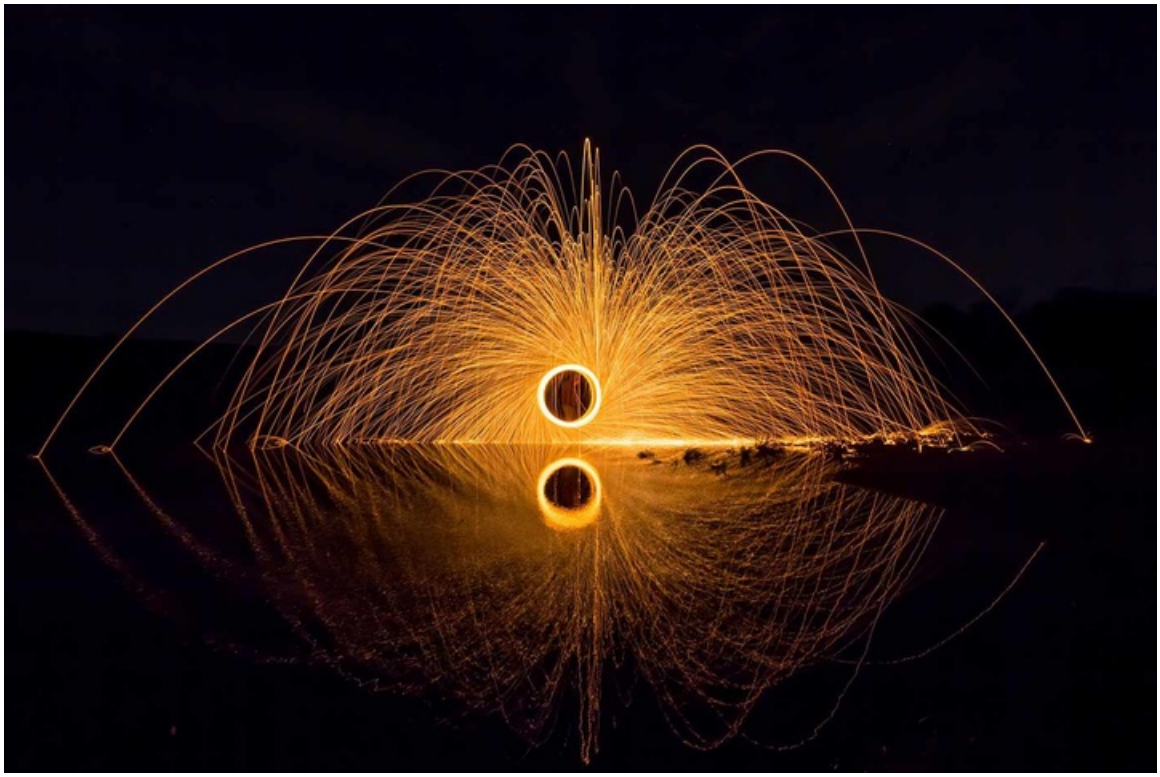
DATA SCIENCE

Data's day of reckoning

We can build a future we want to live in, or we can build a nightmare. The choice is up to us.

By Mike Loukides, Hilary Mason, and DJ Patil. July 31, 2018



Shower of sparks (source: Pixabay)

Check out the session "An ethical foundation for the AI-driven future" at the Strata Data Conference in New York, September 11-13, 2018.

This post is part of a series on data ethics.

Our lives are bathed in data: from recommendations about whom to "follow" or "friend" to data-driven autonomous vehicles. But in the past few years, it has become clear that the products and technologies we have created have been weaponized and used against us. Although we've benefited from the use of data in countless ways, it has also created a tension between individual privacy, public good, and corporate profits. Cathy O'Neil's Weapons of Math Destruction and Virginia Eubanks' Automating Inequality document the many ways that data has been used to harm the broader population.

Data science, machine learning, artificial intelligence, and related technologies are now facing a day of reckoning. It is time for us to take responsibility for our creations. What does it mean to take responsibility for

building, maintaining, and managing data, technologies, and services? Responsibility is inevitably tangled with the complex incentives that surround the creation of any product. These incentives have been front and center in the conversations around the roles that social networks have played in the 2016 U.S. elections, recruitment of terrorists, and online harassment. It has become very clear that the incentives of the organizations that build and own data products haven't aligned with the good of the people using those products.

These issues aren't new to the consumer internet. Other fields have had their days of reckoning. In medicine, widely publicized abuses include the Tuskegee syphilis experiment, the case of Henrietta Lacks (whose cells were used for cancer research without her permission and without compensation), and human experiments performed during World War II by Nazis. The physics community had to grapple with the implications of the atomic bomb. Chemists and biologists have had to address the use of their research for chemical and biological weapons. Other engineering disciplines have realized that shoddy work has an impact on people's lives; it's hard to ignore bridge collapses. As a result, professional societies were formed to maintain and enforce codes of conduct; government regulatory processes have established standards and penalties for work that is detrimental to society.

