



Publication of the Code of Best Practice on data protection for Big Data projects

The Code, prepared by the Spanish Data Protection Agency and ISMS Forum Spain, identifies the main practical questions to be considered when handling large-scale data

Big Data can be defined as the series of technologies, algorithms and systems employed to gather data on a scale and of a variety never seen before in addition to the extraction of valuable data using advanced analytical systems supported by parallel computing.

One of the main uses of Big Data is establishing correlations and creating consumer patterns and profiles. It is of particular interest to countless sectors that undertake online sales and even physical sales via so-called “loyalty cards”, that make it possible to discover the buying habits of a vast number of consumers. Big Data is also an effective tool in sectors such as health, where there are many examples of its effectiveness, for example, in reducing hospital admission times or predicting future illnesses and health risks.

Despite the huge benefits it brings, Big Data also poses obvious risks when it comes to personal data. Imagine, for example, the indiscriminate use of this data without appropriately protecting people's privacy or without adopting the legal, organisational or technical measures required. The Code sets out guidelines for minimizing or eliminating said risks in such projects, pursuant to the provisions of the new European Data Protection Regulation, which will come into force in May 2018.

Basic principles

The Code recommends considering Privacy by Design in order to ensure that data protection guarantees are included from the very beginning

of project planning. Furthermore, it promotes self-regulation by companies managing Big Data projects by producing codes of conduct on this topic (accountability). It also demands that impact assessments are carried out to assess possible risks, amongst others, in the event that health data is being processed.

Legal and technical aspects

Based on the risks that these processes entail in terms of people's privacy, the Code identifies aspects that must be addressed in order for Big Data projects to comply with data protection regulations. On the one hand, it covers the most important legal aspects to be considered, such as the transparency of information provided in advance to those affected or obtaining their consent and the exercise of their rights, or subsequent uses that had not been envisaged when informed consent was obtained. On the other, the Code reviews the different technical and security issues to be considered as part of these projects. Specifically, it identifies the most common privacy strategies: anonymisation, encryption, access control and traceability.