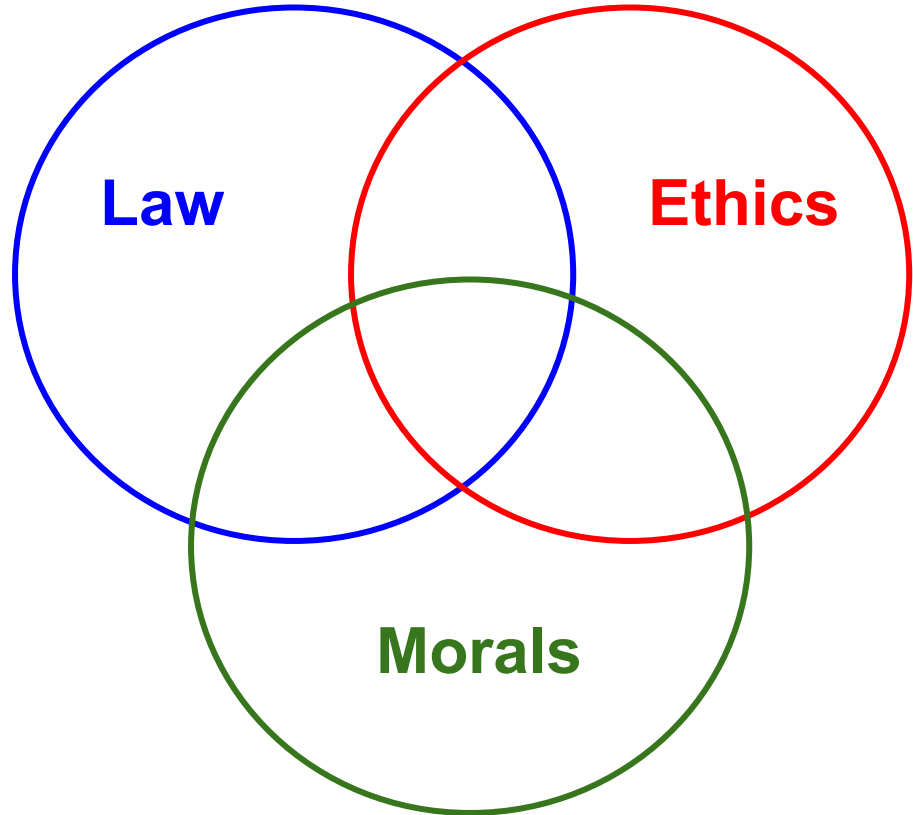


Ethics in CS



Chiara Shah
Emma Willard School
cshah@emmawillard.org

Saturday, April 9th, 2022

NY Celebration of Women in
Computing:
<https://nycwic.acm.org/>



“If this happens, let’s program the car to move 2 inches away from the truck, to reduce risk of collision. I don’t like cyclists anyway!”

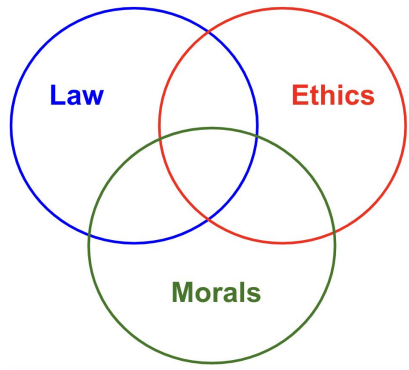
Morality vs. Ethics vs. Law

- Morals
 - individual guiding principles
 - ideas about how to be good
- Ethics
 - guidelines that govern correct behavior in a group (norms)
 - workplaces, schools, religious organizations
- Law
 - regulations that are enforced by a government on its people

Individual *morals* can inform (but not always) a code of *ethics*, which in turn can (but not always) inform *laws*

Morality vs. Ethics vs. Law

	Perspective	Influences	Punishment
Morals	Personal	Will / build it?	Guilt
Ethics	Group	What can we build? Should we build it?	Expulsion Isolation
Law	Government	How can we legally build it?	Fines Prison Community Service



Recent Thoughts From Higher Education

Embedding ethics in computer science curriculum: Harvard initiative seen as a national model

(Harvard School of Engineering and Applied Sciences, 01/29/2019)

- “Embed ethics **across the curriculum** helps computer science students see how **ethical issues can arise from many contexts**, issues ranging from the way social networks facilitate the spread of false information to censorship to machine-learning techniques that empower statistical inferences in employment and in the criminal justice system.”
- Repository of Open Source Course Modules - <https://embeddedethics.seas.harvard.edu/>

