23

DATA ETHICS

Tony Fish

Data ethics centres on how to make better decisions and ensure ethical outcomes. Data ethics is a broad topic encapsulating data and ethical decisions, data-based judgement, bias, the ethics of collecting and analysing data, and the ethics of automation created by data. The data we collected in 2010 could paint a low-quality, black and white, abstract picture of our world, and the data could be used to inform, frame, or guide. In 2020, the data collected can describe our world in high-definition colour, and is being used to mould, form, and shape actions in complicated environments. We are accelerating our collection and use of data. Boards need to get deep and dirty with data and the ethics of judgement based on data. It is complex and often requires new skills.

Professional bias

Boards are biased towards professionals who represent experience in finance, legal, operations, marketing, and sales. These “core functions” are supported by professionals with information, innovation, technology, and human resources expertise. Boards will typically have a deep appreciation

of the expertise in each of the core functions and will have their own experience in each core area to draw on as a reference point. Members feel comfortable understanding and questioning strategy, ratios, metrics, margins, and returns. Often the supporting functions are represented by a single professional and there is little broader expertise amongst the other board members. When data and ethics enter the room, most members feel uncomfortable. Still, there is a chorus of opinions backed from well-known academic institutes, leading consultancies, articles, and journals, but little practical grounding. This paragraph is an observation on the current status and not a criticism. Data, ethics, and data ethics are new, and this chapter details questions we should be reflecting on and asking, as data is now fundamentally a core function.

Data and statistics

“Lies, damned lies and statistics”¹ is an unattributed quote but helpful in the context of discussing data and ethics. The use of statistics to bolster weak arguments is now followed by the use of data to create the outcome you want. Like statistics, which can back any opinion, data can be found to support any recommendation. The most fundamental question a board needs to wrestle with: “How do we know if the decision we are being asked to approve is only based on data that creates that recommendation?” Data is no different to any question a board has been asked to approve previously; however, according to McKinsey, 95% of decisions made in business are based on financial data, which represents, according to Google, 3% of the data available. We accept that there is bias in the data that gets presented in such a board paper derived from this 3% of data, but as the remaining 97% of data becomes available to the board for decision-making, it is not going to get any easier to unpick (to carefully analyse the different elements) if we are being asked to make the right “ethical” decision. It is worth noting that data is not easy to understand, and in many situations, is not understood even by experts and professionals. Data is not oil, sunshine, time, labour, or any other of the analogies, which all fail. Data is data.² Data is closer to the discovery of a new element or quantum particle,

1 “Lies, damned lies and statistics” https://en.wikipedia.org/wiki/Lies,_damned_lies,_and_statistics.

2 “Data is Data” <https://medium.com/@tonyfish/data-is-data-90ba0b803178>.

which has its own unique characteristics. Data itself creates data, and everything is data. Insight, knowledge, and wisdom built on data just create more data. Decisions are just more data. Data means that bias is amplified, which can hide ethical problems. Our systems are built on discrimination.

Ethics

Like data, ethics is equally not easy to understand, and in many situations is not understood. However, as we study and gain a better understanding of human behaviour from more data, it becomes increasingly evident that determining an ethical outcome for two people is complicated, for society it is complex, and for everyone it is impossible. We must be able to explain and justify the rationale for any decisions which affects others. However, linking our judgement and decision-making to a framework that our ethics or culture can be applied to at scale, is a delusional claim. We have to get comfortable that our judgement and decisions, based on any ethical code, hold us to a standard by which we must be prepared to be judged. If we are deemed to be lacking by professional standard, law, or society, directors must be willing to accept the consequences.

Governing data

Data in this context is all data. All data embraces data that comes from the market, customers, partners, supply chain, and internal data from all systems, as shown in the Figure 23.1 below. Data is collected and stored, analysed, and this creates value, insights, and margin, which we observe as an outcome. More data comes from understanding how our actions, which are measured outcomes, impact each of the data collection points. Governance and oversight encompass how we manage the entire system and process from the collection of data to understanding how our decisions impact outcomes. Data enables a leadership team to qualify and quantify new risks beyond the obvious of data security, cyberattacks, and poor processes. Data enables leadership to enquire about any and every aspect of the business that can be measured, which means that whilst once we were unable to determine the impact of a decision, a board can now receive direct feedback from the data that supported a decision, to the outcome, and determine if the delta needs to be addressed.

