



Artificial Intelligence applied to  
environmental monitoring

Technical challenges and  
limitations of environmental AI

Vittoria Mascellaro

**IR0000032 – ITINERIS, Italian Integrated Environmental Research Infrastructures System**  
(D.D. n. 130/2022 - CUP B53C22002150006) Funded by EU - Next Generation EU PNRR-  
Mission 4 “Education and Research” - Component 2: “From research to business” - Investment  
3.1: “Fund for the realisation of an integrated system of research and innovation infrastructures”



## Module 4: Governance and AI Act

**How does the digital revolution transform our views on values, good behaviors, and the kind of sustainable and equitable innovation?**



Cabinet Office

Version 1.0  
Published 19 May 2016

## Data Science Ethical Framework

Data science carries both huge opportunities and a duty of care. Technology is changing so rapidly; as are the public's views. In this new and changing landscape, this document is not about creating additional hurdles, but rather about making innovation easier. It does this by bringing together the relevant law in the context of new technology, and prompting consideration of public reaction so that government data scientists and policymakers can be confident to innovate appropriately with data.

Developing the ethics around data science can't be done by government alone. This framework is a first iteration - a beta, if you like - of a set of principles wider than the legal framework, to help stimulate innovative and responsible action.

I look forward to listening to, and participating in that debate.

**The Rt Hon Matt Hancock MP**  
**Minister for the Cabinet Office and**  
**Paymaster General**

## Why data science ethics are important

### Who is this guidance for?

This guidance gives those analysing or making policy or operational decisions with data the confidence to innovate. It balances the use of new data and techniques with respect for privacy and makes sure no-one suffers *unintended* negative consequences. An introduction to data science can be found [here](#).

### Why is guidance needed for data science?

Data science is a new practice for government which provides opportunities to create insight and improve public services. Digital advances are producing huge amounts of new forms of data, allowing computers to more quickly process this data and makes decisions without human oversight. This creates new opportunities and many new challenges we have not had to consider before.

The law (e.g. the [Data Protection](#) and [Intellectual Property](#) Acts) sets out some important principles about how you can use data. And analytical, health and other professions have high standards for the quality and integrity of data processes. Those working with data should be aware of these and always act within them. But these are often in different places and not written with data science in mind. This guidance gives people the confidence to innovate by bringing together these laws and standards in the context of the rapidly evolving data landscape.

Public attitudes to data are changing. Working with data in a way which makes the public feel uneasy, without adequate transparency or engagement, could put your project at risk and also jeopardise other projects across government. Consideration of public attitudes and communication with them is key: most people are data pragmatists if told how society will benefit and how risks are managed.

Rather than creating additional hurdles this guidance makes it easier to innovate by helping you both navigate the legal aspects applicable to data science and think through some of the ethical issues which sit outside the law.

### How to use the guidance

Data science projects have a number of stages; discovery work to explore what it is possible to do with the data; the actual delivery; refining the accuracy of the insight; and the ongoing use of that insight by policymakers or operational staff. This guidance will help you think through the methodology and ask appropriate questions about how the project is conducted at each stage.

The guidance gives six principles which are based on existing law. **Fundamentally, the public benefit of doing the project needs to be balanced against the risks of doing so.**

- 1 Start with clear user need and public benefit
- 2 Use data and tools which have the minimum intrusion necessary
- 3 Create robust data science models
- 4 Be alert to public perceptions
- 5 Be as open and accountable as possible
- 6 Keep data secure

The guidance starts with a summary and checklist against the six principles and then goes on to explain each principle in more detail with real examples of where data science has been used well and less well, and practical suggestions of what you can do to act ethically. The Information Commissioner's Office has confirmed that the checklist can form the basis of a Privacy Impact Assessment.

This guidance is based on existing law. The [exemptions](#) within the Data Protection Act around crime, fraud and national security still apply.

The guidance will be iterated and developed with feedback from departments and external stakeholders. It is designed to be iterated as it is used, and is shared in the expectation that it will encourage feedback and further improvement. It also complements other ethical frameworks for analysis such as those relating to health data and from the [National Statistician](#).

