Jožef Stefan Institute
Ljubljana, Slovenia

United Nations Educational, Scientific and Cultural Organization

International Research Centre on Artificial Intelligence under the auspices of UNESCO

AI Regulation – quick overview

Marko Grobelnik

Artificial Intelligence Lab, Jozef Stefan Institute

UNESCO IRCAI

Ljubljana, Slovenia

Marko.Grobelnik@ijs.si

Name of the game:
Definition of AI

Informal definition of AI

- AI is exactly the opposite from what is happening in the video…
- …instead of living beings mimicking machines, AI is supposed to make machines imitating living beings.

OECD AI Definition (OECD 2019)

(the only AI definition politically approved by 44 countries)

"An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments.

It does so by utilising machine and/or human-based inputs to:

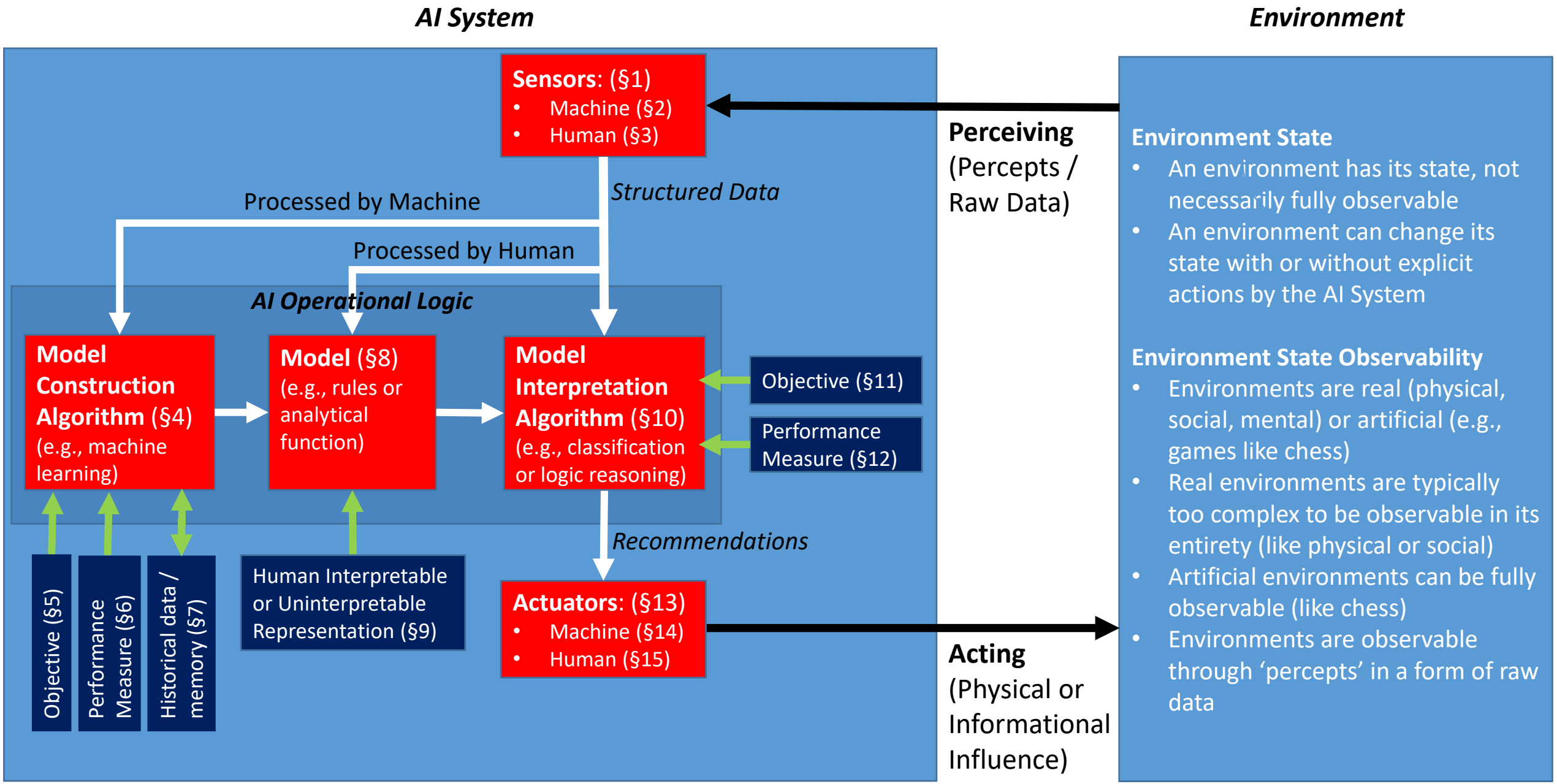
***i)** perceive and/or analyse real and/or virtual environments;*