Figure 23.1 Interface to ecosystem

Schools of thought

Essentially all ethical and moral thinking has the same goal, and there are several diverse ways to help reach a positive impact by considering what is best for others and seeking mutual benefit. Selecting a model for a specific context increases the chances of a successful outcome, which is a positive impact. There are numerous books in which moral philosophy and philosophers are categorised; however, there are three core schools: virtue ethics, consequentialist ethics, and deontological or duty-based ethics (Table 23.1). Each approach provides a different way to understand ethics and has implications for the decision-making process. If I asked the question “What is the best way to achieve a healthy life?” you could respond with one of three approaches: good nutrition, through exercise, or through spiritual discipline. Each is vital but inadequate by itself. It is in bringing these and other strategies together that we can create better outcomes from a complex decision. In ethics, no one school answers all the issues, concerns, and problems raised by a diverse society; therefore, many schools need to be considered to reach an ethical decision.

Data ethics

- Do we have a data ontology and/or data dictionary? What standard does it meet?
- What confidence do we have in the lineage and provenance of data?
- What bias could be in the data, what bias is in the data set, what are the implications?
- What data do you get from third parties, what is their model, and have you asked q2 and q3 of the sources of external data?
- Do you need the data you have, is the data valid, how have you determined that the data is useful?
- What data is being used to support human and automated decision-making?

Ethics and data

- Which framework (from the Table 23.1) are individual leaders using, and are we applying all frameworks equally?
- What framework does our company and our values align to and why?
- Whose ethics are we guided by, and how do we know?
- Who is accountable for the outcome of our ethical choice?
- What level of transparency are we using, and is it working for us?

Algorithms and automation based on data ethics

- Do we know the true origin of the algorithms used in our company?
- Do we know what data was used to create/determine the algorithms?
- How are we sure that the data set used would apply today?
- How are we made aware of the unintended/unimagined consequences?
- What process and procedures do we have in place to test and qualify algorithms and automated decision-making, and are these biased towards certain re-enforcing conformational outcomes?

Table 23.1 Ethics and philosophy

	Virtue Ethics	Consequentialist Ethics	Deontological Ethics
Overall-philosophy	How to Live Your Life?	Is It Good?	Is It Right?
Summary	What kind of person/ company do I/we want to be in the current situation/ context?	What impact is my/ our behaviour having on the world?	What are my ethical principles telling me, we should do?
Key Questions	What virtues bring me/ us closer to my/our goal; which vices prevent me/us from achieving it?	Am I/we doing more good or harm by my/our behaviour?	What duties do I/we owe and how do I/ we decide between conflicting duties?
Informing Ethical Decisions:	Is my/our behaviour consistent with being a moral person?	Is my/our behaviour making the world a better place?	What does reason require of me/us regarding my/our treatment of others? Are my/our reasons consistent and coherent?
Key guiding principles	Aspiring to a set of virtues and avoiding a set of vices. Integrity is a primary value and finding the right balance within and between values.	Actions aim at bringing about the greatest good for the most significant number of people. Benevolence is a primary value.	Having a duty to others based on ethical principles. Respecting the autonomy of others is a primary value.
Original Philosophers	Aristotle (384–322 BCE) Alasdair MacIntyre (1929)	David Hume (1711–1776) Jeremy Bentham (1748–1832) John Stuart Mill (1806–1873)	John Locke (1632–1704) Immanuel Kant (1724–1804)

Analysis and insights based on data ethics

- What tools are we using to create analysis and insights based on data?
- Are our tools part of the system or independent of the system?
- Who is checking the analysis and insights, and how are they being checked?

- What happens to analysis and insight that does not align to decisions, strategy, outcomes, or rewards?
- Who controls and determines the process by which the analysis and insight is conducted and presented?

Governance and oversight for data ethics

- Do we have diverse data ethics skills and experience on this board?
- Is data ethics working for us?
- Who is “us” in this context?
- Whom are we asking about data ethics, bias, and outcomes?
- How do we know if the decision we are being asked to approve is based on data that could only create that recommendation?
- What checks and balances do we have in place to understand and determine the filters and direction we are being guided towards?
- What delegated authority is there that enables ethical questions to be asked and decided on outside of our visibility, especially regarding marketing, sales, privacy, terms, model, and the use of data?
- How are we sure, and reassured, our decisions based on data have a positive impact on our customers, partners, and society?
- Are we willing to be judged by a professional standard, law, or society, and do we accept the consequences?

It is critical that we can explain and justify the rationale for decisions which affect others made with data. We must be comfortable that our judgement and decisions, based on our chosen ethical code, hold us to a standard by which we must be prepared to be judged. If we are thought to be lacking by professional standard, law, or society, we must be willing to accept the consequences.