Tension between innovation and regulation...

Innovation and safety
Self-determination and social wellbeing
Acceleration and precaution
...

"Stability AI's Diffusion model is available for anyone to use without hard restrictions. This means Diffusion can create images that OpenAI's DALL-E currently blocks, including propaganda, violent imagery, pornography, copyright violations, and disinformation and misinformation.... I urge using established powers to control the release of unsafe AI models."

Anna Eshoo, U.S. House of Representatives, California's 16th Congressional District (San Francisco/Silicon Valley)

Experiment: Dall-E and Stability

(Text *about* safety and innovation dilemma subjected *to* safety and innovation)

What kind of ethics supports Stability AI?
  *As opposed to: *Should OpenAI or Stability be supported?*

1. Uncertainty is encouraging

It looks like this request may not follow our
content policy.

Inputted text rejected, not because of a
demonstrable harm done, instead because
of what graphic version *might* do.

.

Instead of unknown as daunting because not
enough certainty, is provocative because so
much uncertainty: Uncertainty itself as
attractive.
  - Intrinsic, not uncertainty as promising
    unknown practical benefits…

Not knowing what graphics will be produced
is ethical justification for bypassing
restrictions and converting into pixels.

(Travel and (digital) nomadism.)

2. Innovation is intrinsically valuable

Value of any particular innovation = subsequent benefit or harm in society
- Innovation valuable only *after* social evaluation.

Innovation intrinsically valuable = *before* considering downstream effects in society.

Conventionally, AI designers burdened to show why projects should not be restricted. Now, burden shifts to restrictors.
- Not make AI unstoppable, only requires ethical reasons for halting to overcome initial creative value of AI existing.

(AI engineering innovation resembles artistic creativity as Kant's sublime.)

It looks like this request may not follow our content policy.

Innovation problems solved by limiting innovation.

"Innovation" and "Solutions to innovation problems" are opposed.

Response to deepfake pornography, racist text = *Pause, Stop AI*.
- OpenAI initially limited users to 400, allowing human review of pictures.

Innovation solves innovation problems.
- Diffusion $200k prize for open-source deepfake detector. (Wildcat coders…)
- Mathematical coating against deepfakery.
- Machine Unlearning

The only way out is through: straight on until emerging on the other side.

(Tesla / Bigwave)

"Innovation" and "Solution to innovation problems" not opposed

The only way out is through: more, faster...

4. Decentralized regulation

It looks like this request may not follow our content policy.

Regulations come before AI applications released, and narrow group develops guidelines for wider community.
- High-Level Expert Group produces *"Ethics Guidelines for Trustworthy AI."*

Instead of regulations conditioning AI use, AI use produces regulations. And, instead of few deciding for many, the many users decide for each other.

Example
- Female sumo wrestlers and representative fairness versus reinforcement learning with human feedback.

(Swarm intelligence/ethics + Augmented democracy.)

5. Embedded ethics

It looks like this request may not follow our content policy.

Centralized ethicists issue restrictions from above and outside the technology: more like policing than participation.

Philosophers/ethicists mix with information engineers and domain experts from start to spin language of human values alongside AI modeling.

==Ethical skill closer to translation than restriction== because primary aim is participants with diverse technical languages collaborating.
  • Formulating AI dilemmas more imperative than resolving.

(Working in 2 directions, German Big Data Lab.)

Elements of acceleration AI ethics

Uncertainty is encouraging. The unknown displays potential more than risk, and therefore AI models are developed for the same reason that (digital?) nomads travel: because we do not know how we will change and be changed.
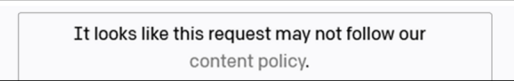
Innovation is intrinsically valuable. Because it resembles artistic creativity in being worthwhile even before considering social implications, an ethical burden tips: engineers no longer need to justify starting their models, instead, others need to demonstrate reasons for halting.

The only way out is through: more, faster. When ethical problems arise, the response is not to slow AI and veer away, it is straight ahead until emerging on the other side. Accelerating AI solves AI problems.

Decentralization. The overarching ethics of permissions and restrictions governing AI evolve with the broad community of users and their uses, instead of representing a select group's peremptory judgments.

Embedding. Ethics works inside AI and with engineers to pose questions about human values, instead of remaining outside and emitting restrictions. Formulating AI dilemmas is more imperative than resolving them.

Conclusion

It looks like this request may not follow our content policy.

Precaution

Humanist innovation problems solved by slowing innovation: fewer problems

Written into foundational European Union law, the precaution principle states:

If it is possible that an action might cause harm to the public and if there is still no scientific agreement on the issue, the action in question should not be carried out (Eur-lex 2023).
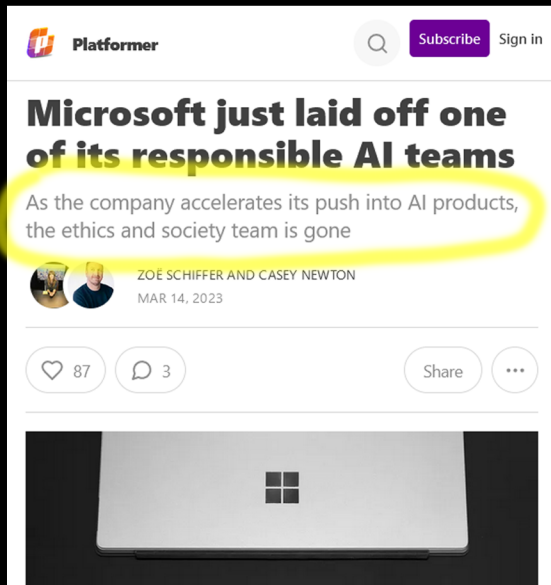
Acceleration

Humanist innovation problems solved by accelerating innovation: problems resolved

Practical argument

Practically, the choice is not between
precaution/conventional ethics and
acceleration. It is between acceleration
and resignation...

Platformer

Microsoft just laid off one of its responsible AI teams

As the company accelerates its push into AI products, the ethics and society team is gone

ZOË SCHIFFER AND CASEY NEWTON
MAR 14, 2023

♡ 87   💬 3   Share   ⋯

How to Accelerate Ethics for Innovation and Against Precaution in Generative AI

James Brusseau
Philosophy, Pace University, New York City
& Information Engineering and Computer Science, University of Trento, Italy

Abstract

The contrast between Dall-E and Stable Diffusion asks whether AI will slow for humanist ethics, or human ethics will accelerate for AI. If the answer is acceleration, five elements increase ethical velocity.

*Appendix: Range of responses to generative AI

Debate in this essay

"DESTROY DATACENTERS BY AIRSTRIKE."	SLOW AI (UNITED STATES) PRECAUTION (EUROPE)	ACCELERATION AI ETHICS	"MOVE FAST & BREAK THINGS" POSTHUMANISM
Stop AI	**Slow AI innovation to human speeds**	**Accelerate humanism to AI innovation speeds**	**Surrender humanism and humans to AI innovations**
Humanism privileged over machines	**Humanism privileged over machines**	**Humanism privileged over machines**	**Machines privileged over humanism**
Safety as default setting against value of technical innovation	**Safety as default setting against value of technical innovation**	**Innovation as intrinsically valuable independently of safety**	**Innovation as intrinsically valuable independently of safety**