


# The Complete Beginner's Guide To Big Data Everyone Can Understand

Bernard Marr Contributor 

Follow

Mar 14, 2017, 03:34am EDT

 This article is more than 6 years old.

**Big Data** is a term that has been around for some time now but there is still confusion about what it actually is. The concept is continuing to evolve and to be reconsidered, as it remains the driving force behind many ongoing waves of digital transformation, including artificial intelligence, data science and the Internet of Things (IoT).

With that in mind I thought it was time to write a beginner's guide to what Big Data means in 2017. In a similar way to my beginner's guides to [Blockchain](#) and [FinTech](#), this will be jargon-free and aims to explain the core concepts and ideas to anyone regardless of background knowledge.

It all starts with the exponential explosion in the amount of data we have generated since the dawn of the digital age. This is largely due to the rise of computers, the Internet and technology capable of capturing information from the real, physical world we live in, and converting it to digital data.

In 2017, we generate data whenever we go online, when we carry our GPS-equipped smartphones, when we communicate with our friends through social media or chat applications, and when we shop. You could say we leave digital footprints with everything we do that involves a digital transaction, which is almost everything.

On top of this, the amount of machine-generated data is rapidly growing too. Data is generated and shared when our "smart" home devices communicate with each other or with their home servers. Industrial machinery in plants and factories around the world is increasingly equipped with sensors that gather and transmit data. Soon, self-driving cars will take to the streets, beaming a real-time, four-dimensional maps of their surroundings back home from wherever they go.

## What can Big Data do?

This ever-growing stream sensor information, photographs, text, voice and video data, is the foundation of Big Data which we can now use in ways that were not possible even a few years ago. Right now, Big Data projects are helping to:



Shutterstock

---

**Forbes Daily: Get our best stories, exclusive reporting and essential analysis of the day's news in your inbox every weekday.**

Email address

Sign Up

By signing up, you accept and agree to our [Terms of Service](#) (including the class action waiver and arbitration provisions), and [Privacy Statement](#).

- 
- **Cure disease and prevent cancer** – Data-driven medicine involves analyzing vast numbers of medical records and images for patterns which can help spot disease early and develop new medicines.
  - **Feed the hungry** – Agriculture is being revolutionized by data which can be used to maximize crop yields, minimize the amount of pollutants released into the ecosystem and optimize the use of machines and equipment
  - **Explore distant planets** – NASA analyzes millions of data points and uses them to model every eventuality to land its Rovers on the surface of Mars and plan future missions.
  - **Predict and respond to natural and man-made disasters** – Sensor data can be analyzed to predict where earthquakes are likely to strike next, and patterns of human behavior give clues which help aid organizations give relief to survivors. Big Data technology is also used to monitor and safeguard the flow of refugees away from war zones around the world.
  - **Prevent crime** – Police forces are increasingly adopting data-driven strategies based on their own intelligence and public data sets in order to deploy resources more efficiently and act as a deterrent where one is needed.

- Make our everyday lives easier and more convenient – Shopping online, crowdsourcing a ride or a place to stay on holiday, choosing the best time to book flights and deciding what movie to watch next are all easier thanks to Big Data.

## **How does Big Data work?**

Big Data works on the principle that the more you know about anything or any situation, the more reliably you can gain new insights and make predictions about what will happen in the future. By comparing more data points, relationships will begin to emerge that were previously hidden, and these relationships will enable us to learn and inform our decisions.

Most commonly this is done through a process which involves building models, based on the data we can collect, and then running simulations, tweaking the value of data points each time and monitoring how it impacts our results. This process is automated – today’s advanced analytics technology will run millions of these simulations, tweaking all the possible variables until it finds a pattern – or an insight – that helps solve the problem it is working on.

Increasingly, data is coming to us in an unstructured form, meaning it cannot be easily put into structured tables with rows and columns. Much of this data is in the form of pictures and videos – from satellite images to photographs uploaded to Facebook or Twitter – as well as email and instant messenger communications and recorded telephone calls. To make sense of all of this, Big Data projects often use cutting edge analytics involving [artificial intelligence](#) and [machine learning](#). By teaching computers to identify what this data represents– through image recognition or natural language processing, for example – they can learn to spot patterns much more quickly and reliably than humans.

A strong trend over the last few years has been a move towards the delivery of Big Data tools and technology through an “as-a-service” platform. Businesses and organizations rent server space, software systems and processing power from third-party cloud service providers. All of the work is carried out on the service provider’s systems, and the customer simply pays for whatever was used. This model is making Big Data-driven discovery and transformation accessible to any organization and cuts out the need to spend vast sums on hardware, software, premises and technical staff.

## **Big Data concerns**

Today, Big Data gives us unprecedented insights and opportunities, but it also raises concerns and questions that must be addressed:

- Data privacy – The Big Data we now generate contains a lot of information about our personal lives, much of which we have a right to keep private. Increasingly we are asked to strike a balance between the amount of personal data we divulge, and the convenience that Big Data powered apps and services offer. Who do we allow to have access to this data?

- Data security – Even if we decide we are happy for someone to have our data for a particular purpose, can we trust them to keep it safe? Is the existing legal framework up to the job of regulating data use at this scale?
- Data discrimination – When everything is known, will it become acceptable to discriminate against people based on data we have on their lives? We already use credit scoring to decide who can borrow money, and insurance is heavily data-driven. We can expect to be analyzed and assessed in greater detail, and care must be taken that this isn't done in a way which contributes to making life more difficult for those who already have fewer resources and access to information.

Facing up to these challenges is part of “Big Data,” too. They are certainly a major part of the debate around the use of Big Data in academic circles. However they must also be addressed by those who want to take advantage of Big Data in business. Failure to do so can leave businesses vulnerable and lead to financial disaster as well as huge fines.

When people first started talking about Big Data it was sometimes dismissed as a fad – the latest trendy technology term which would be talked about for a while then quietly forgotten about when the next big thing came along. This hasn't proven to be the case yet – in fact, while newer buzzwords have popped up, Big Data is still the driving force behind just about all of them. The amount of data available to us is only going to increase, and analytics technology will become more capable. So if Big Data is capable of all of this today – just imagine what it will be capable of tomorrow.

*Follow me on [Twitter](#) or [LinkedIn](#). Check out my [website](#) or some of my other work [here](#).*



**Bernard Marr**

Follow

Bernard Marr is an internationally best-selling author, popular keynote speaker, futurist, and a strategic business & technology advisor to governments... **Read More**

Editorial Standards

Reprints & Permissions

---

ADVERTISEMENT