

QuantumBlack, AI by McKinsey

Data ethics: What it means and what it takes

Every company must establish its own best practices for managing its data. Here are five pitfalls to avoid based on our conversations with experts and early adopters.

by Alex Edquist, Liz Grennan, Sian Griffiths, and Kayvaun Rowshankish



Now more than ever, every company is a data company. By 2025, individuals and companies around the world will produce an estimated 463 exabytes of data each day,¹ compared with less than three exabytes a decade ago.²

With that in mind, most businesses have begun to address the *operational* aspects of data management—for instance, determining how to build and maintain a data lake or how to integrate data scientists and other technology experts into existing teams. Fewer companies have systematically considered and started to address the *ethical* aspects of data management, which could have broad ramifications and responsibilities. If algorithms are trained with biased data sets or data sets are breached, sold without consent, or otherwise mishandled, for instance, companies can incur significant reputational and financial costs. Board members could even be held personally liable.³

So how should companies begin to think about ethical data management? What measures can they put in place to ensure that they are using consumer, patient, HR, facilities, and other forms of data appropriately across the value chain—from collection to analytics to insights?

We began to explore these questions by speaking with about a dozen global business leaders and data ethics experts. Through these conversations, we learned about some common data management traps that leaders and organizations can fall into, despite their best intentions. These traps include thinking that data ethics does not apply to your organization, that legal and compliance have data ethics covered, and that data scientists have all the answers—to say nothing of chasing short-term ROI at all costs and looking only at the data rather than their sources.

In this article, we explore these traps and suggest some potential ways to avoid them, such as adopting new standards for data management, rethinking

governance models, and collaborating across disciplines and organizations. This list of potential challenges and remedies is not exhaustive; our research base was relatively small, and leaders could face many other obstacles, beyond our discussion here, to the ethical use of data. But what's clear from our research is that data ethics needs both more and sustained attention from all members of the C-suite, including the CEO.

Potential challenges for business leaders

There is a dynamic body of literature on data ethics. Just as the methods companies use to collect, analyze, and access data are evolving, so will definitions of the term itself. In this article, we define data ethics as data-related practices that seek to preserve the trust of users, patients, consumers, clients, employees, and partners. Most of the business leaders we spoke to agreed broadly with that definition, but some have tailored it to the needs of their own sectors or organizations (see sidebar, “What is data ethics?”). Our conversations with these business leaders also revealed the unintended lapses in data ethics that can happen in organizations. These include the following:

Thinking that data ethics doesn't apply to your organization

While privacy and ethical considerations are essential whenever companies use data (including artificial-intelligence and machine-learning applications), they often aren't top of mind for some executives. In our experience, business leaders are not intentionally pushing these thoughts away; it's often just easier for them to focus on things they can “see”—the tools, technologies, and strategic objectives associated with data management—than on the seemingly invisible ways data management can go wrong.

In a 2021 McKinsey Global Survey on the state of AI, for instance, only 27 percent of some 1,000 respondents said that their data professionals

¹ Jeff Desjardins, “How much data is generated each day?” World Economic Forum, April 17, 2019.

² IBM Research Blog, “Dimitri Kanevsky translating big data,” blog entry by IBM Research Editorial Staff, March 5, 2013.

³ Leah Rizkallah, “Potential board liability for cybersecurity failures under Caremark law,” *CPO Magazine*, February 22, 2022.

What is data ethics?

We spoke with about a dozen business leaders and data ethics experts. In their eyes, these are some characteristics of ethical data use:

It preserves data security and protects customer information. The practitioners we spoke with tend to view cybersecurity and data privacy as part and parcel of data ethics. They believe companies have an ethical responsibility (as well as legal obligations) to protect customers' data, defend against breaches, and ensure that personal data are not compromised.

It offers a clear benefit to both consumers and companies. "The consumer's got to

be getting something" from a data-based transaction, explained an executive at a large financial-services company. "If you're not solving a problem for a consumer, you've got to ask yourself why you're doing what you're doing." The benefit to customers should be straightforward and easy to summarize in a single sentence: customers might, for instance, get greater speed, convenience, value, or savings.

