

Atos Breakout session
12 09 2019 – 10.30

Ethics by Design, an approach to
trustworthy, social responsible AI

Celestino Güemes Seoane



Ethics by Design, an approach to trustworthy, social responsible AI

Abstract:

- ▶ Any technology raises, sooner or later, complex ethical dilemmas. However, in the case of AI and other automation-related technologies, these issues can be quite pervasive, and complex to tackle, like the cases of digital bias or the lack of transparency of many popular Deep Learning modules.
- ▶ What is important is the impact that is already becoming evident in many fields. Discrimination of facial recognition systems, issues with policing & justice systems that show clear biases, the unclear limits of accountability of autonomous systems or even safety issues create clear direct harm. But what is worse it is that they may have deeper, indirect effects in the way people behave, a slow erosion of human rights that in many societies are considered part of a core "social contract".
- ▶ Problem is not only a technical problem of how to detect data biases, or how to explain black box models, but much profound in the way that solutions that are based in AI are designed, built and operated. This demands new technical and methodological approaches.

Ethics by Design, an approach to trustworthy, social responsible AI



- ▶ **"Ethics by design"** is a multidisciplinary approach to development of AI & Automation intensive systems that has Ethics as a "first principle" in the whole solution lifecycle.
- ▶ "multidisciplinary approach" as it combines different kind of elements: tools, methodologies, governance models, cultural change management and regulation compliance.
- ▶ "AI & automation intensive systems": even if some concepts may be generic to any kind of development, the focus is in systems with a strong AI part, or that rely on autonomous models of working.
- ▶ "Ethics as first principle": like many other topics (security, accessibility, privacy) it needs to be considered even before the solution exists. It is something that you can't "bolt on" after the fact.
- ▶ "whole solution lifecycle": end to end view, from the initial definition of the use case, to the exploitation and operation of the system.

Ethics by Design, an approach to trustworthy, social responsible AI

Presenter:

Celestino Güemes Seoane

L System/Solution Business Architecture
Senior Expert

