times story about how companies can leverage their data included an anecdote about a father learning that his teenage daughter was pregnant due to Target sending her coupons for baby items in an apparent act of premonition. Although the story about the teenager may be [apocryphal](https://www.predictiveanalyticsworld.com/machinelearningtimes/target-really-predict-teens-pregnancy-inside-story/) — even if it did happen, it would most likely have been coincidence, not predictive analytics that was responsible for the coupons, according to Target’s process detailed by *The New York Times* story — there is a real risk to privacy in light of this predictive project. After all, if a company’s marketing department predicts who’s pregnant, they’ve ascertained medically sensitive, unvolunteered data that only healthcare staff are normally trained to appropriately handle and safeguard.

#### INSIGHT CENTER

###### [AI and Equality](https://hbr.org/insight-center/ai-and-equality)

Designing systems that are fair for all.

Mismanaged access to this kind of information can have huge implications on someone’s life. As one concerned citizen [posted online](https://books.google.com/books?id=dGgs3JKUFQIC&pg=PA64&lpg=PA64&dq=imagine+that+a+pregnant+womans+job+is+shaky,+and+the+state+disability+isnt+set+up+right+yet&source=bl&ots=9EUg63i8CL&sig=ACfU3U2kOhMb2a2o3hy8olQel8MevI6CBg&hl=en&sa=X&ved=2ahUKEwjwwP2KwcjsAhVLJt8KHQceDfgQ6AEwAHoECA4QAg#v=onepage&q=imagine%20that%20a%20pregnant%20womans%20job%20is%20shaky%2C%20and%20the%20state%20disability%20isnt%20set%20up%20right%20yet&f=false), imagine that a pregnant woman’s “job is shaky, and [her] state disability isn’t set up right yet…to have disclosure could risk the retail cost of a birth (approximately $20,000), disability payments during time off (approximately $10,000 to $50,000), and even her job.”

This isn’t a case of mishandling, leaking, or stealing data. Rather, it is the generation of *new* data — the indirect discovery of unvolunteered truths about people. Organizations can predict these powerful insights from existing innocuous data, as if creating them out of thin air.

So are we ironically facing a downside when predictive models perform *too well*? We know there’s a cost when models predict incorrectly, but is there also a cost when they predict *correctly*?

Even if the model isn’t highly accurate, per se, it may still be confident in its predictions for a certain group of pregnant individuals. Let’s say that 2% of the female customers between age 18 and 40 are pregnant. If the model identifies customers, say, three times more likely than average to be pregnant, only 6% of those identified will actually be pregnant. That’s a lift of three. But if you look at a much smaller, focused group, say the top 0.1% likely to be pregnant, you may have a much higher lift of, say, 46, which would make women in that group 92% likely to be pregnant. In that case, the system would be capable of revealing those women as very likely to be pregnant.

The same concept applies when predicting sexual orientation, race, health status, location, and your intentions to leave your job. Even if a model isn’t highly accurate in general, it can still reveal with high confidence — for a limited group — things like sexual orientation, race, or ethnicity. This is because, typically, there is a small portion of the population for whom it is easier to predict. Now, it may only be able to predict confidently for a relatively small group, but even just the top 0.1% of a population of a million would mean 1,000 individuals have been confidently identified.

It’s easy to think of reasons why people wouldn’t want someone to know these things. As of 2013, [Hewlett-Packard was predictively scoring its more than 300,000 workers](http://analytics-magazine.org/predictive-analytics-the-privacy-pickle-hewlett-packards-prediction-of-employee-behavior/) with the probability of whether they’d quit their job — HP called this the Flight Risk score, and it was delivered to managers. If you’re planning to leave, your boss would probably be the *last* person you’d want to find out before it’s official.