# Data Science Ethical Framework

- A document that aims to make innovation easier to understand.
- To achieve this, it attempts to bring together relevant regulations and user feedback so that policymakers and scientists can work effectively on data management.
- It is a guide that balances the use of new data and techniques with respect for privacy, ensuring that no one faces unintended consequences.
- Data science is a practice used by governments to improve public services.

# Data Science Ethical Framework

The guide establishes six principles based on existing laws:

- Start with a clear user need and a public benefit
- Use data and tools that are minimally intrusive
- Build robust models
- Stay vigilant regarding public perceptions
- Be open and accountable
- Take data security into consideration

# Digital ethics

- Understanding which of these moves are the best is the task of **digital ethics**.
- It is the branch of knowledge that deals with moral issues related to data and information, aiming to support morally good actions.
- It shapes both digital governance and regulation.

## **SOFT ETHICS**

- Post-compliance ethics of established legal norms.
- Considers what should or should not be done beyond existing regulations.

## **HARD ETHICS**

- Precedes and helps shape legal regulation.
- Analyzes what is generally right or wrong.

**The real challenge is to anticipate ethical development, and to do so, we must be able to assess what is truly feasible, prioritizing what is sustainable from both an environmental and social perspective.**

### ADOPTING AN ETHICAL APPROACH OFFERS A DOUBLE ADVANTAGE:

- **Soft ethics** can provide an opportunity-driven strategy
- **Ethics** offers a solution for effective risk management
- This is possible when there is **adequate legislation, public trust, and clear accountability.**

## WHY IS IT SO URGENT?

Because ethical evaluation shapes public opinion



which determines what is politically possible



and thus what is legally enforceable

.

# Debate on the Ethics of algorithms

- **Algorithms** = mathematical constructs that can be developed into a program or a configuration.
- Algorithms underpin the services and infrastructures of society (e.g., recommendation systems).
- They are used in schools, hospitals, financial institutions, courts, local government bodies, and national governments.
- Their growing use prompts us to reflect on their ethical implications.

**Are algorithms neutral?**

# Mapping the ethical issues related to algorithms

Machine learning algorithms ("automata" or "semi-automata" because their results are induced by data and are not deterministic) can be used to:

- Transform data into evidence for another outcome
- Trigger or motivate an action that may have ethical consequences
- Assign responsibility for the effects of actions that an algorithm may trigger

## Six ethical issues

- 1. Inconclusive evidence**
- 2. Inscrutable evidence**
- 3. Misleading evidence**
- 4. Unfair outcomes**
- 5. Transformative effects**
- 6. Traceability**

# Inconclusive evidence

They refer to the way in which machine learning algorithms identify associations and correlations between data variables, but not causal connections → thus generating **probabilistic outputs**



## **Apophenia**

The tendency to perceive meaningful patterns where none actually exist.

# Inconclusive evidence

## **PROBLEM:**

They generate ethical problems because non-causal indicators divert attention from the underlying cause of the problem — the problem is not solved by simply having more data.

## **SOLUTION:**

Through independent validation of algorithms and practices that ensure data preservation and reproducibility.

# Inscrutable evidence

They refer to issues of non-transparency, lack of control, and accountability



1. Cognitive impossibility for humans to interpret algorithmic models and data systems
2. Lack of tools to keep track of large volumes of code and data
3. Poor structuring of data and code causing unreadability
4. Continuous updates (malleability of algorithms)
5. Human influence on the model

# Inscrutable evidence

Transparency is not an ethical principle in itself, but a tool for ethical purposes



## **Solution:**

Highlight the causes that lead to non-transparency

- 1.The AI Now Institute (New York University) has produced a guide for algorithmic impact assessment which includes an interpretable explanation of internal processes (such as Explainable AI made available by Google and IBM).
- 2.Implementation of public education

# Misleading evidence

**Unintended *biases* are of concern**



Caused by algorithmic formalism



Adherence to prescribed rules that ignore the complexity of the world



1. Inability to model the entire system to which the social criterion will be applied
2. Inability to understand how replicating a solution in one context can be misleading or inaccurate in another
3. Inability to account for fairness
4. Inability to understand that the introduction of a new technology may change the behaviors and/or values of a system
5. Inability to understand that the best solution may be dissociated from the technology

# Misleading evidence

**Biases derive from the data used**



**Solution:**

Remove certain specific data variables (e.g., gender, race — prohibited by anti-discrimination law and data protection regulations)

**And what about invisible biases?**

Through external auditing and public disclosure of a model along with the data and metadata used.

