

# Glossary

---

## A

### Administrative data

Data collected by government agencies or other organizations, not for research purposes but for record-keeping (keeping track of registrants or transactions).

### Algorithm

A set of instructions given to a computer to perform a specific task.

### Artificial intelligence (AI)

The ability of a computer to perform tasks commonly associated with intelligent beings, such as the ability to reason or to learn from past experience. An example is the smart personal assistant Siri.

## B

### Big data

A massive volume of both structured and unstructured data that is too large to process using traditional software techniques. One of its uses is to obtain insights into consumer behavior by analyzing massive volumes of data collected from social media.

### Blinded experiment

An experiment in which information is withheld from participants in order to reduce bias. A blind can be imposed on any component of an experiment, including subjects, data, etc.

### Blockchain

An open, distributed ledger that can record transactions between two parties efficiently and in a verifiable and permanent way.

## C

### Coding scheme

A standard that tells the user's machine which character represents which set of bytes. Without this, the machine might interpret the given bytes as a different character than intended.

## D

### Data landscape

An organization's data storage options, processing capabilities and analytics, as well as the applications present in its data environment.

### Data value chain

The evolution of data from collection through analysis and dissemination to the final impact of data on decision-making.

### Data visualization tools (DVTs)

Tools used to represent data visually in order to communicate information clearly and efficiently. Examples include statistical graphics and plots.

### Database

An electronic system that allows data to be easily accessed, manipulated and updated.

### Database management system (DBMS)

System software for creating and managing databases. A DBMS makes it possible for end users to create, read, update and delete data in a database, and serves as an interface between the database and end users.

---

# I

**Infographics**

Graphic visual representations of information or knowledge intended to present information quickly and clearly.

**Internet of things (IoT)**

An extension of internet connectivity into everyday physical objects, which can then communicate with other objects over the internet, and can be remotely monitored. For example, in the future your car might instruct your house to turn on the heating and lights when you are 5 miles from home.

# M

**Mathematical modeling**

An abstract model that uses mathematical language to describe the behavior of a system. It presents knowledge of that system in usable form.

# O

**Open data**

Data that can be freely used, shared and built on by anyone, anywhere, for any purpose.

# P

**Programming**

The process of designing an executable computer program for accomplishing a specific computing task. Programming involves tasks such as analysis, generating algorithms and the implementation of algorithms in a chosen programming language.

**Prototype**

A working example through which a new model or a new version of an existing product can be derived.

# R

**Real-time data**

Information that is delivered immediately after collection – for example, traffic Global Positioning Systems (GPS) that show drivers what is going on around them.

# S

**Secondary data**

Research data that has previously been gathered and can be accessed by researchers (as opposed to primary data, which is data collected directly from its source).

**Software**

A set of instructions that tells a computer what to do.