As another example, facial recognition technologies can serve as a way to track location, decreasing the fundamental freedom to move about without disclosure, since, for example, publicly-positioned security cameras can identify people at specific times and places. I certainly don’t sweepingly condemn face recognition, but know that CEO’s at both [Microsoft](https://www.geekwire.com/2020/microsoft-ceo-satya-nadella-warns-fallout-without-strong-national-law-facial-recognition/) and [Google](https://www.huffpost.com/entry/facial-recognition-google_n_869583) have come down on it for this reason.

In yet another example, a consulting firm was modeling employee loss for an HR department, and noticed that they could actually model employee deaths, since that’s one way you lose an employee. The HR folks responded with, “Don’t show us!” They didn’t want the liability of potentially knowing which employees were at risk of dying soon.

[Research](https://www.cam.ac.uk/research/news/digital-records-could-expose-intimate-details-and-personality-traits-of-millions) has shown that predictive models can also discern other personal attributes — such as race and ethnicity — based on, for example, Facebook likes. A concern here is the ways in which marketers may be making use of these sorts of predictions. As Harvard professor of government and technology Latanya Sweeney [put it](https://news.harvard.edu/gazette/story/2013/04/seeking-fairness-in-ads/), “At the end of the day, online advertising is about discrimination. You don’t want mothers with newborns getting ads for fishing rods, and you don’t want fishermen getting ads for diapers. The question is when does that discrimination cross the line from targeting customers to negatively impacting an entire group of people?” Indeed, a [study](https://news.harvard.edu/gazette/story/2013/04/seeking-fairness-in-ads/) by Sweeney showed that Google searches for “black-sounding” names were 25% more likely to show an ad suggesting that the person had an arrest record, even if the advertiser had nobody with that name in their database of arrest records.

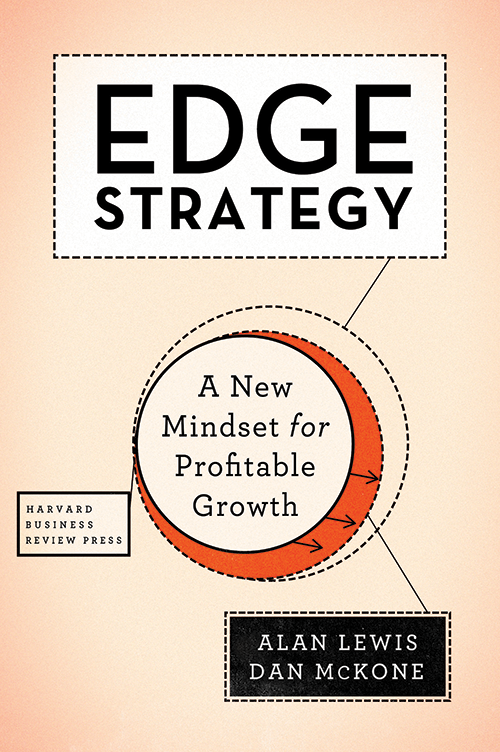
“If you make a technology that can classify people by an ethnicity, someone will use it to repress that ethnicity,” [says Clare Garvie](https://www.nytimes.com/2019/05/15/business/facial-recognition-software-controversy.html), senior associate at the Center on Privacy and Technology at Georgetown Law.

Which brings us to China, where the government applies facial recognition to [identify and track members of the Uighurs, an ethnic group](https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html) systematically oppressed by the government. This is the first known case of a government using machine learning to profile by ethnicity. This flagging of individuals by ethnic group is designed specifically to be used as a factor in discriminatory decisions — that is, decisions based at least in part on a protected class. In this case, members of this group, once identified, will be treated or considered differently on the basis of their ethnicity. One Chinese start-up valued at more than $1 billion said its software could recognize “sensitive groups of people.” Its website said, “If originally one Uighur lives in a neighborhood, and within 20 days six Uighurs appear, it immediately sends alarms” to law enforcement.