Ethics and security training

In many fields, ethics is an essential part of professional education. This isn't true in computer science, data science, artificial intelligence, or any related field. While courses on ethics exist at many schools, the ideas taught in ethics classes often aren't connected to existing projects or course work. Students may study ethical principles, but they don't learn how to implement those principles in their projects. As a result, they are ill-prepared for the challenges of the real world. They're not trained to think about ethical issues and how they affect design choices. They don't know how to have discussions about projects or technologies that may cause real-world harm.

Software security and ethics frequently go hand in hand, and our current practices for teaching security provide an example of what not to do. Security is usually taught as an elective, isolated from other classes about software development. For example, a class on databases may never discuss SQL injection attacks. SQL injection would be addressed in classes on security, but not in the required database course. When a student submits a project in a database course, its vulnerability to hostile attack doesn't affect the grade; an automated grading system won't even test it for vulnerabilities. Furthermore, a database course might not discuss architectural decisions that limit damage if an attacker gains access—for example, storing data elements such as names and social security numbers in different databases.

Teaching security in an elective is better than not teaching it at all; but the the best way to produce programmers who really understand security is to incorporate it into assignments and grading within the core curriculum, in addition to teaching it in electives. The core curriculum ensures that everyone can recognize and deal with basic security problems; the electives can go into greater depth, and tie together issues from different disciplines ranging from physical security to cryptography. Security as an afterthought doesn't work in product development; why do we expect it to work in education? There is no industry in which security lapses haven't led to stolen data, affecting millions of individuals. Poor security practices have led to serious vulnerabilities in many consumer devices, from smart locks to smart light bulbs.

Ethics faces the same problem. Data ethics is taught at many colleges and universities, but it's isolated from the rest of the curriculum. Courses in ethics help students think seriously about issues, but can't address questions like getting informed consent in the context of a real-world application. The White House report "Preparing for the Future of Artificial Intelligence" highlights the need for training in both ethics and security:

> Ethical training for AI practitioners and students is a necessary part of the solution. Ideally, every student learning AI, computer science, or data science would be exposed to curriculum and discussion on related ethics and security topics. However, ethics alone is not sufficient. Ethics can help practitioners understand their responsibilities to all stakeholders, but ethical training should be augmented with technical tools and methods for putting good intentions into practice by doing the technical work needed to prevent unacceptable outcomes.

Ethics and security must be at the heart of the curriculum, not only as electives, or even isolated requirements. They must be integrated into every course at colleges, universities, online courses, and programming boot camps. They can't remain abstract, but need to be coupled with "technical tools and methods for putting good intentions into practice." And training can't stop upon graduation. Employers need to host regular forums and offer refresher courses to keep people up-to-date on the latest challenges and perspectives.

Developing guiding principles

The problem with ethical principles is that it's easy to forget about them when you're rushing: when you're trying to get a project finished on a tight, perhaps unrealistic, schedule. When the clock is ticking away toward a deadline, it's all too easy to forget everything you learned in class—even if that class connected ethics with
solutions to real-world problems.

Checklists are a proven way to solve this problem. A checklist, as described by Atul Gawande inThe Checklist Manifesto, becomes part of the ritual. It's a short set of questions that you ask at the start of the project, and at every stage as you move toward release. You don't go to the next stage until you've answered all the questions affirmatively. Checklists have been shown to reduce mistakes in surgery; they're used very heavily by airline pilots, especially in emergencies; and they can help data professionals to not forget ethical issues, even when they are under pressure to deliver.

Here's a checklist we've developed for developers working on data-driven applications:

❑ Have we listed how this technology can be attacked or abused?

❑ Have we tested our training data to ensure it is fair and representative?

❑ Have we studied and understood possible sources of bias in our data?

❑ Does our team reflect diversity of opinions, backgrounds, and kinds of thought?

❑ What kind of user consent do we need to collect and use the data?

❑ Do we have a mechanism for gathering consent from users?

❑ Have we explained clearly what users are consenting to?

❑ Do we have a mechanism for redress if people are harmed by the results?

❑ Can we shut down this software in production if it is behaving badly?

❑ Have we tested for fairness with respect to different user groups?

❑ Have we tested for disparate error rates among different user groups?

❑ Do we test and monitor for model drift to ensure our software remains fair over time?

❑ Do we have a plan to protect and secure user data?

Feel free to modify this checklist to fit your situation and use it in your projects.