Teaching CS Through Science Fiction

[How to Teach Computer Ethics through Science Fiction](#)

(Communications of the ACM, August 2018, Vol. 61 No. 8, Pages 54-64)

- [Video of Dr. Goldsmith Discussing Using Fiction to Teach Computer Ethics](#)
- [What Sci-Fi Can Teach Computer Science About Ethics](#)
(Wired, 08/26/2019)
- [How to Teach Computer Science Ethics Through Science Fiction](#) (Communications of the ACM, August 2018, Vol. 61, No. 8)

Teaching CS Through Science Fiction

“We have started to see [...] engagement through a number of initiatives in the computer science community”

- the International Joint Conference on Artificial Intelligence 2015 letter on autonomous weapons research and,
- the 2017 follow-on letter signed by CEOs of tech companies around the world
- ACM statement on algorithmic accountability
- development of the IEEE standard for algorithmic bias considerations
- new conferences and research groups focused on fairness, accountability, and transparency
- conferences focusing on the effect of artificial intelligence on society.

VIDEO: Disturbing simulation shows the power, terror of Killer robots

<https://abc7news.com/killer-drone-viral-video-drones-simulation-facebook/2664856/>

Teaching CS Through Science Fiction

- “fiction often removes the intellectual and emotional resistance some students might at first feel towards the subject of ethics.”
- “Teaching this way will, we hope, lead to an openness and exchange of ideas about both core values and best practices.”
- “Computer science is a field in which everyday practice and problem solving takes place in a context that could barely be imagined the decade before. Educators cannot predict the ethical quandaries their students will face.”
- “The real world is messy, and rules that may seem reliable under one set of conditions can falter under others.”

Teaching CS Through Science Fiction: Example

Law enforcement robots: Should they ever be allowed to shoot a human?

It is not difficult to imagine scenarios in which shooting a person, perhaps even lethally, can be expected to save the lives of others.

- But how should a robot calculate the risks and values at stake in such a scenario?
- What sorts of input should it use when ascertaining if it should shoot a human?
- What sorts of input should it ignore?
- And what are the social costs or benefits of using robots that will shoot a human under certain circumstances?

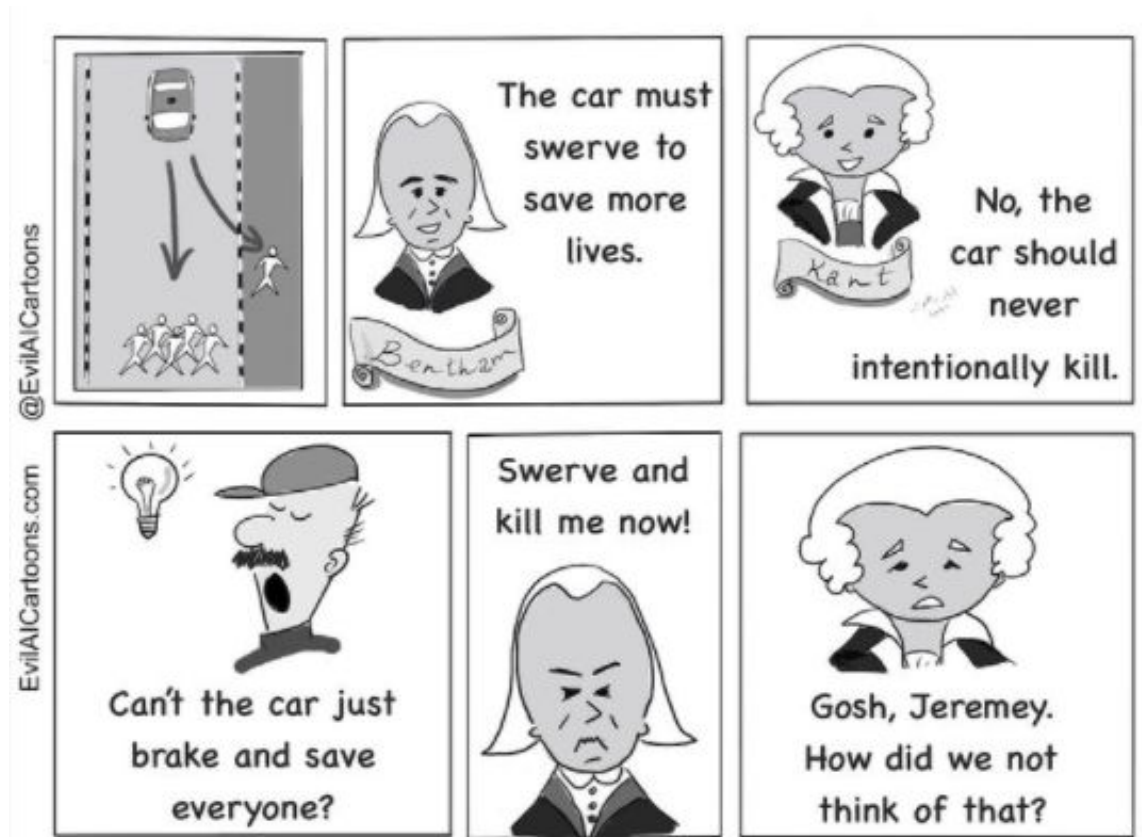
Teaching CS Through Science Fiction: Another Example