It offers customers some measure of agency. "We don't want consumers to be surprised," one executive told us. "If a customer receives an offer and says, 'I think I got this because of how you're using my data, and that makes me uncomfortable. I

don't think I ever agreed to this,' another company might say, 'On page 41, down in the footnote in the four-point font, you did actually agree to this.' We never want to be that company."

It is in line with your company's promises. In data management, organizations must do what they say they will do—or risk losing the trust of customers and other key stakeholders. As one senior executive pointed out, keeping faith with stakeholders may mean turning down certain contracts if they contradict the organization's stated data values and commitments.

actively check for skewed or biased data during data ingestion. Only 17 percent said that their companies have a dedicated data governance committee that includes risk and legal professionals. In that same survey, only 30 percent of respondents said their companies recognized equity and fairness as relevant AI risks. AI-related data risks are only a subset of broader data ethics concerns, of course, but these numbers are striking.

Thinking in silos: Legal, compliance, or data scientists have data ethics covered

Companies may believe that just by hiring a few data scientists, they've fulfilled their data management obligations. The truth is data ethics is everyone's domain, not just the province of data scientists or of legal and compliance teams. At different times, employees across the organization—from the front

line to the C-suite—will need to raise, respond to, and think through various ethical issues surrounding data. Business unit leaders will need to vet their data strategies with legal and marketing teams, for example, to ensure that their strategic and commercial objectives are in line with customers' expectations and with regulatory and legal requirements for data usage.

As executives navigate usage questions, they must acknowledge that although regulatory requirements and ethical obligations are related, adherence to data ethics goes far beyond the question of what's legal. Indeed, companies must often make decisions before the passage of relevant laws. The European Union's General Data Protection Regulation (GDPR) went into effect only in May 2018, the California Consumer Privacy Act has been in effect only since

January 2020, and federal privacy law is only now pending in the US Congress. Years before these and other statutes and regulations were put in place, leaders had to set the terms for their organizations' use of data—just as they currently make decisions about matters that will be regulated in years to come.

Laws can show executives what they *can* do. But a comprehensive data ethics framework can guide executives on whether they *should*, say, pursue a certain commercial strategy and, if so, how they should go about it. One senior executive we spoke with put the data management task for executives plainly: “The bar here is not regulation. The bar here is setting an expectation with consumers and then meeting that expectation—and doing it in a way that’s additive to your brand.”

Chasing short-term ROI

Prompted by economic volatility, aggressive innovation in some industries, and other disruptive business trends, executives and other employees may be tempted to make unethical data choices—for instance, inappropriately sharing confidential information because it is useful—to chase short-term profits. Boards increasingly want more standards for the use of consumer and business data, but the short-term financial pressures remain. As one tech company president explained: “It’s tempting to collect as much data as possible and to use as much data as possible. Because at the end of the day, my board cares about whether I deliver growth and EBITDA.... If my chief marketing officer can’t target users to create an efficient customer acquisition channel, he will likely get fired at some point—or at least he won’t make his bonus.”

Looking only at the data, not at the sources

Ethical lapses can occur when executives look only at the fidelity and utility of discrete data sets and don’t consider the entire data pipeline. Where did the data come from? Can this vendor ensure that the subjects of the data gave their informed consent for use by third parties? Do any of the market data contain material nonpublic

information? Such due diligence is key: one alternative data provider was charged with securities fraud for misrepresenting to trading firms how its data were derived. In that case, companies had provided confidential information about the performance of their apps to the data vendor, which did not aggregate and anonymize the data as promised. Ultimately, the vendor had to settle with the US Securities and Exchange Commission.⁴

A few important building blocks

These data management challenges are common—and they are by no means the only ones. As organizations generate more data, adopt new tools and technologies to collect and analyze data, and find new ways to apply insights from data, new privacy and ethical challenges and complications will inevitably emerge. Organizations must experiment with ways to build fault-tolerant data management programs. These seven data-related principles, drawn from our research, may provide a helpful starting point.

Set company-specific rules for data usage

Leaders in the business units, functional areas, and legal and compliance teams must come together to create a data usage framework for employees—a framework that reflects a shared vision and mission for the company’s use of data. As a start, the CEO and other C-suite leaders must also be involved in defining data rules that give employees a clear sense of the company’s threshold for risk and which data-related ventures are OK to pursue and which are not.

