

# ETHICAL AI ASSESSMENT: CONCEPTS

Worker's oversight on AI systems at use in the workplace is essential to ensure only legitimate, fair, transparent and safe systems are used. A human-centric approach to the development, use and implementation of AI systems will ensure they act as designed, protecting workers and consumers. To address these issues, see the infographic below, which describes the key concepts, requirements and principles of ethical AI use. This is followed by a checklist, used to assess each step in how the AI systems that are used in your workplace can be assessed. Not everything is an AI system, they are more complex than traditional automated data processes. These differences can be seen when comparing the two methods:

## Automation

- Designed for simple, repetitive tasks
- Cannot evolve or learn
- Rule-based operation
- Generates data but cannot analyze it

## Artificial Intelligence

- Designed for more complex, non-repetitive tasks
- Designed to analyze and react to its environmental data
- Evolves or learns based on previous and current data
- Helps organizations analyze data
- Can identify pattern

## PRINCIPLES

AI systems need to follow 4 key principles to make them.

Respect for  
human  
autonomy

Prevention  
of harm

Fairness

Explicability

## Key Requirements

The principles set 7 key requirements for systems, that they must meet to be considered trustworthy.

Human agency and oversight

Technical robustness and safety

Privacy and data governance

Transparency

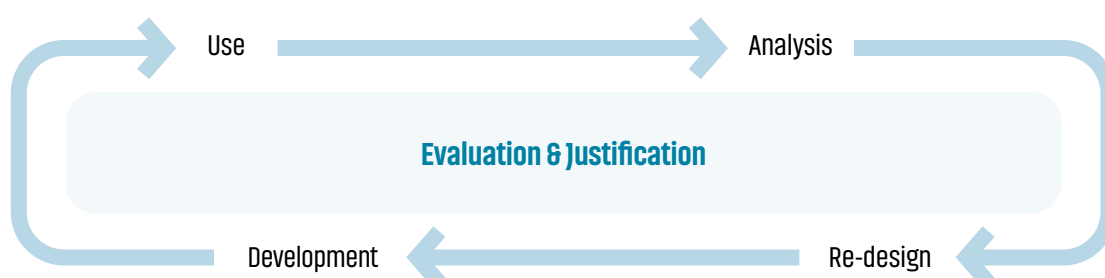
Diversity, non-discrimination and fairness

Societal and environmental wellbeing

Accountability

## AI Lifecycle

These principles need to be assessed in different parts of the AI Lifecycle as these systems are constantly developed and subject to evaluation and oversight.



# ETHICAL AI ASSESSMENT: CHECKLIST

**How & what to assess regarding the AI being used in your workplace: How can you make sure that the 7 key requirements are met? This checklist can be used to assess if the AI in your workplace respects the principles set for ethical systems.**

## 1. HUMAN AGENCY AND OVERSIGHT

- ☐ Humans can easily discern when they are interacting with the AI system vs. a human
- ☐ Humans can easily discern when and why AI system is taking action or making decisions.
- ☐ Improvements will be made regularly to meet human needs and technical standards
- ☐ AI system decisions
  - ☐ Are explained clearly
  - ☐ Are reversible
  - ☐ Can be overridden by humans

## 2. TECHNICAL ROBUSTNESS AND SAFETY

- ☐ Providing understandable security methods
- ☐ Resilience to attacks
- ☐ Making AI system robust, valid and reliable

## 3. PRIVACY AND DATA GOVERNANCE

- ☐ Respecting privacy and data rights, GDPR compliance<sup>1</sup>
- ☐ Only necessary data is collected

## 4. TRANSPARENCY

- ☐ The purpose, limitations, and biases of the AI system are explained in plain language
- ☐ Data sources have unambiguous respected sources, and biases are known and explicitly stated
- ☐ Algorithms and models are proportionate to the aim and verifiable

- ☐ Confidence and context are presented for humans to base decisions on
- ☐ Transparent justification for recommendations and outcomes is provided
- ☐ Straightforward and interpretable monitoring systems are provided

## 5. DIVERSITY, NON-DISCRIMINATION AND FAIRNESS

- ☐ Identifiable and discriminatory biases are removed in the data collection phase
- ☐ System's purpose, constraints, requirements and decisions are analysed and addressed in a clear and transparent manner
- ☐ Applications using AI system is user-centric and accessible by design
- ☐ Stakeholders are consulted throughout system life cycle

## 6. SOCIETAL AND ENVIRONMENTAL WELL-BEING

- ☐ Impact of development, deployment, use process and entire supply chain, are assessed.
- ☐ Effects on social agency are carefully monitored and considered.
- ☐ System is assessed from a societal perspective: institutions, democracy and society at large.

## 7. ACCOUNTABILITY

- ☐ Auditability
- ☐ Minimising and reporting negative Impact
- ☐ Documenting trade-offs
- ☐ Ability to redress

<sup>1</sup> To familiarise yourself on what data is protected by the GDPR, please see [What is personal data? | European Commission \(europa.eu\)](https://european-commission.europa.eu/what-is-personal-data/)