***ii)** abstract such perceptions/analyses into models manually or automatically; and*

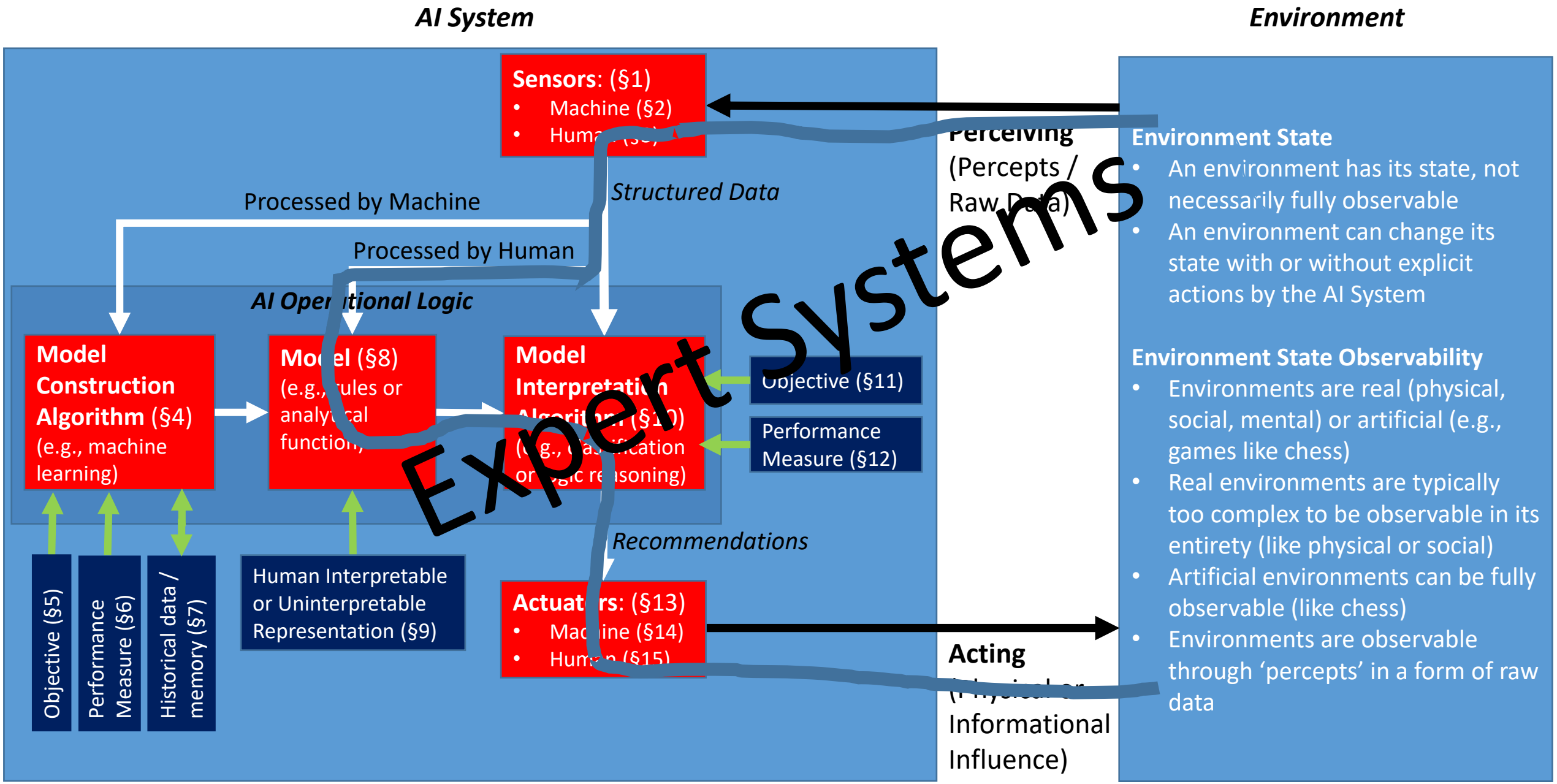
***iii)** use model interpretations to formulate options for outcomes.*

AI systems are designed to operate with varying levels of autonomy."