Such rules can improve and potentially speed up individual and organizational decision making. They should be tailored to your specific industry, even to the products and services your company offers. They should be accessible to all employees, partners, and other critical stakeholders. And they should be grounded in a core principle—for example, “We do not use data in any way that we cannot link to a better outcome for our customers.” Business

⁴ “SEC charges App Annie and its founder with securities fraud,” US Securities and Exchange Commission, September 14, 2021.

Leaders must come together to create a data usage framework that reflects a shared vision and mission for the company's use of data.

leaders should plan to revisit and revise the rules periodically to account for shifts in the business and technology landscape.

Communicate your data values, both inside and outside your organization

Once you've established common data usage rules, it's important to communicate them effectively inside and outside the organization. That might mean featuring the company's data values on employees' screen savers, as the company of one of our interview subjects has done. Or it may be as simple as tailoring discussions about data ethics to various business units and functions and speaking to their employees in language they understand. The messaging to the IT group and data scientists, for instance, may be about creating ethical data algorithms or safe and robust data storage protocols. The messaging to marketing and sales teams may focus on transparency and opt-in/opt-out protocols.

Organizations also need to earn the public's trust. Posting a statement about data ethics on the corporate website worked for one financial-services organization. As an executive explained: "When you're having a conversation with a government entity, it's really helpful to be able to say, 'Go to our website and click on Responsible Data Use, and you'll see what we think.' We're on record in a way that you can't really walk back." Indeed, publicizing your company's data ethics framework may help increase the momentum

for powerful joint action, such as the creation of industry-wide data ethics standards.

Build a diverse data-focused team

A strong data ethics program won't materialize out of the blue. Organizations large and small need people who focus on ethics issues; it cannot be a side activity. The work should be assigned to a specific team or attached to a particular role. Some larger technology and pharmaceutical companies have appointed chief ethics or chief trust officers in recent years. Others have set up interdisciplinary teams, sometimes referred to as data ethics boards, to define and uphold data ethics. Ideally, such boards would include representatives from, for example, the business units, marketing and sales, compliance and legal, audit, IT, and the C-suite. These boards should also have a range of genders, races, ethnicities, classes, and so on: an organization will be more likely to identify issues early on (in algorithm-training data, for example) when people with a range of different backgrounds and experiences sit around the table.

One multinational financial-services corporation has developed an effective structure for its data ethics deliberations and decision making. It has two main data ethics groups. The major decisions are made by a group of senior stakeholders, including the head of security and other senior technology executives, the chief privacy officer, the head of the consulting arm, the head of strategy, and the heads of brand, communications,

and digital advertising. These are the people most likely to use the data.

Governance is the province of another group, which is chaired by the chief privacy officer and includes the global head of data, a senior risk executive, and the executive responsible for the company's brand. Anything new concerning data use gets referred to this council, and teams must explain how proposed products comply with the company's data use principles. As one senior company executive explains, "It's important that both of these bodies be cross-functional because in both cases you're trying to make sure that you have a fairly holistic perspective."

As we've noted, compliance teams and legal counsel should not be the only people thinking about a company's data ethics, but they do have an important role to play in ensuring that data ethics programs succeed. Legal experts are best positioned to advise on how your company should apply existing and emerging regulations. But teams may also want to bring in outside experts to navigate particularly difficult ethical challenges. For example, a large tech company brought in an academic expert on AI ethics to help it figure out how to navigate gray areas, such as the environmental impact of certain kinds of data use. That expert was a sitting but not voting member of the group because the team "did not want to outsource the decision making." But the expert participated in every meeting and led the team in the work that preceded the meetings.

Engage champions in the C-suite

Some practitioners and experts we spoke with who had convened data ethics boards pointed to the importance of keeping the CEO and the corporate board apprised of decisions and activities. A senior executive who chaired his organization's data ethics group explained that while it did not involve the CEO directly in the decision-making process, it brought all data ethics conclusions to him "and made sure he agreed with the stance that we were taking." All these practitioners and experts agreed that having a champion or two in the C-suite can signal the importance of data ethics to the rest of the

organization, put teeth into data rules, and support the case for investment in data-related initiatives.

