# 5 Principles of Data Ethics for Business Professionals

## 1. Ownership

The first principle of data ethics is that an individual has ownership over their personal information. Just as it's considered stealing to take an item that doesn't belong to you, it's unlawful and unethical to collect someone's personal data without their consent.

Some common ways you can obtain consent are through signed written agreements, digital privacy policies that ask users to agree to a company's terms and conditions, and pop-ups with checkboxes that permit websites to track users' online behavior with cookies. Never assume a customer is OK with you collecting their data; always ask for permission to avoid ethical and legal dilemmas.

## 2. Transparency

In addition to owning their personal information, data subjects have a right to know how you plan to collect, store, and use it. When gathering data, exercise transparency.

For instance, imagine your company has decided to implement an algorithm to personalize the website experience based on individuals' buying habits and site behavior. You should write a policy explaining that cookies are used to track users' behavior and that the data collected will be stored in a secure database and train an algorithm that provides a personalized website experience. It's a user's right to have access to this information so they can decide to accept your site's cookies or decline them.

Withholding or lying about your company's methods or intentions is deception and both unlawful and unfair to your data subjects.

## 3. Privacy

Another ethical responsibility that comes with handling data is ensuring data subjects' privacy. Even if a customer gives your company consent to collect, store, and analyze their personally identifiable information (PII), that doesn't mean they want it publicly available.

PII is any information linked to an individual's identity. Some examples of PII include:

- Full name
- Birthdate
- Street address
- Phone number
- Social Security card
- Credit card information
- Bank account number
- Passport number

To protect individuals' privacy, ensure you're storing data in a secure database so it doesn't end up in the wrong hands. Data security methods that help protect privacy include dual-authentication password protection and file encryption.

For professionals who regularly handle and analyze sensitive data, mistakes can still be made. One way to prevent slip-ups is by de-identifying a dataset. A dataset is de-identified when all pieces of PII are removed, leaving only anonymous data. This enables analysts to find relationships between variables of interest without attaching specific data points to individual identities.

Related: Data Privacy: 4 Things Every Business Professional Should Know

4. Intention

When discussing any branch of ethics, intentions matter. Before collecting data, ask yourself why you need it, what you'll gain from it, and what changes you'll be able to make after analysis. If your intention is to hurt others, profit from your subjects' weaknesses, or any other malicious goal, it's not ethical to collect their data.

When your intentions are good—for instance, collecting data to gain an understanding of women's healthcare experiences so you can create an app to address a pressing need—you should still assess your intention behind the collection of each piece of data.

Are there certain data points that don't apply to the problem at hand? For instance, is it necessary to ask if the participants struggle with their mental health? This data could be sensitive, so collecting it when it's unnecessary isn't ethical. Strive to collect the minimum viable amount of data, so you're taking as little as possible from your subjects while making a difference.

Related: 5 Applications of Data Analytics in Health Care

5. Outcomes

Even when intentions are good, the outcome of data analysis can cause inadvertent harm to individuals or groups of people. This is called a disparate impact, which is outlined in the Civil Rights Act as unlawful.

In Data Science Principles, Harvard Professor Latanya Sweeney provides an example of disparate impact. When Sweeney searched for her name online, an advertisement came up that read, "Latanya Sweeney, Arrested?" She had not been arrested, so this was strange.

"What names, if you search them, come up with arrest ads?" Sweeney asks in the course. "What I found was that if your name was given more often to a Black baby than to a white baby, your name was 80 percent more likely get an ad saying you had been arrested."

It's not clear from this example whether the disparate impact was intentional or a result of unintentional bias in an algorithm. Either way, it has the potential to do real damage that disproportionately impacts a specific group of people.

Unfortunately, you can't know for certain the impact your data analysis will have until it's complete. By considering this question beforehand, you can catch any potential occurrences of disparate impact.