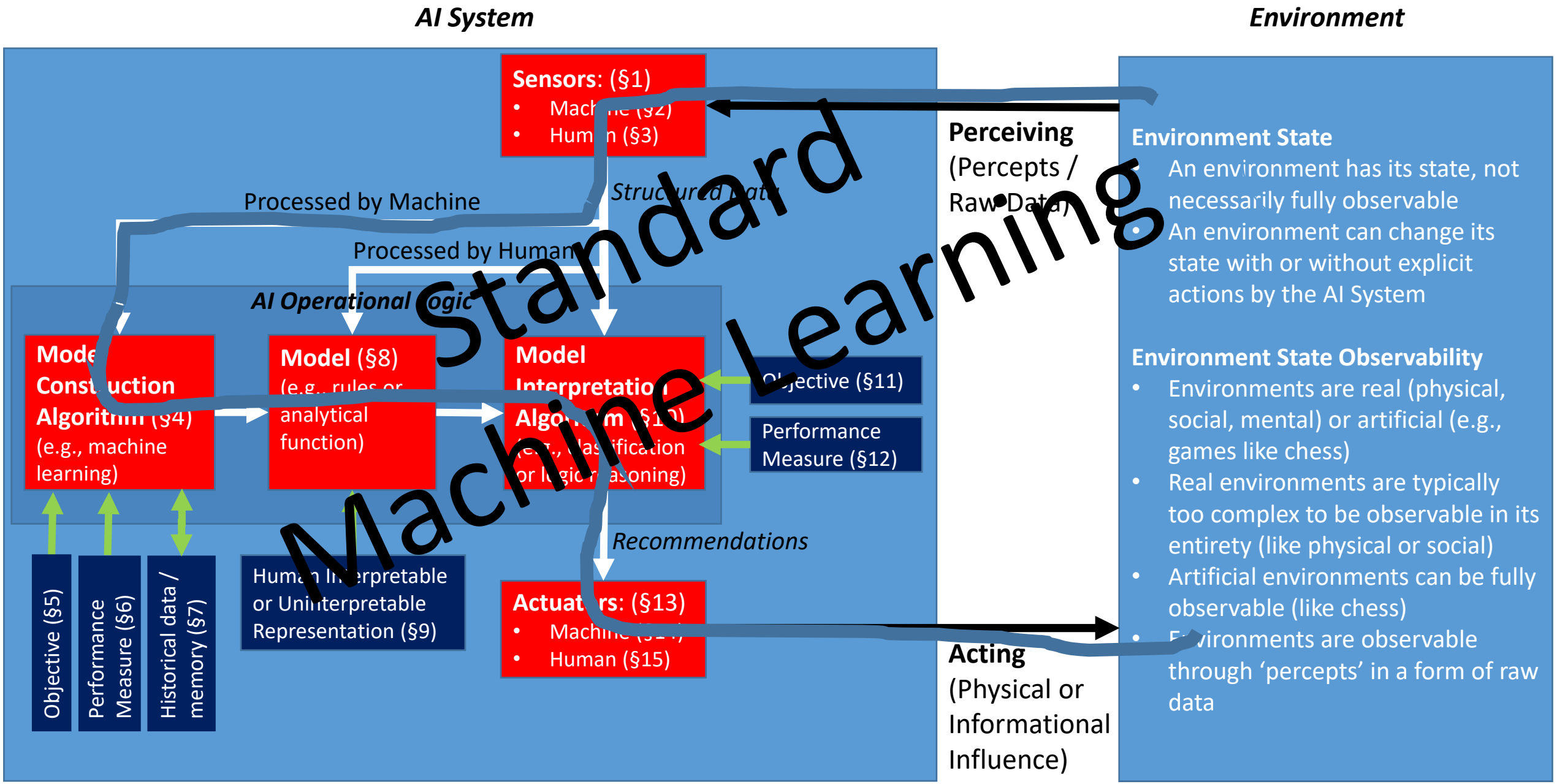
AI System as defined by OECD



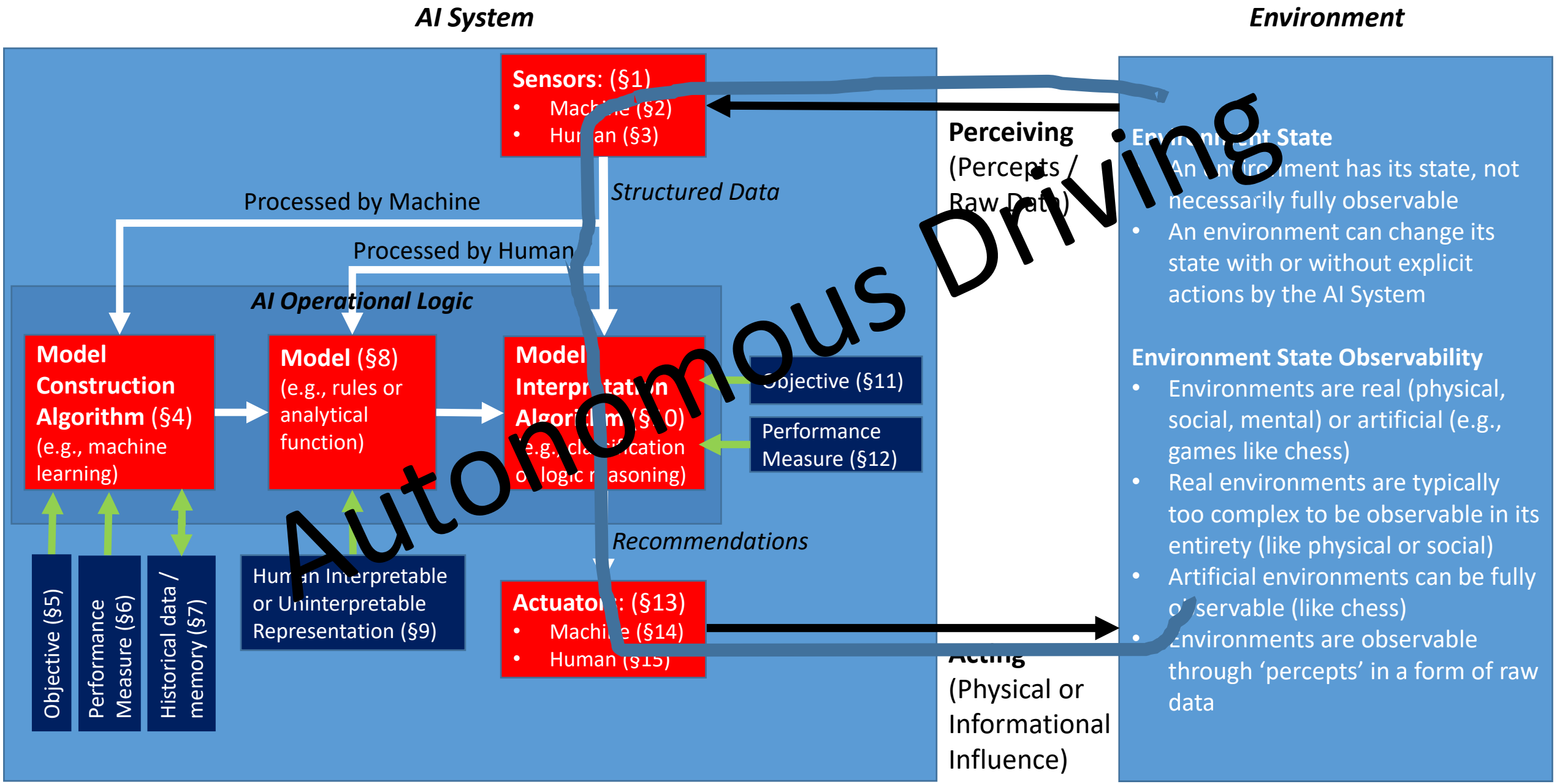
AI System as defined by OECD



AI System as defined by OECD



AI System as defined by OECD



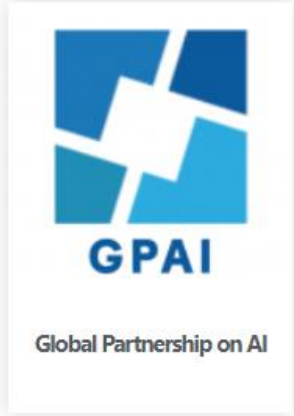
International organizations involved in AI policy documents (https://globalpolicy.ai/)

Human rights (binding)

General (binding)