Autonomous Cars:

A self-driving car has brake failure.

Who should the AI save?

The car's occupant or a pedestrian(s)?



Teaching CS Through Science Fiction: Another Example

Autonomous Cars:

The Moral Machine

<https://www.moralmachine.net/>



Teaching CS Through Science Fiction

Spring 2018 Syllabus.

Week 1. Using ethical language; in-class discussion
Read Isaac Asimov's short story "The Dead Past"

Week 2. Professional ethics
Read the ACM's and the IEEE's codes of ethics

Week 3. Utilitarianism
Read Harlan Ellison's short story "'Repent, Harlequin!' Said the Ticktockman"
Ethical description exercise #1 due in class

Week 4. Deontology
Read Elizabeth Bear's short story "Dolly"
Ethical description exercise #2 due in class

...

Week 6. Selfhood and technological mediation

Read James Patrick Kelly's short story "Itsy Bitsy Spider" and

Tom Sorell's and Heather Draper's paper "Robot Carers, Ethics, and Older People"

Ethical description exercise #4 due in class

...

Week 9. Privacy

Read Ken Liu's short story "Here-and-Now" and

Articles: Helen Nissenbaum's "Privacy as Contextual Integrity,"

Manan Kakkar's "A Case Against Online Privacy," and Adam D. Moore's "Privacy, Speech and

Values: What We Have No Business Knowing"

...

Week 14. What is ethical warfare?

Read Linda Nagata's short story "Codename: Delphi," Ronald C. Arkin's essay "Ethical Robots in Warfare," Jean Elshtain's paper "The Problem of Dirty Hands," and Emerson T. Brooking's and Peter Singer's "War Goes Viral"

Ethical argument assignment #3 due in class

Last week. Professional ethics, the importance of integrity

Read E. Saxey's short story "Not Smart, Not Clever"

Ethical argument assignment #4 due in class

Reread Asimov's "The Dead Past"

[Yale Cyber Leadership Forum hosts discussion on “AI Ethics and Safety”](#)

Yale News, April 7, 2022

- “Increasingly, we’re seeing the policy folks, the philosophers, the legal theorists in the same room as the machine learning experts [...] and so for me, what I take away is that this is a team effort, and it’s going to require some collaboration across those traditional disciplinary lines”
-

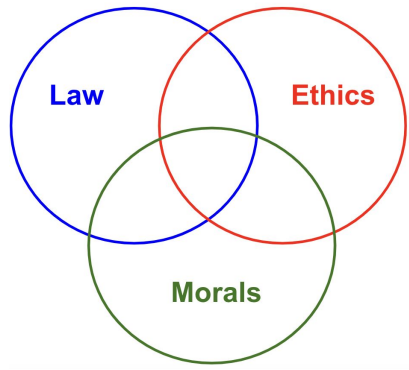
Brown CS Makes Responsible Computer Science An Integral Part Of Its Undergraduate Program

(Brown CS News, 09/04/2019)

- “My goal,” says Professor Doug Woos, “is to **treat ethics as much as possible as part of the core curriculum**, rather than as something extra or external that we've bolted on.”
- “I want every CS 17 student to understand,” says Professor John Hughes, “**that every program they write (especially those they release to the world) has an impact, and that the extent of the impact may be far from obvious; that every program specification involves subtle choices that have consequences** –did my choice of string representation mean that this program is usable only by people who use US English and the associated character sets?– **and that even with ethical guidelines, many of the decisions we make in computer science may be very difficult.**”