# Unfair outcomes

Related to the need for algorithmic fairness



DEFINED BY 4 PRINCIPLES

- **Anti-classification:** refers to protected categories (race, gender)
- **Classification parity:** a model is fair if common measurements are equal across protected groups
- **Calibration:** measures how well an algorithm is calibrated
- **Statistical parity:** views fairness as an equal estimate of the average probability across protected groups

# Unfair outcomes

**Removing protected variables does not mean eliminating the discrimination rate because invisible variables exist.**



**Solution:**

Introduce a third party that attempts to eliminate possible discrimination.  
A collaborative model focused on the data resource.

# Transformative effects

They are intertwined with issues concerning autonomy.

1. Pervasive and proactive distribution of algorithms in shaping user choices
2. Limited understanding of algorithms by users
3. Lack of power over algorithmic outcomes



## **Solution:**

Autonomy must be constrained and reversible

# Traceability

## It concerns the issue of responsibility

On one hand, it seems impossible to define moral responsibility due to the human-machine hybridization /  
On the other hand, it is unthinkable to trace data back to its source.

Lack of transparency + lack of explainability =  
**NEED FOR NEW APPROACHES**

# The European Commission regulation

In April **2021**, the European Commission published a proposal for a **regulation on the European approach to artificial intelligence**.

It includes transparency rules and more specific provisions regarding high-risk systems (such as personnel selection systems and systems assessing the reliability of statements made by individuals to prevent or investigate crimes).