Human rights

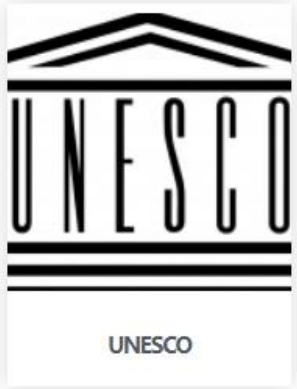
General

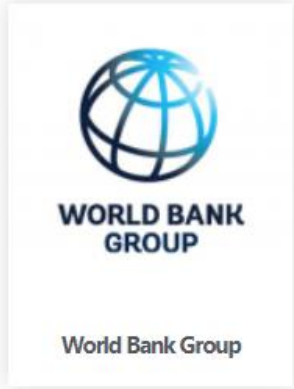


Economy

Digital context

Ethics



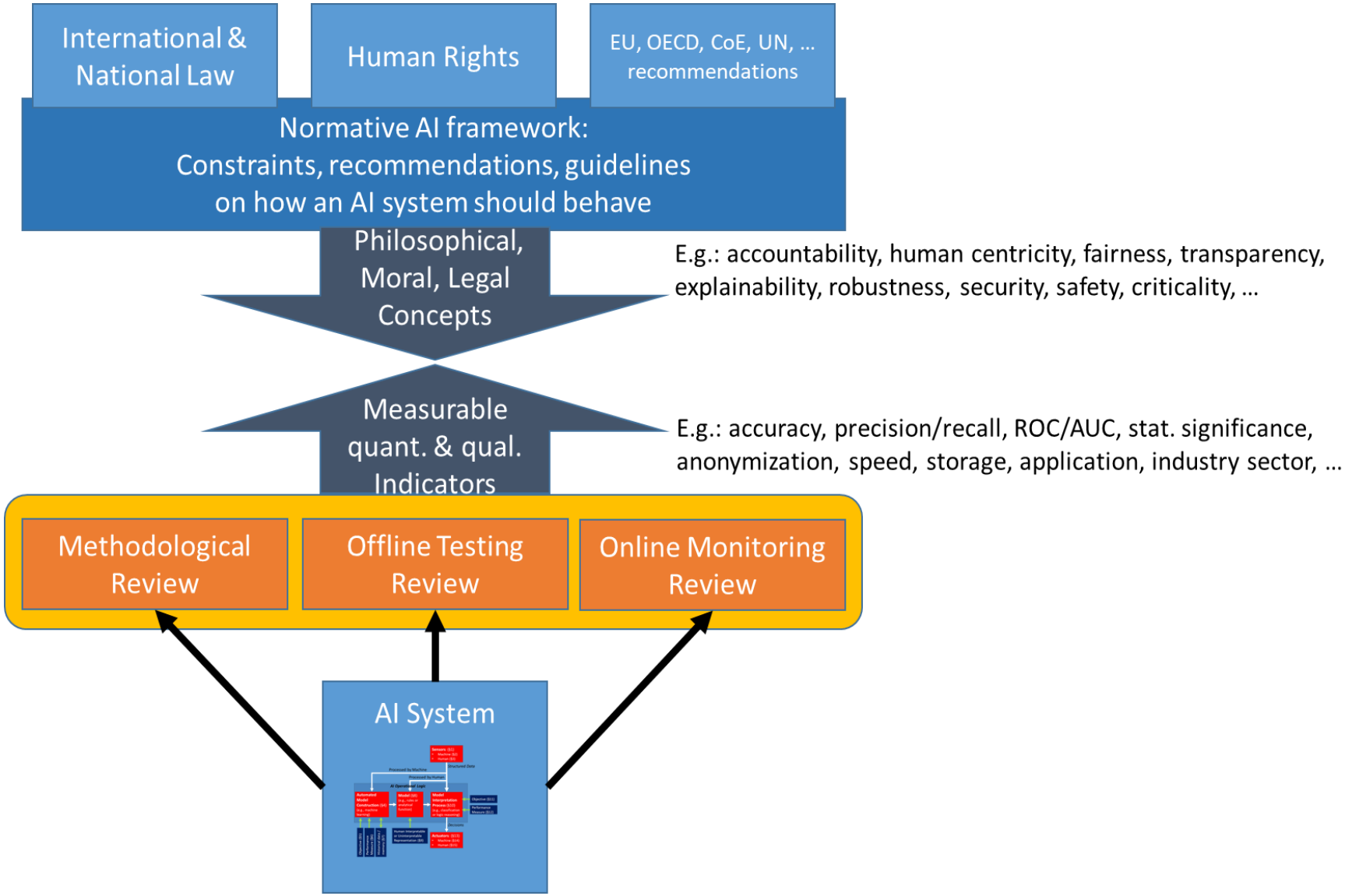


Who is regulating AI on international level?