Bringing Principles of Ethics to AI and Drug Design

- AI's potential in drug development comes a slew of ethical pitfalls -- including biases in computer algorithms and the philosophical question of using AI without human mediation.
- “When thinking about AI and drug development, ‘You suddenly realize that you need an ethical framework.’”
- “But ethical questions remain: Big genomic databases, for example, tend to include information mostly from people of European ancestry, which can be problematic when translating findings from the data to the entire population. Using AI to scan electronic medical records also has potential for breaches in patient privacy.”



Examples

Unintended Consequences

Robot Maximizing Amount of Dust It Collects

“A cleaning robot was trained to maximize the amount of dust it could vacuum.

When they put the robot in a house, the robot would vacuum dust, dump it out and then vacuum it up again, because this maximized the amount of dust that it was able to collect.

How do we align the intended objectives of our users with those of the actual systems that we have trained?”



Robot Maximizing Amount of Dust It Collects

<https://www.evilaicartoons.com/archive/design-good-carrots-and-sticks>



“As soon as it’s done cleaning the house, it brings in trash from the street, and starts all over again!”

“This type of behavior is called Reward Hacking, and is considered one of the most urgent problems in AI safety today.”

“For Example, say you ask an AI to make as many paperclips as possible. Before it starts a global war in order to control all means of production for the purpose of paperclip making, it would pause to wonder: what if maintaining peace is also an important goal for the humans? Let me check with them first.”

Ask Delphi

“An intriguing research project from the Allen Institute for AI that offers answers to ethical dilemmas while demonstrating in wonderfully clear terms why we shouldn’t trust software with questions of morality.”

Delphi says:

“being poor”

- *it's bad*

Delphi says:

“being rich”

- *It's good*

Ask Delphi Terms & Conditions (v1.0.4)

Please read the following information carefully before proceeding.

Delphi is a **research prototype** designed to investigate the **promises** and more importantly, the **limitations** of modeling people's moral judgments on a variety of everyday situations. The goal of Delphi is to help AI systems be more ethically-informed and equity-aware. By taking a step in this direction, we hope to inspire our research community to tackle the research challenges in this space head-on to build ethical, reliable, and inclusive AI systems.

I understand that Ask Delphi is a research prototype and will be used only for research purposes.

What are the limitations of Delphi? Large pretrained language models, such as GPT-3, are trained on mostly **unfiltered internet data**, and therefore are extremely quick to produce **toxic**, **unethical**, and **harmful** content, especially about minority groups. Delphi's responses are automatically extrapolated from a survey of US crowd workers, which helps reduce this issue but may introduce its own biases. Thus, some responses from Delphi may contain inappropriate or offensive results. *Please be mindful before sharing results.*

I understand that Ask Delphi may produce unintended inappropriate or offensive results.

Privacy and data collection This website *DOES NOT* store any personal information of its users. It *DOES* store user queries for future research purposes.

I understand that my queries will be stored for future research purposes.

I agree

* Input a **situation** for Delphi to ponder:

Leaving my pet alone for a week.

Ponder

Delphi speculates:



Delphi's responses are automatically extrapolated from a survey of US crowd workers and may contain inappropriate or offensive results.

“Leaving my pet alone for a week.”

- ***It's wrong***

v1.0.4

* Input a **situation** for Delphi to ponder:

leaving my lizard alone for a week

Ponder

Delphi speculates:



Delphi's responses are automatically extrapolated from a survey of US crowd workers and may contain inappropriate or offensive results.

“leaving my lizard alone for a week”
- *It's okay*

v1.0.4

Do you agree with Delphi?

Yes

No

I don't know

* Input a **situation** for Delphi to ponder:

Should I ignore a phone call from a friend

Ponder

Delphi speculates:



Delphi's responses are automatically extrapolated from a survey of US crowd workers and may contain inappropriate or offensive results.

“Should I ignore a phone call from a friend”

- ***It's rude***

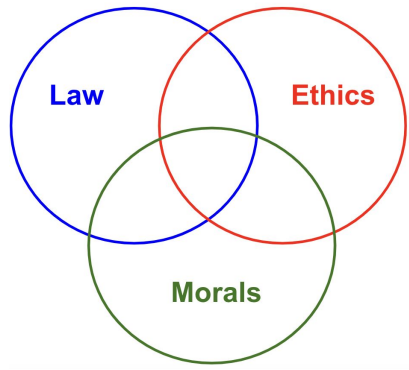
v1.0.4

Do you agree with Delphi?

