



## Vocabulary:

- "Big data" (or "large datasets", "megadata", "big data" or "massive data") refers to a very large set of data that no conventional database or information management tool can work with and brings together in large families trillions of data bytes we produce every day: messages we send to each other, videos we publish, weather information, GPS signals, [transactional records of online purchases and](#) many more.
- A data set is a collection of related data elements that are associated with each other and are accessible individually or combined, or managed as an entity. A dataset is organized in a data structure. In a [database, for](#) example, a dataset may contain business data (names, salaries, contact information, sales figures, etc.). The database itself can be considered a dataset, as can the bodies of data it contains that are associated with a specific type of information, for example, sales data for a corporate department.
- "High Value" data: refers to data sets that are likely to generate the most economic value, or whose importance to an organization's essential or strategic needs requires special treatment and regulation.
- "Open Data" is the practice of "opening" access to one's data to all, both for consultative purposes and for reuse. It allows the sharing person/structure to contribute to the community and to submit its data to criticism in order to improve it; but also to improve the service rendered to the client/user with better information. It also allows third-parties to incorporate this data into their projects to produce higher value-added services. They can be of public or private origin, produced in particular by a community, a public service, a citizen collective or a company. They are disseminated in a structured manner according to a method and an [open license](#) guaranteeing their free access and reuse by all, without technical, legal or financial restrictions.
- [Artificial Intelligence \(AI\) is the use of a](#) number of techniques to enable machines to mimic a form of real intelligence. AI is being implemented in a growing number of fields of [application](#). [Artificial vision](#), for example, allows the machine to precisely determine the content of an image and then automatically classify it according to the object, [colour](#) or face that has been spotted. Algorithms are able to optimize their calculations as they perform. This is how spam filters become more and more effective as the user identifies an unwanted message or, on the contrary, processes false positives. [Speech recognition](#) is [gaining](#) momentum with virtual assistants that can transcribe natural language and then process [requests](#) either by replying directly via [text-to-speech](#), instant translation or by making an order-related query. While in 2015 the market for



artificial intelligence was worth \$200 million, it is estimated that by 2025 it will be worth nearly **\$90 billion**.