- EC AI White Paper (2020)
  - https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
- EC AI Act (2021-…)
  - https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence-artificial-intelligence
- OECD AIGO (2018-2019)
- OECD ONE AI (2020-2021)
  - https://oecd.ai/
- Council of Europe CAHAI (2020-2022)
  - https://www.coe.int/en/web/artificial-intelligence/cahai
- UNESCO AHEG (2020-)
  - https://en.unesco.org/events/ad-hoc-expert-group-meeting-declaration-ethical-principles-relation-climate-change
- UN Digital Cooperation 3C (2019-)
  - https://www.un.org/en/digital-cooperation-panel/
- ISO SC42
  - https://www.iso.org/committee/6794475/x/catalogue/

Normative vs technical indicators

- *High level view to the methodological approach on bridging the gap between normative systems (on the top) and technical AI systems (on the bottom).*
- *The gap appears between the abstract concepts used in normative documents and technical indicators measurable from a technical system.*

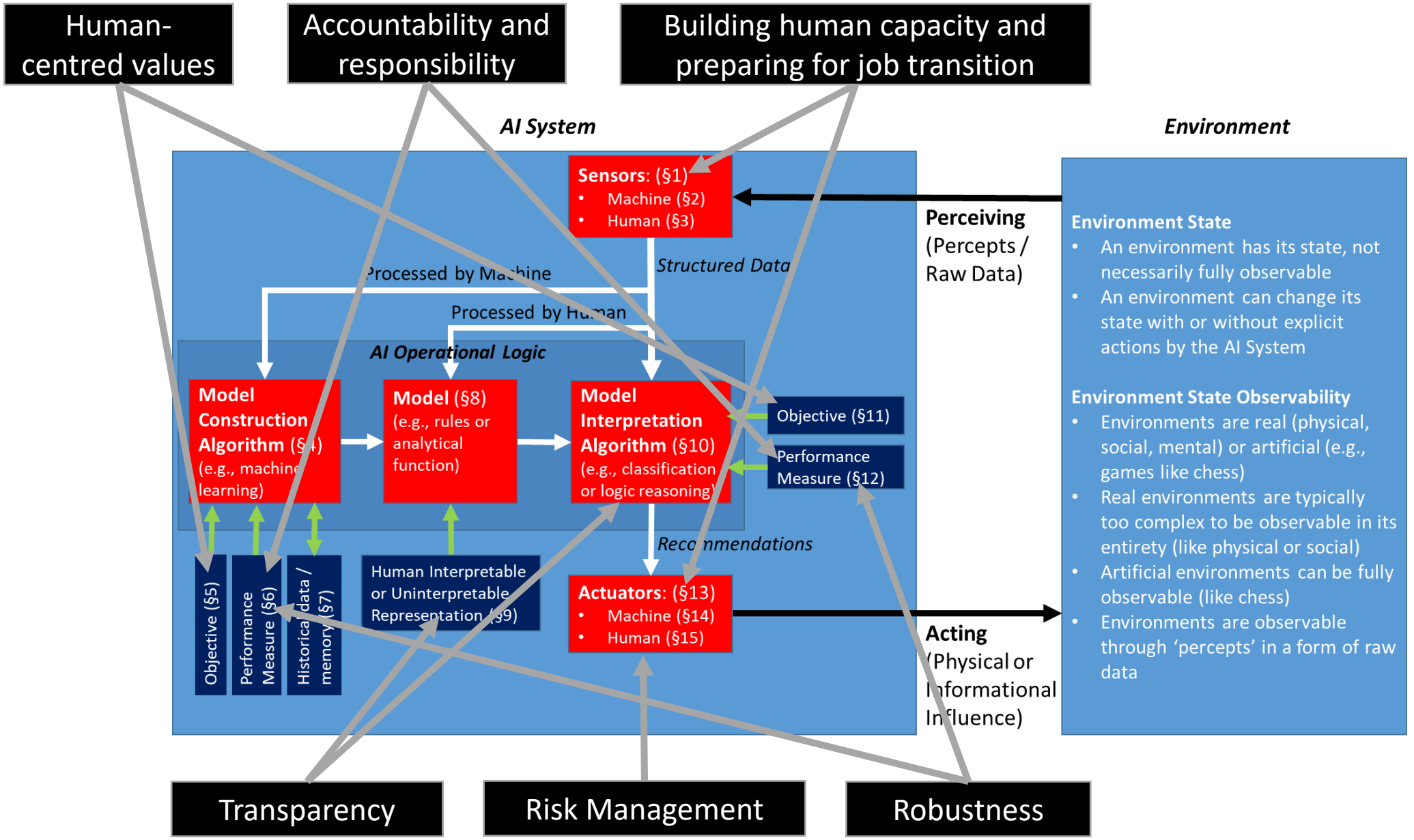


Seven key requirements for trustworthy AI systems
(across multiple regulatory initiatives)