Indeed, corporate boards and audit committees can provide the checks needed to ensure that data ethics are being upheld, regardless of conflicting incentives. The president of one tech company told us that its board had recently begun asking for a data ethics report as part of the audit committee's agenda, which had previously focused more narrowly on privacy and security. "You have to provide enough of an incentive—a carrot or a stick to make sure people take this seriously," the president said.

Consider the impact of your algorithms and overall data use

Organizations should continually assess the effects of the algorithms and data they use—and test for bias throughout the value chain. That means thinking about the problems organizations might create, even unwittingly, in building AI products. For instance, who might be disadvantaged by an algorithm or a particular use of data? One technologist we spoke with advises asking the hard questions: "Start your meetings about AI by asking, 'Are the algorithms we are building sexist or racist?'"

Certain data applications require far greater scrutiny and consideration. Security is one such area. A tech company executive recalled the extra measures his organization took to prevent its image and video recognition products and services from being misused: "We would insist that if you were going to use our technology for security purposes, we had to get very involved in ensuring that you debiased the data set as much as possible so that particular groups would not be unfairly singled out." It's important to consider not only what types of data are being used but also what they are being used for—and what they could potentially be used for down the line.

Think globally

The ethical use of data requires organizations to consider the interests of people who are not in the room. Anthropologist Mary Gray, the senior principal

researcher at Microsoft Research, raises questions about global reach in her 2019 book, *Ghost Work*. Among them: Who labeled the data? Who tagged these images? Who kept violent videos off this website? Who weighed in when the algorithm needed a steer?

Today's leaders need to ask these sorts of questions, along with others about how such tech work happens. Broadly, leaders must take a 10,000-foot view of their companies as players in the digital economy, the data ecosystem, and societies everywhere. There may be ways they can support policy initiatives or otherwise help to bridge the digital divide, support the expansion of broadband infrastructure, and create pathways for diversity in the tech industry. Ultimately, data ethics requires leaders to reckon with the ongoing rise in global inequality—and the increasing concentration of wealth and value both in geographical tech hubs and among AI-enabled organizations.⁵

Embed your data principles in your operations

It's one thing to define what constitutes the ethical use of data and to set data usage rules; it's another to integrate those rules into operations across the organization. Data ethics boards, business unit leaders, and C-suite champions should build a common view (and a common language) about how data usage rules should link up to both the company's data and corporate strategies and

to real-world use cases for data ethics, such as decisions on design processes or M&A. In some cases, there will be obvious places to operationalize data ethics—for instance, data operations teams, secure-development operations teams, and machine-learning operations teams. Trust-building frameworks for machine-learning operations can ensure that data ethics will be considered at every step in the development of AI applications.

Regardless of which part of the organization the leaders target first, they should identify KPIs that can be used to monitor and measure its performance in realizing their data ethics objectives. To ensure that the ethical use of data becomes part of everyone's daily work, the leadership team also should advocate, help to build, and facilitate formal training programs on data ethics.

Data ethics can't be put into practice overnight. As many business leaders know firsthand, building teams, establishing practices, and changing organizational culture are all easier said than done. What's more, upholding your organization's data ethics principles may mean walking away from potential partnerships and other opportunities to generate short-term revenues. But the stakes for companies could not be higher. Organizations that fail to walk the walk on data ethics risk losing their customers' trust and destroying value.

Alex Edquist is an alumna of McKinsey's Atlanta office; **Liz Grennan** is an associate partner in the Stamford, Connecticut, office; **Sian Griffiths** is a partner in the Washington, DC, office; and **Kayvaun Rowshankish** is a senior partner in the New York office.

The authors wish to thank Alyssa Bryan, Kasia Chmielinski, Ilona Logvinova, Keith Otis, Marc Singer, Naomi Sosner, and Eckart Windhagen for their contributions to this article.

Designed by McKinsey Global Publishing
Copyright © 2022 McKinsey & Company. All rights reserved.

Find more content like this on the
McKinsey Insights App



Scan • Download • Personalize



⁵ For more on the concentration of value among AI-enabled firms, see Marco Iansiti and Karim R. Lakhani, *Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World*, Boston: Harvard Business Review Press, 2020.