**RECOMMENDED PRACTICES TO IMPLEMENT IMPROVEMENTS IN THE DATA ANALYTICAL PROCESSES OF THE ORGANIZATION**

The objective of this document is to list the best practices explained at a general level, which should serve as a basis to propose improvements that can be implemented to the data analytics processes. These practices can be adapted by the evaluator to the characteristics and needs of the organization.

This document is divided into four parts: General, Data Acquisition, Data Analysis and Use of Inferences. The General section includes high level practices, proposed to propose solutions from the perspective of analytical governance, personnel experience and regulatory compliance. The section Acquisition of data covers the necessary practices to carry out a correct management of the data of the organization. The Data Analysis section focuses on improving the management of resources for the analysis of information as such. The Use of inferences section includes good practices to extract the maximum value from the data from its analysis.

**GENERAL**

***1.1. Governance***

Define a governance framework that works before, during and after the period in which the activities corresponding to analytics are carried out (acquisition, analysis and use of data). It must encompass both risk assessment and management, as well as appropriate responsibility assignments. Likewise, the main focus of these activities must be the fulfilment of the needs of the main stakeholders and the goals of the organization.

***1.2. Experience***

Make sure that all the people selected to perform the data analysis are properly trained and trained to do so, and that they have the necessary experience in real world systems management, especially those that are necessary for data analytics.

***1.3. Compliance***

Make sure that all the activities carried out are aligned with the laws corresponding to the country where they are carried out.

**DATA ACQUISITION**

***2.1. The domain of the problem***

It must be very clear about the details of the problem to be solved and what is the analysis of data that will be done to achieve this objective.

***2.2. The data sources***

Understand each data source, including:

a. The origin of the data

b. The purpose of the creation of the data

c. The meaning of the data

d. The quality of the data at the time of creation

e. The quality of the data and information at the time of use

***2.3. Data fusion***

If the data are to be merged from several sources, evaluate the compatibility of the various collections, records and data elements, taking into account the provenance, purposes, meaning and quality of the data, and the possible impact of data coincidence errors and erroneous assumptions.

***2.4. Data scan***

In case of cleaning the data, evaluate the reliability of the processes for that purpose and the possible impacts of erroneous assumptions and erroneous changes.

***2.5. Identity protection***

Apply the best techniques to assign pseudonyms or hide identities of people or organizations, in case the information to be analysed is very sensitive.

***2.6. Data security***

Minimize the risks arising from the acquisition, storage, access, distribution and retention of data, and manage the inevitable risks.

**ANALYSIS OF DATA**

***3.1. Experience***

Ensure that all those involved in data analytics activities meet the following requirements:

a. They must be professionally qualified

b. They must have adequate training in specific tools and processes

c. They should be familiar with the reality that the data projects.

***3.2. The nature of the tools***

Understand the pros and cons of the data analysis tools that are considered for their use.

***3.3. The nature of the data processed by the tools***

Understand the conclusions that can be drawn from the data processed by analytical tools, as well as the extent to which the data are consistent with these conclusions, that is, cautiously measure the quality of the data.

***3.4. The suitability of the tool and the data***

Have the ability to select the right tool for each type of data or data set.

***3.5. Inappropriate data***

Do not apply data analysis without having verified the quality of these data, their alignment with the objectives of analysis and the coherence of these with the real world.

***3.6. Logically understandable fundamentals***

Do not use tools that produce information that is not very comprehensible and lacks transparency.

**USE OF INFERENCES**

***4.1. The impacts***

Identify the impact on the stakeholders that are part of the analysis, taking into account the quality of this analysis and the process to carry it out.

***4.2. Evaluation***

Analyse the advantages and disadvantages of the material impacts caused by analytical decisions. It is also important to analyse the reliability and quality of the process by which such decisions were reached.

***4.3. Alignment with real events***

To extract a sample of the results of the analysis to observe its relation with the real world and thus to be able to verify the reliability of the analyzed data. Compare with external sources, preferably with information issued by state organizations.

***4.4. Contingency***

Design, implement and maintain protection and mitigation measures, accompanied by controls that ensure that they comply with the provisions, based on the potential impacts of the inferences drawn.

***4.5. Proportionality***

Consider the reasonableness of the decisions taken from the analysis, before assuming a commitment to them.

***4.6. Discussion***

Ensure that decisions based on analytics are transparent and that interested parties have access to information to exercise their right to express their opinions or debate.

***4.7. Effective communication***

In the event of a decision based on analytics that negatively impacts one or more interested parties, communicate in the appropriate manner and provide the opportunity to object to said decision.

***4.8. Post-implementation review***

Ensure that actions and results are audited, and that adjustments are made to reflect the findings.

***4.9. Justification understandable at the human level***

Do not make decisions based on inferences generated automatically by analytical tools, especially those that may have a negative impact on the assets of the organization or the stakeholders.

***4.10. Haste actions***

Do not make decisions based on analytics without having met with all the interested parties in which the necessary debates are carried out to reach a final action.

***4.11. Automated decision making***

Do not delegate any decision that has potentially negative effects to a device without first having an assessment demonstrating the reliability of the device.

# Bibliography

Clarke, R. (2017). Guidelines for the responsible application of data analytics. *Computer Law & Security Review*.