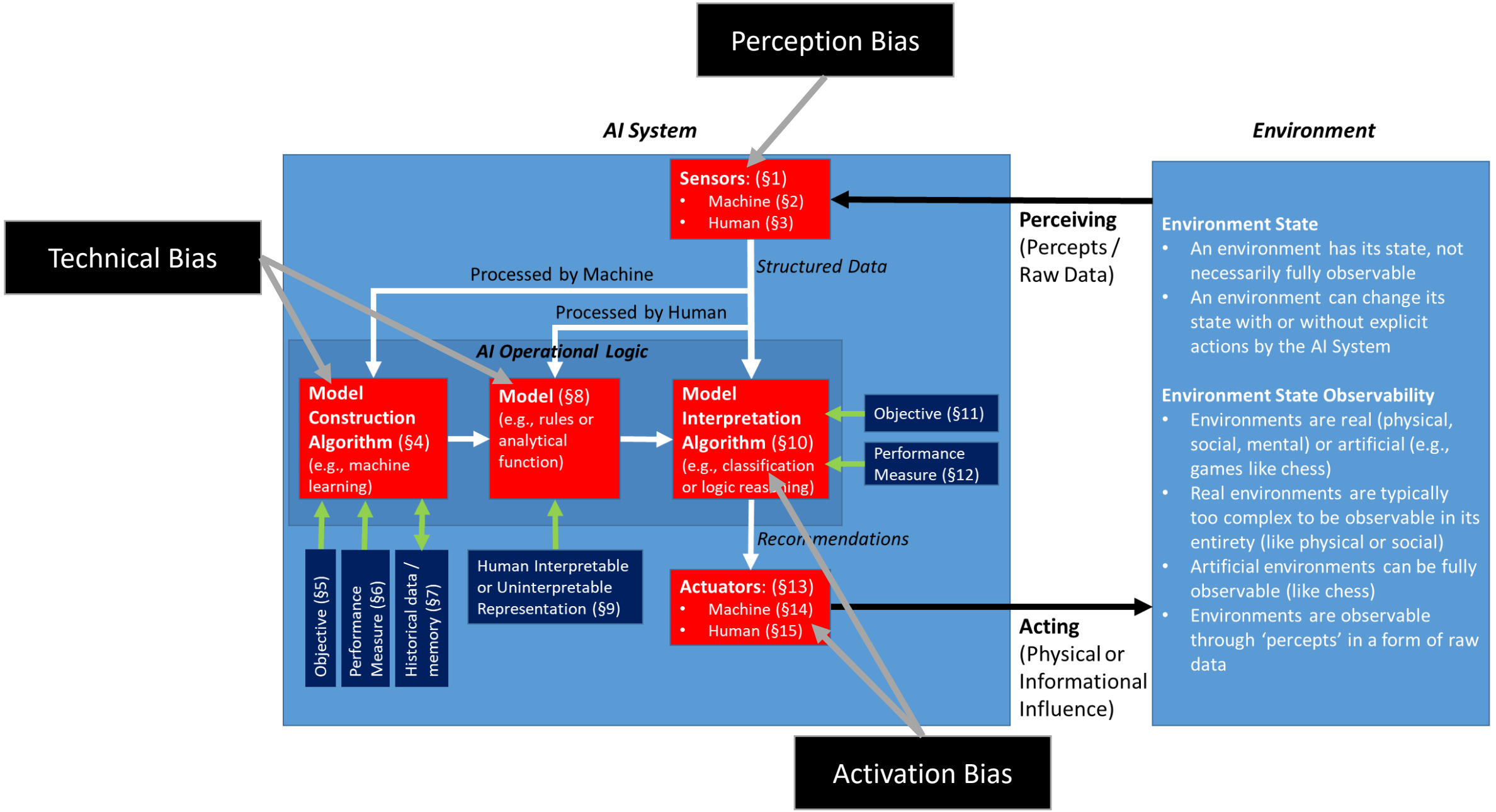
- **Human agency and oversight**: AI systems should empower human beings.
- **Technical Robustness and safety**: AI systems need to be resilient and secure.
- **Privacy and data governance**: respecting privacy and data protection & data governance mechanisms
- **Transparency**: the data, system and AI business models should be transparent & tracable. Humans need to be aware that they are interacting with an AI system.
- **Diversity, non-discrimination and fairness**: Unfair bias must be avoided.
- **Societal and environmental well-being**: AI systems should benefit all human beings, including future generations.
- **Accountability**: Mechanisms should be put in place to ensure responsibility and accountability for AI systems and their outcomes.