The Fairness, Accountability, and Transparency in Machine Learning group (FAT/ML) advocates a similar approach. Their Principles for Accountable Algorithms and a Social Impact Statement for Algorithms suggests assessing the social impact statement of a project at least three times during its life: during design, pre-launch, and post-launch. Working through a social impact statement requires developers to think about the ethical consequences of their projects and address any problems that turn up. In a similar vein, the Community Principles on Ethical Data Practices, which arose out of the Data For Good Exchange (D4GX), provides a set of values and principles that have been gathered through community discussion. They're a great start for any group that wants to create its own checklist. And Cathy O'Neil has proposed auditing machine learning algorithms for fairness.

Building ethics into a data-driven culture

Individual responsibility isn't sufficient. Ethics needs to be part of an organization's culture. We've seen many organizations recognize the value of developing a data-driven culture; we need to ensure ethics and security become part of that culture, too.

Security is gradually becoming a part of corporate culture: the professional, financial, legal, and reputational consequences of being a victim are too large to ignore. Organizations are experimenting with bug-bounty programs, sharing threats with each other, and collaborating with government agencies. Security teams are no longer simply corporate naysayers; they're charged with preventing serious damage to an organization's reputation and to finances.

Integrating ethics into corporate culture has been more challenging. A single team member may object to an approach, but it's easy for an individual to be overruled, and if there's no support for ethical thinking within the organization, that's likely to be where it ends. Ethical thinking is important with or without corporate support, but it's more likely to make a difference when ethical action is a corporate value. Here are some ideas for building ethics into culture:

- An individual needs to be empowered to stop the process before damage is done. Toyota and W. Edwards Deming pioneered the use of the andon cord to improve quality and efficiency. Anyone who saw a problem could pull the cord, which would halt the production line. Senior managers as well as production line operators would then discuss the issue, make improvements, and restart the process.

  Any member of a data team should be able to pull a virtual "andon cord," stopping production, whenever they see an issue. The product or feature stays offline until the team has a resolution. This way, an iterative process can be developed that avoids glossing over issues.
- Anyone should be able to escalate issues for remediation without fear of retaliation. There needs to be an escalation process for team members who don't feel their voice has been heard. The U.S. Department of State has a dissent channel where any diplomat can make sure the Secretary of State hears their concerns. In health care, a path to escalate legal and ethical issues is required by law. For health care plans in the U.S., there is a compliance officer who reports directly to the board of directors.

Data-driven organizations need a similar model that allows people to escalate issues without the fear of reprisal. An escalation process could be implemented in several forms. For example, companies could work with an organization such as the Electronic Frontier Foundation (EFF) to develop a program that accepts and investigates whistleblower reports. The problem would be kept from public scrutiny unless specific criteria are violated. A similar approach could be implemented under an existing or new agency (e.g., a Consumer Data Protection Agency).

- An ethical challenge should be part of the hiring process. When hiring, companies frequently assess whether a candidate will be a "cultural fit." Interviewers ask questions that help them understand whether a candidate will work well with other team members. However, interviewers rarely ask questions about the candidate's ethical values.

  Rather than asking a question with a right/wrong answer, we've found that it's best to pose a problem that lets us see how the candidate thinks about ethical and security choices. Here's a question we have used:

  > *Assume we have a large set of demographic data. We're trying to evaluate individuals and we're not supposed to use race as an input. However, you discover a proxy for race with the other variables. What would you do?*

  This kind of question can start a dialogue about how to use the proxy variable. What effects does it have on people using the product? Are we making recommendations, or deciding whether to provide services? Are we implementing a legal requirement, or providing guidance about compliance? Discussing the question and possible answers will reveal the candidate's values.

- Product reviews must ask questions about the product's impact. Environmental impact statements predict the impact of construction projects on the public. We've already mentioned FAT/ML's proposed Social Impact Statements as an example of what might be done for data. In the social sciences and the biomedical industry, Institutional Review Boards (IRBs) assess the possible consequences of experiments before they're performed.

  While both environmental impact statements and IRBs present problems for data products, data teams need to evaluate the impact of choices they make. Teams need to think about the consequences of their actions before releasing products. We believe that using a checklist is the best approach for ensuring good outcomes.

- Teams must reflect diversity of thought, experiences, race, and background. All too often, we hear about products that are culturally insensitive or overtly racist. One notorious example is an automated passport control system that doesn't let an individual proceed until a good digital image is captured. People of Asian ancestry reported that the system kept asking them to open their eyes, even though their eyes were open. Many cringe-worthy examples are well documented; they can often be traced to a lack of data or a lack of insight into the diversity of the population that will be impacted.

  While there's no general solution to these problems of cultural sensitivity, diversity and inclusion are a tremendous help. Team members should be from the populations that will be impacted. They'll see issues well before anyone else. External peer reviews can help to reveal ethical issues that your team can't see. When you're deeply involved with a project, it can be hard to recognize problems that are obvious to outsiders.