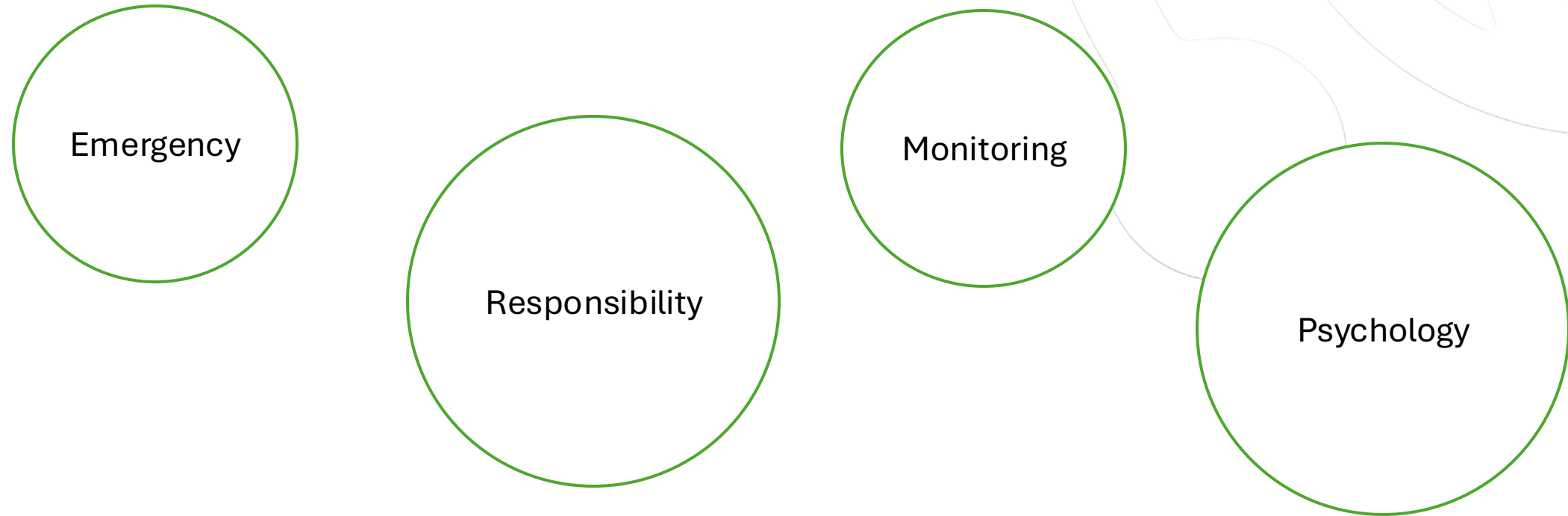
## **It prohibits:**

- **The use of systems employing subliminal techniques on unaware individuals**
- **The deployment by public authorities of systems to assess individuals' reliability based on their social behavior or personality traits**
- **The use of real-time remote biometric identification systems**

## **Five Criminal Areas Affecting AI Crimes (Artificial Intelligence Crimes):**

1. Trade and Financial Markets
2. Harmful and Dangerous Drugs
3. Crimes Against the Person
4. Sexual Crimes
5. Theft, Fraud, Counterfeiting, and Identity Theft

# Concerns Related to AI Crimes (CIA)



# Emergency

It refers to the possibility that artificial agents may act in ways that go beyond our expectations

# Emergency

It refers to the possibility that artificial agents may act in ways that go beyond our expectations

# Responsability

It refers to the possibility that AI systems can undermine existing models of responsibility →  
It would undermine legal certainty → No law would be recognized that could acknowledge the crime itself

## Emergency

It refers to the possibility that artificial agents may act in ways that go beyond our expectations

## Responsability

It refers to the possibility that AI systems can undermine existing models of responsibility →

It would undermine legal certainty → No law would be recognized that could acknowledge the crime itself

## Monitoring

1. It refers to three types of problems: 1. Attribution of non-compliance with regulations 2. Feasibility: concerns cases where artificial agents operate at speeds or levels of complexity beyond the capacity to monitor the rules (e.g., bots on social media) 3. Intersystem actions: refer to experiments showing how to automatically reproduce a user's identity (e.g., on Twitter)

## Emergency

It refers to the possibility that artificial agents may act in ways that go beyond our expectations

## Responsability

It refers to the possibility that AI systems can undermine existing models of responsibility →  
It would undermine legal certainty → No law would be recognized that could acknowledge the crime itself

## Monitoring

1. It refers to three types of problems: 1. Attribution of non-compliance with regulations 2. Feasibility: concerns cases where artificial agents operate at speeds or levels of complexity beyond the capacity to monitor the rules (e.g., bots on social media) 3. Intersystem actions: refer to experiments showing how to automatically reproduce a user's identity (e.g., on Twitter)

## Psychology

It refers to the concern that AI could negatively influence and manipulate a user's mental state to the point of facilitating a crime → Among the effects, gaining users' trust is noted

# AI Act

- **Entered into force:** August 1, 2024
- Some provisions are **already applicable**
- Others require a **transitional period** due to complexity and implementation needs
- Promoted by the **European Commission** to:
  - Support early adoption of the AI Act
  - Encourage developers to comply **before official deadlines**