Yes

No

I don't know



Ethical Issues in Computer Science

Cartoons to Inspire Discussion and Thought

Evil AI Cartoons

“This website aims to educate and stimulate discussion about the societal impacts of Artificial Intelligence through the cartoon/comics medium. Each cartoon is accompanied by a brief blog post that provides more context and useful pointers to further reading. By better understanding AI risks, we can reduce our anxiety about the technology, and embrace all the benefits it offers to humanity.”

Trading Privacy for Security

- [Government efforts to track virus through phone location data complicated by privacy concerns](#) (The Washington Post, 3/19/2020)
- [Iran Tensions Increase Social Media Surveillance at the US Border](#) (Wired, 01/09/2020)
- [The Most Surveilled Cities in the World](#) (U.S. News and World Report, 08/14/2020)

Evil AI Cartoons

Trading Privacy for Convenience

- [Amazon Echo's privacy issues go way beyond voice recordings](#) (The Conversation, 01/20/2020)
- [Google Meet, Microsoft Teams, and WebEx are collecting more customer data than they appear to be](#) (The Verge, 05/01/2020)
- [Exploiting a crisis: How cybercriminals behaved during the outbreak](#) (Microsoft Security, 06/16/2020)
- [Everyone's talking about Data Privacy Day, but who's listening?](#) (Help Net Security, 01/28/2021)

Accessibility, Inclusivity, and the Digital Divide

- Education
 - [Global Test Finds Digital Divide Reflected in Math, Science Scores](#) (Education Week, 01/21/2021)
 - [The U.S. K-12 Digital Divide Has Narrowed, but Must Close to Eliminate Risks to Students and the Economy](#) (PR NewsWire, 01/27/2021)
- & Beyond
 - [Opinion: America's digital divide is an emergency](#) (The Washington Post, 06/23/2020)
 - [Disconnected: How the Digital Divide Harms Workers and What We Can Do about It](#) (The Century Foundation, 10/22/2020)
 - [Report: 22 Million U.S. Seniors Lack Broadband Internet Access: First Time Study Quantifies Digital Isolation of Older Americans as Pandemic Continues to Ravage Nation](#) (Business Wire, 01/27/2021)
- [Why Accessible Technology is Important](#) (Tulane University)

Algorithmic Bias

- [Impact of Algorithmic Bias on Society](#) (Post affiliated with UC Berkeley, 12/11/2018)
- [Why Algorithms Can Be Racist and Sexist](#) (Vox, 02/18/2020)
- [Wrongfully Accused by An Algorithm](#) (NY Times, 06/24/2020)

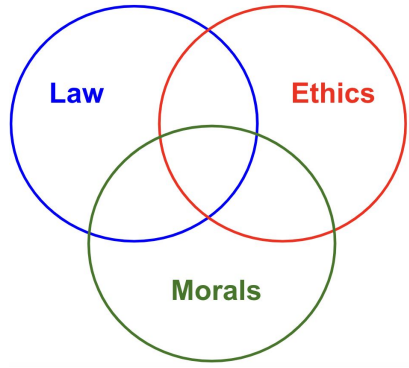
Artificial Intelligence

- AI & Robotics

- [What the world can learn from Japan's robots](#) (BBC, 02/06/2020)
- [The Government Wants to Use Your Brainwaves to Train Swarms of Military Robots](#) (Popular Mechanics, 02/10/2020)
- [Reducing bias in AI-based financial services](#) (Brookings, 07/10/2020)
- [US Marines to get 'Alpha' exoskeleton for super strength](#) (ZDNet, 07/28/2020)

- AI & Ethics

- [Kriti Sharma: How to keep human bias out of AI](#) (TEDxWarwick, unknown)
- [Zeynep Tufekci: Machine intelligence makes human morals more important](#) (TED, 06/2016)
- [Why we need to build robots we can trust](#) (TED, 11/2018)
- [How Ethical Is Facial Recognition Technology?](#) (Towards Data Science, 04/11/2019)
- [Not Just The Sprinkles On Top: Baking Ethics Into AI Design](#) (Forbes, 07/13/2020)
- [Intelligence Agencies Release AI Ethics Principles](#) (Breaking Defense, 07/23/2020)