Implementing the differential treatment of an ethic group based on predictive technology takes the risks to a whole new level. Jonathan Frankle, a deep learning researcher at MIT, [warns that this potential extends beyond China](https://www.vanityfair.com/news/2019/04/china-created-a-racist-artificial-intelligence-to-track-muslims). “I don’t think it’s overblown to treat this as an existential threat to democracy. Once a country adopts a model in this heavy authoritarian mode, it’s using data to enforce thought and rules in a much more deep-seated fashion… To that extent, this is an urgent crisis we are slowly sleepwalking our way into.”

It’s a real challenge to draw the line as to which predictive objectives pursued with machine learning are unethical, let alone which should be legislated against, if any. But, at the very least, it’s important to stay vigilant for when machine learning serves to empower a preexisting unethical practice, and also for when it generates data that must be handled with care.

#### Readers Also Viewed These Items

* 

#### [Edge Strategy: A New Mindset for Profitable Growth](https://hbr.org/product/edge-strategy-a-new-mindset-for-profitable-growth/15026?ab=at_art_idp_v1x2_s01)

Book

[Buy Now](https://hbr.org/product/edge-strategy-a-new-mindset-for-profitable-growth/15026?ab=at_art_idp_v1x2_s01)

# A Modern Trolly Problem: Smart Cars and Ethical Programming

## Link

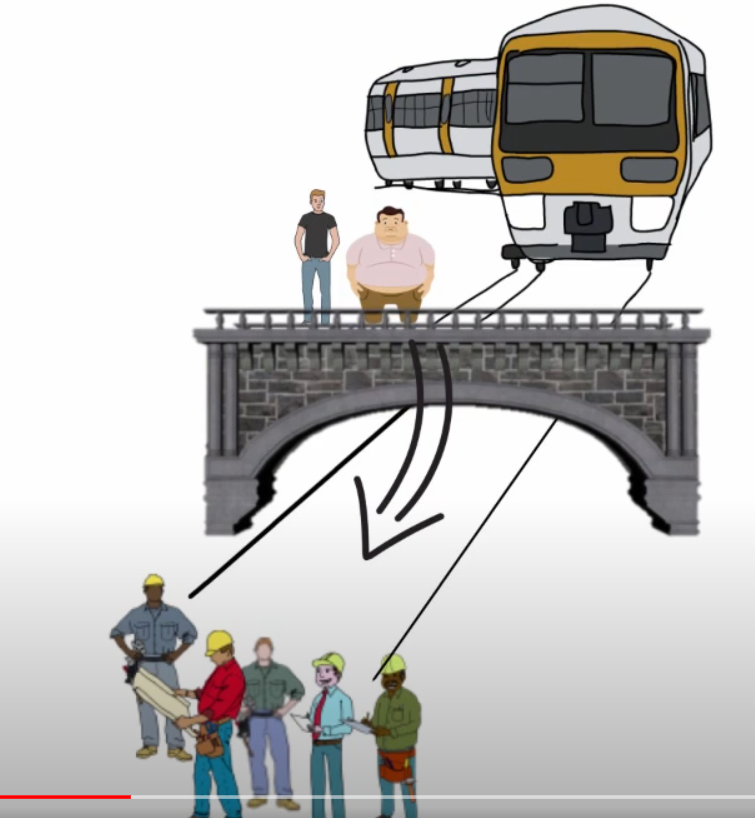
https://www.youtube.com/watch?v=w6AtGjf5wtc

## Notes

* ethical dilemma on a trolly program.
* Would you pull a lever to change the train track to hit 5 people or 1 person?



* 90% said they would pull the lever to sacrifice 1 to save 5
* What if you are on a overpass. The only way to stop the train is to push the fat man off the bridge. What would you do?



* 10% would push the man off the bridge
* people see that it is ok to pull the lever but not ok to push the man. The fatman variant is not realistic.
* Ethics is not a utilitarian choice. They focused on gender, race etc that ethical dilemmas are a sticky subject.
* More types of transportation
  + Driverless Cars - how can it handle an ethical dilemma?
  + Would it be ethical to program the car differently depending on the passenger?