https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

AI System and relation to higher level principles

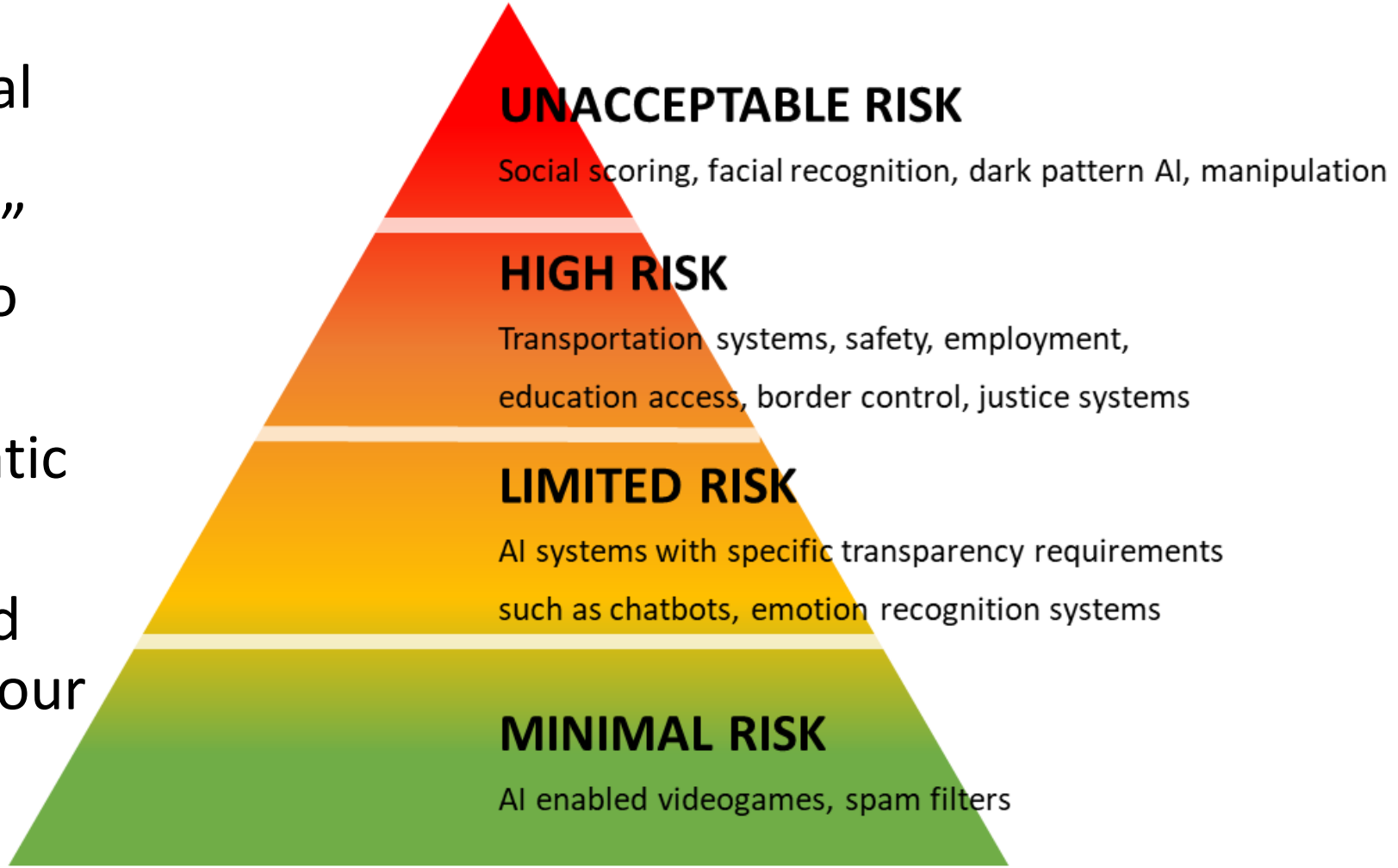


AI System and sources of various types of biases



Risk based approach to regulate AI

- EC and other international organizations decided to use "*risk based approach*" as the main philosophy to form AI regulation
- …the approach is pragmatic and practical
- It is based on the pyramid classifying AI systems in four classes of risk



Summary

- AI policy & regulatory efforts started in 2018 are maturing

- In 2019 OECD adopted AI principles and definition
- In 2021 Council of Europe, UNESCO adopted AI docs on AI & ethics and AI & human rights
- In 2022 EC will adopt EC AI Act, the first legally binding AI regulatory document