

Externalities & Dual-Use Technologies

CS 221 Artificial Intelligence: Principles & Techniques

Learning Objectives

- Identify externalities & dual-use technologies
- Understand the distinction between externalities and dual-use technologies
- Apply theoretical background on externalities and dual-use technologies to identify these in the real world

Externalities

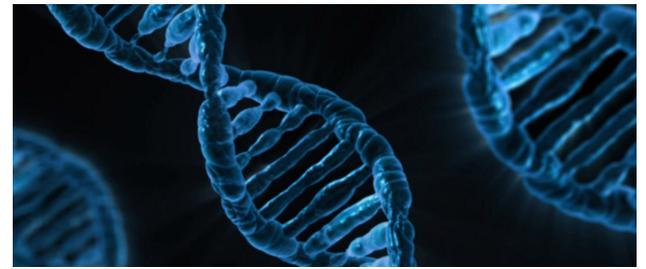
Externalities

An externality is a consequence (cost) or a benefit that arises from one party's action and impacts another party.

Key points about externalities:

- Externalities can be a result of either the production or the consumption of a good or service
- Externalities can be positive or negative
- The cost or benefit can be private – impacting an individual or organization – or social – impacting society as a whole

Example: Ancestry Testing



Positive Externalities

- Connect individuals with their biological family members
- Inform individuals on their genetic background
- Provide genetic screening for health risks

Negative Externalities

- Genetic information sold to third parties
- Mishandling of genetic information and data (FTC, 2023)

Dual-Use Dilemma in Technology

Dual Use Dilemma in Technology - Definition

A conundrum describing the reality that any piece of technology has the potential to be used for harm as well as for good.

Adapted from Seumas Miller & Michael J. Selgelid: Ethical and Philosophical Consideration of the Dual-use Dilemma in the Biological Sciences

Example: The Manhattan Project



Oppenheimer's academic research did not start with the intention of creating an atomic bomb.

Arms race against the Nazis under immense pressure by the American government.

Academic Freedom? Should we have the right to uninhibitedly participate in intellectual inquiry?

But, what about the effect? Nearly 230,000 people died in Hiroshima and Nagasaki. Who's to take responsibility for this tragedy?

Sometimes the thing you intend your technology to do is not the only thing it can do or will be used for.

Four scenarios to consider in dual-use

1. Intended outcomes

“follow an instruction in a prompt and provide a detailed response” - from OpenAI

2. Unintended but foreseen outcomes

Delivers false information with an authoritative tone.

3. Unintended but foreseeable outcomes

Replaces human workers (e.g., ChatGPT is getting more and more competent at writing code.

4. Unforeseen and possibly impossible to have foreseen outcomes

Professing love to a user - from an NYT report

The perils of machine learning in designing new chemicals and materials

[Sadasivan Shankar](#) ✉ & [Richard N. Zare](#) ✉

[Nature Machine Intelligence](#) **4**, 314–315 (2022) | [Cite this article](#)

4601 Accesses | 4 Citations | 20 Altmetric | [Metrics](#)

Machine Learning and Chemicals

Research on machine learning model that can identify toxicity in chemicals

Positive: “less than 1% of chemicals under commercial use in the U.S. has undergone toxicity characterization.”

Negative: bad actors can use the model to develop targeted toxins.

Individual vs Institutional Responsibility

Dual use technologies are created within a collective (e.g., academia, company, military)

Pressure to publish, perform, profit.

Individuals can't see the big picture.