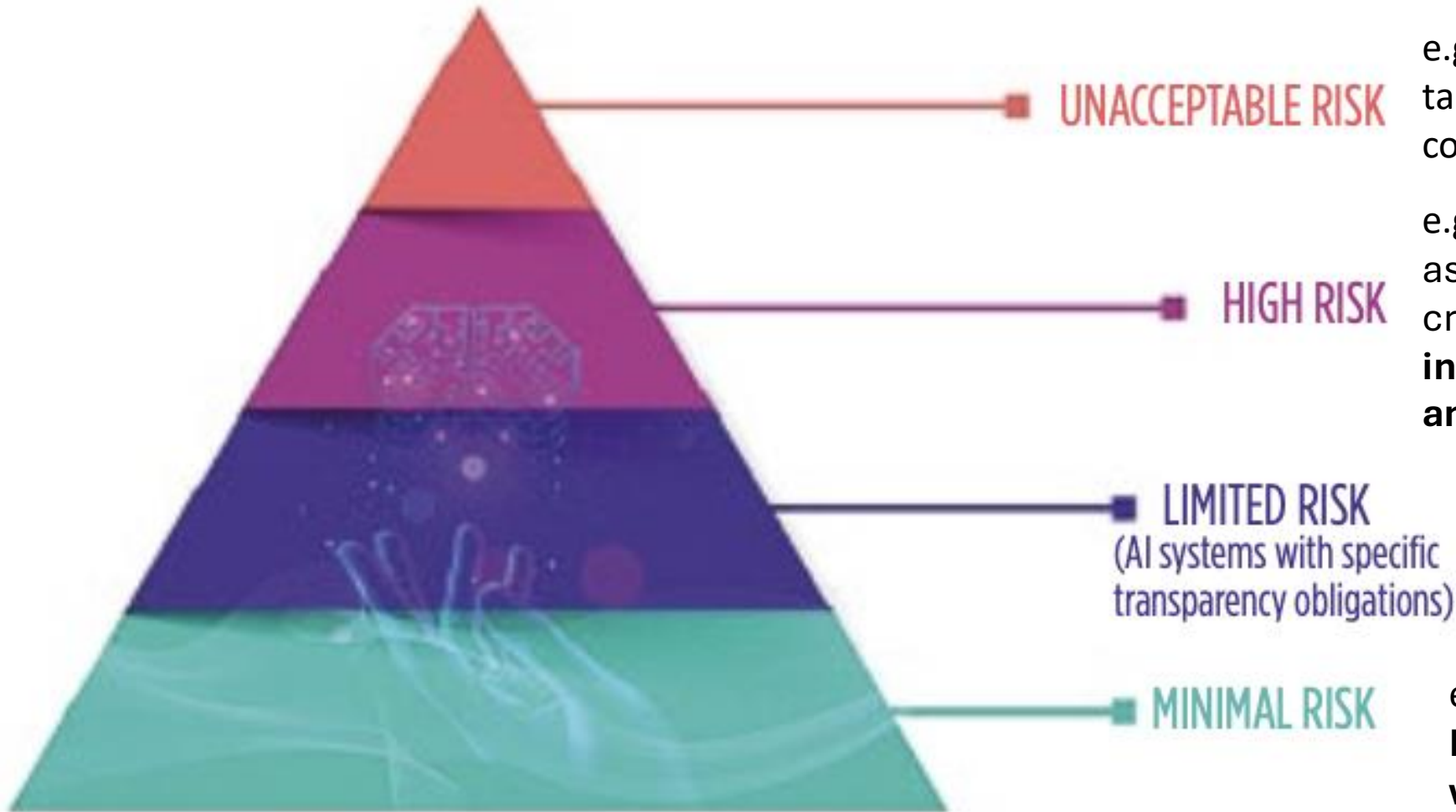
## Pillar 1: Knowledge Sharing through the AI Pact Network

- Access point to the AI Pact community
- Exchange of **best practices**, experiences, and expertise
- Updates and guidance on **AI Act implementation**

## Pillar 2: Facilitating Business Commitments

- Framework to **accelerate adoption** of AI Act measures
- Encourages AI providers/operators to:
  - Prepare in advance
  - Adopt compliance strategies early

# AI Risk Classification in the AI Act



e.g. Social scoring, advertising messages targeted at children, psychological conditioning → **forbidden**

e.g. credit scoring, personnel selection, assisted surgery, police operations, critical infrastructures → **Permission in compliance with AI requirements and prior conformity assessment**

e.g. chatbot → **Permission in compliance with transparency and information obligations**

e.g. videogames, anti-spam systems → **Permission without obligations but with a recommended code of conduct**

# AI Act and enviromental monitoring

Currently, the AI Act does not contain specific articles on environmental monitoring through artificial intelligence. The existing provisions address environmental impact in a limited and non-binding manner.



BREAK