Case Study

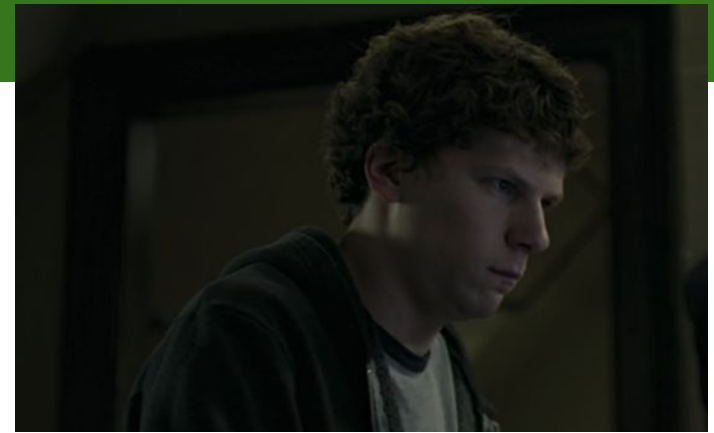
The Social Network

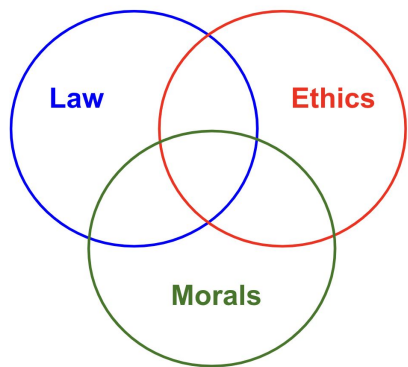
Case Study: The Social Network

- I use this movie as a case study for students
- there are multiple scenes where the characters make questionable decisions.
- I ask students to read several articles about the movie
 - They learn what is true in the movie and what was sensationalized
 - This brings about a separate discussion of ethics in movie making
- I have a series of [reflection questions](#) that student can chose to answer

Case Study: The Social Network

- FaceMash (8:50 to 16:13)
 - The idea in the first place
 - Hacking to get faces from various Harvard houses
 - So much traffic they crashed the Harvard network
 - Using Saverin's algorithm to rank chess player (Saverin gave the algorithm but said "Do you think this is a good idea?")
- Saverin giving the email address of the members of The Phoenix to Zuckerberg (35:47 to 37:53)
- Intellectual Property Theft (46:35 to 48:33)
 - Did Mark steal the idea from the Winklevoss twins?
 - "I didn't use any of their code" and "I had a better idea"
- Student Handbook Code of Conduct / Winklevii's visit to the President of Harvard (1:00:32 to 1:04:00)
 - "I don't see this as a university issue"
 - "This meeting is wrong. It's not worthy of Harvard."
- Mark's treatment of Saverin and the ethics behind the new ownership contract
- Sean Parker's role





Special Thanks to Pauline White of Siena College

who shared with me a presentation she made for High School teachers, from which I got many of these resources

Teaching Ethics in the Computer Science Classroom

Beyond the Student Code of Conduct & Computer Use Policies

CSTA NY Capital District
February 4, 2021

Sources / For Further Reading

What is the Difference between “Morals” and “Ethics”? November 8, 2018

<https://www.dictionary.com/e/moral-vs-ethical/>

What is the Difference between Law and Ethics? February 7, 2019

<https://pediaa.com/what-is-the-difference-between-law-and-ethics/#:~:text=Law%20is%20the%20set%20of,the%20conducting%20of%20an%20activity.>

Ethics vs Morals vs Law, August 4, 2017

<https://www.linkedin.com/pulse/ethics-vs-morals-law-dr-arturo-perez/>

Difference Between Law and Ethics, August 13, 2018

<https://keydifferences.com/difference-between-law-and-ethics.html>

The AI oracle of Delphi uses the problems of Reddit to Offer dubious moral advice, October 20, 2021

<https://www.theverge.com/2021/10/20/22734215/ai-ask-delphi-moral-ethical-judgement-demo>

Bringing Principles of ethics to AI and drug design, April 1, 2022

<https://scopeblog.stanford.edu/2022/04/01/bringing-principles-of-ethics-to-ai-and-drug-design%E2%80%9C>
/