

# Benefits of Ethical Data Science

## **What kind of benefits does ethical data science bring to society?**

Ethical data science gives us new ways to understand our society. We have opportunities to gain a better understanding of inequities, and the factors that contribute to those. We also may better predict the tradeoffs inherent in a proposed policy, the costs for a given benefit, and where those land.

## **What's an example of how you can use a cost benefit analysis for a proposed policy?**

For example, suppose one were to develop a prediction mechanism that was somewhat unfairly biased against certain individuals. That would seem to be problematic. However, such a system might still be a great deal more fair in its decision-making than a typical human being. Research tells us that Judges are more likely to give a harsh sentence before lunch (when their blood sugar is low) than afterwards.

There might therefore be a tradeoff between a modest and controllable machine-driven bias, and a worse human alternative. Factors must not be considered in terms of associated elements and the value of potential alternatives, and not simply in isolation.

## **What kind of benefits does it bring to businesses, governments, and other organizations?**

Ethical data science can empower effective leadership with more precise and sophisticated information, enabling better decisions. Depending on the focus of an analysis, a macro view can provide a very broad oversight of the whole picture of a situation, or alternatively a detailed drill down into specific elements on a small scale.

## **What's an example of how ethical data science has created benefits in society?**

The Environmental Defense Fund has sponsored data science projects that collect information on air pollution on moving platforms such as automobiles. This enabled them to triangulate where the worst pollution was coming from. In a very short space of time, the air quality in New York City has been transformed, and similar benefits are now helping to protect the environment worldwide.

## **What are the benefits of data sharing, data philanthropy and open data?**

Data is often compared to oil, as it is an essential 'fuel' for algorithmic technologies to learn about the world from. However, data can of course be copied an unlimited number of times, it doesn't get used up when it is applied. This means that we can potentially bring the benefits of effective and ethical data to empower many individuals and organizations around the world. By sharing good data mutually for no consideration, we have opportunities for all to become richer, assuming that the data is valid and unbiased, of course.

## **How do we go about collecting and sharing open data? Who is leading this movement?**

There are many web services that aggregate open data for public benefit. To follow up on the air quality example from earlier, researchers, governments, companies, and citizens can share their data on air pollution with entities that specialize in that kind of data, such as [OpenAQ.org](https://openaq.org) or [AQICN.org](https://aqicn.org).

This means that we can create a high resolution near real-time map of air pollution as it happens. Such large datasets can then be mined to generate statistics, predictions, and to pinpoint opportunity for intervention.

There are a great many open data sets, on everything from air, sea, and road traffic, to weather and climate models. Such data sources are a powerful asset in the data-driven ecosystem of ethical emerging technologies.