- Corporations must make their own principles clear. Google's "Don't be evil" has always been a cute, but vague, maxim. Their recent statement, Artificial Intelligence at Google: Our Principles, is much more

specific. In a similar vein, the face recognition startup Kairos has said that they won't do business with law enforcement companies. Kairos' CEO writes that "the use of commercial face recognition in law enforcement or government surveillance of any kind is wrong."

However, it's important to realize that advocating for corporate ethical principles has consequences. Significant internal protest, and the resignation of several developers in protest over Google's defense contracts, were needed to get their AI principles in place. Kairos is probably leaving a lot of money on the table. It's also important to realize that organizations frequently point to their ethical principles to divert attention from unethical projects.

Over the past few years, we've heard a lot about software startups that begin with a "minimal viable product," and adhere to Facebook's slogan, "move fast and break things." Is that incompatible with the approach we've just described? This is a false choice. Going fast doesn't mean breaking things. It is possible to build quickly and responsibly.

The lean/agile methodology used in many startups is a good way to expose ethical issues before they become problems. Developers start with a very simple product ("the simplest thing that could possibly work," according to Ward Cunningham's seminal phrase), demo it to users, get feedback, develop the next version, and repeat. The process continues for as many iterations as needed to get a product that's satisfactory. If a diverse group of users tests the product, the product development loop is likely to flush out systematic problems with bias and cultural insensitivity. The key is testing the product on a truly diverse group of users, not just a group that mirrors the expected customer base or the developers' backgrounds.

Regulation

In some industries, ethical standards have been imposed by law and regulation. The Nuremberg Code was developed in response to Nazi atrocities. It focuses on individual consent to participation in an experiment. After the Tuskegee syphilis experiments became public knowledge, the code was put into law in the 1974 National Research Act and the 1975 Declaration of Helsinki. This push to codify ethical guidelines established the role of the institutional review board (IRB), and was adopted widely in the U.S via the Common Rule.

In other industries, other regulatory bodies enforce ethical standards These include the U.S. Federal Trade Commission (FTC), which oversees commerce; the Nuclear Regulatory Commission (NRC), which oversees nuclear power plants; the Federal Food and Drug Administration (FDA), which oversees the safety of pharmaceuticals; and, most recently, the Consumer Finance Protection Bureau (CFPB), which oversees bankers and lenders on behalf of consumers.

The European Union's General Data Protection Regulation (GDPR) takes an aggressive approach to regulating data use and establishing a uniform data policy. In June 2018, California passed a digital privacy law similar to GDPR, despite the reservations of many online companies. One challenge of developing a policy framework is that the policy development process nearly always lags the pace of innovation, and isn't agile enough to keep policy iterative. By the time a policy has been formulated and approved, it almost always lags behind technology; but it's impossible for policy makers to iterate quickly enough to catch up with the newest technology. Another problem is that the committees that make policy often lack experts with the necessary technical background. That can be good; technologists are too easily influenced by "what technology wants." But policies created by people who are technologically uninformed are frequently out of touch with reality: look at the debate over back doors to encryption protocols.

Some have argued that organizations using data should adopt the Institutional Review Board (IRB) model

from the biomedical industry. Unfortunately, while there are many positive aspects of the IRB, this isn't a viable approach. IRBs are complex, and they can't be agile; it's very difficult for IRBs to adapt to new ideas and technologies. It's why the Obama Administration pushed for nearly eight years to update the Common Rule's models for consent to be consistent with digital technologies and to enable data mining.

Building our future

For some time, we've been aware of the ethical problems that arise from the use and abuse of data. Public outcry over Facebook will die down eventually, but the problems won't. We're looking at a future in which most vehicles are autonomous; we will be talking to robots with voices and speech patterns that are indistinguishable from humans; and where devices are listening to all our conversations, ready to make helpful suggestions about everything from restaurants and recipes to medical procedures. The results could be wonderful—or they could be a nightmarish dystopia.

It's data's day of reckoning. The shape of the future will depend a lot on what we do in the next few years. We need to incorporate ethics into all aspects of technical education and corporate culture; we need to give people the freedom to stop production if necessary, and to escalate concerns if they're not addressed; we need to incorporate diversity and ethics into hiring decisions; and we may need to consider regulation to protect the interests of individual users, and society as a whole.

Above all, talk about ethics! In "It's time for data ethics conversations at the dinner table," Natalie Evans Harris and others write that "we need to be having difficult conversations about our individual and collective responsibility to handle data ethically." This is the best single thing you can do to further data ethics: talk about it in meetings, at lunch, and even at dinner. Signing a data oath, or agreeing to a code of conduct, does little if you don't live and breathe ethics. Once you are living and breathing it, you will start to think differently about the code you write, the models you build, and the applications you create. The only way to create an ethical culture is to live it. The change won't take place magically, nor will it be easy—but it's necessary.

We can build a future we want to live in, or we can build a nightmare. The choice is up to us.

Article image: Shower of sparks (source: Pixabay).

Tags: Data Ethics.