

Topics:

- Artificial Intelligence is Disruptive
- Social Impacts
- The Digital Economy
- Values and Bias

The Social Impact of Artificial Intelligence

“Once the rockets are up, who cares where they come down? That’s not my department,” says Wernher von Braun.

[Lehrer, 1965]

Never in the history of humanity have we allowed a machine to autonomously decide who should live and who should die Before we allow our cars to make ethical decisions, we need to have a global conversation to express our preferences to the companies that will design moral algorithms, and to the policymakers that will regulate them.

[Awad, 2018]

Our vision is that computer scientists can and should play a key role in helping address societal and environmental challenges in pursuit of a sustainable future, while also advancing computer science as a discipline.

[Gomes, 2019]

Artificial Intelligence is Disruptive

- Artificial intelligence is a transformational set of ideas, algorithms, and tools.
- AI is both a science and a technology.
- AI systems are now increasingly deployed at scale in the real world.
- Significant impact across almost all forms of human activity
- Including the economic, social, psychological, healthcare, legal, political, government, scientific, technological, manufacturing, military, media, educational, artistic, transportation, agricultural, environmental, and philosophical spheres

Social Impacts

- Impacts both beneficial and harmful
- Must consider AI risks and harms.
- Ethical, governance and regulatory issues abound.
- Consider use of AI in social media: benefits and harms.
- Platforms allow people to connect and form social groups in healthy and unhealthy ways.
- Optimize users' feeds to maximize engagement to increase advertising revenue
- Maximizing engagement leads to adversarial behavior and polarization.
- Can rights to privacy, fair treatment, accountability and trustworthy systems be guaranteed?

The Digital Economy - 1

- The uses of AI have developed rapidly in the **digital economy**, with the rise of digital communication and the Internet.
- The largest global corporations rely heavily upon AI applications.
- Those companies are centered more on the use of information than on the production of material goods.
- The shift from matter to information is characterized as **dematerialization**.
- This **atoms-to-bits** transformation allows transactions with less friction and more speed.
- Streaming music easier, quicker, cheaper, and more material and energy efficient than what used to happen: going to a store to buy a record or CD.

- Digitalization leads to a general temporal speedup of society and the economy.
- It shrinks distances through telecommunication.
- We all live now in McLuhan's **global village** [1962].
- Digital revolution eliminates many intermediaries between the producers and consumers of goods and services.
- People (and jobs) such as retail clerks, bank tellers, and travel agents become redundant - the process of **disintermediation**.

The Digital Economy - 3

- Digital revolution and AI are transforming the global economy.
- This transformation has both benefits and harms for people.
- The benefits, and the harms, are very unevenly distributed in the new economy.
- A **winner-take-all** dynamic: the most powerful corporations (and societies) use their power to increase their dominance, often establishing a monopolistic, or oligopolistic, market position.
- AI and machine learning algorithms, relying on tracking and modeling users, are central to this dynamic .
- Zuboff [2019] characterized the new economy as **surveillance capitalism**, epitomized by the large-scale harvesting of personal data online to facilitate targeted monitoring and advertising for commercial and political purposes.

- **Human attention**, selective concentration on available information, is a critical and limited resource.
- Attention is a psychological **and** an economic issue.
- Herb Simon created the key concept of the **attention economy**.

In an information-rich world, the wealth of information means a dearth of something else: a scarcity of whatever it is that information consumes. What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it.

[Simon, 1971]

- Turning human attention into a commodity requires monitoring users, triggering privacy concerns.
- Corporations, and other actors, know a lot about us, using that knowledge to manipulate our attention, our thoughts, and our actions.

- Learning systems, trained on large datasets, produce outputs that reflect any bias present in the training sets.
- Datasets are acquired in the past, so using them to predict future outcomes propagates any bias from the past to the future.
- What if the future will not, or should not, resemble the past?
- Consider policing, recommending healthcare treatment,

- Notice two different meanings of **bias**.
- In machine learning, **bias** has a neutral technical meaning, “the tendency to prefer one hypothesis over another”.
- The **no-free-lunch theorem** implies that any effective learning algorithm *must* have a bias in that sense.
- In ordinary language use, human **bias** has a negative connotation, meaning “prejudice in favor of or against one thing, person, or group compared with another”.

- Training sets for **facial recognition**, usually acquired without informed consent, typically do not represent people equitably, causing misclassification, often with harmful effects.
- **Large language models**, pre-trained on vast text corpora, when prompted often produce new text that is racist, sexist, or otherwise demeaning of human dignity.
- An AI-based decision system, or a generative AI system, inherently reflects certain implicit values, or preferences.

Facial Recognition

Selinger and Leong [2021] define four forms of facial recognition:

- **facial detection** finds the location of faces in images
- **facial characterization** finds features of individual faces, such as approximate age, emotions, what the person is looking at
- **facial verification** determines whether the person matches a single template
- **facial identification** is used to identify each person in an image from a database of faces.

Facial identification, usually considered the most problematic

- perfect and pervasive facial identification → people know they are constantly under surveillance → self-censorship.
- Given a database + facial detection + facial verification → facial identification. A false-positive rate of 1 in 10 million results in 800 people on earth who match a particular face.
- When facial identification makes mistakes, they usually do not affect all groups equally.

- Whose values are incorporated into AI systems?
- Typically, the values of the designer or owner of the system, or the values implicit in a deep learning training set.
- Can those values be made explicit?
- Is it possible to ensure those are democratic values, avoiding discrimination and prejudice?
- Can systems be designed that respect privacy, dignity, equity, diversity, and inclusion?
- What values (whose values) should be incorporated into AI systems?