



Digital
Futures

No. 23

Teaching Machines Right from Wrong

Applying artificial intelligence (AI) to complex decisions has clear benefits. But it also increasingly means automating ethical choices that can alter human lives.

The better machines get at finding and learning from patterns in Big Data, the more important it is to think critically about how and whether to use those patterns.



By 2018, smart machines will supervise **over 3 million workers worldwide**.



21% of consumers in an FTC study had **confirmed errors** on their credit reports.



2014: **the first annual** Fairness, Accountability, and Transparency in Machine Learning **conference**.



A private university **encouraged 20-25 students to drop out** based on AI predictions of poor grades.

[Link to Sources](#)

Real-world examples of **misused AI algorithms** abound. These are just a few:

- Women who weren't pregnant — or weren't ready to reveal it — received special offers of baby products and "congratulatory" messages.
- People with minority ethnic names received a disproportionate number of ads implying they had criminal records.
- Guests at a party learned a ride-hailing company kept track of customers who stayed out all night and went home in the wee hours.



Ethical-Edge Cases

Credit scoring algorithms designed to evaluate lending risk are now commonly used to gauge reliability and trustworthiness, determining whether someone should get a job or apartment.

Insurance underwriting algorithms determine the extent, price, and type of coverage someone can get, with little room for disagreement.

Healthcare algorithms could be used to penalize the currently healthy for their probability of future illness.

Algorithms often use zip codes as proxy for (illegal) racial profiling in major decisions, such as employment and law enforcement.

Self-driving cars will have to learn how to react in an accident situation when every possible outcome is bad.

What Should We Do About It?

All machine learning contains assumptions and biases of the humans who create it — unconscious or otherwise. To ensure fairness, business leaders must insist that AI be built on a strong ethical foundation.

We can:

- Monitor algorithms for neutrality and positive outcomes.
- Support academic research into making AI-driven decisions more fair, accountable, and transparent.
- Create human-driven overrides, grievance procedures, and anti-bias laws.
- Include ethics education in all employee training and development.

Above all, we must consider this a human issue, not a technological one. AI is only as unbiased a tool as we make it. It's our responsibility to keep it on the ethical straight and narrow.