Quaderni di Comunità

Persone, Educazione e Welfare
nella società 5.0

Community Notebook

People, Education, and Welfare in society 5.0

n. 1/2025

HUMAN-CENTRIC APPROACH
TO ARTIFICIAL INTELLIGENCE

edited by
Marco Filoni, Filippo Maria Giordano, Giorgio Grimaldi

eurilink
University Press

Iscrizione presso il Registro Stampa del Tribunale di Roma
al n. 172/2021 del 20 ottobre 2021

Eurilink University Press Srl
Via Gregorio VII, 601 - 00165 Roma
www.eurilink.it - ufficiostampa@eurilink.it
ISBN: 979 12 80164 90 2
ISSN: 2785-7697 (Print)
ISSN: 3035-2525 (Online)

Prima edizione, luglio 2025
Progetto grafico di Eurilink

INDICE

EDITORIALE
Marco Filoni, Filippo Maria Giordano, Giorgio Grimaldi 15

RUBRICA *EDUCATION* 31

1. Intelligenza artificiale generativa nella didattica: verso un approccio umano-centrico
Michele Baldassarre, Francesco Pio Sarcina, Anna Maria Cuzzi 33

2. Approccio plurale all'intelligenza artificiale: sfide etiche e formative nelle istituzioni scolastiche
Sara Pellegrini, Riccardo Sebastiani, Patrizia Ninassi, Emanuela Lampis 59

3. L'intelligenza artificiale nell'educazione: un'analisi degli studenti italiani
Antonio Opromolla 89

RUBRICA *EMPOWERMENT* 97

1. Approccio umanocentrico all'intelligenza artificiale: sfide etiche, sociali ed economiche
Riccardo Mancini, Sara Pellegrini, Riccardo Sebastiani, Debora Glori 99

2. Bridging expectations and realities: the future socio-economic impact of AI
Viviana Condorelli, Fiorenza Beluzzi 121

3. Balancing Innovation and Equity: an Analysis of the European AI Act
Sergio Pappagallo 127

4. La protezione dei dati personali, l'intelligenza artificiale e i traduttori automatici
Federica De Stefani 135

5. Participatory Approaches For The Transition From Automation To Artificial Intelligence (AI): A Case Study
Sara Calicchia, Chiara Colagiacomo, Angela Bagnato, Roberta Pistagni, Bruno Papaleo, Francesca Grosso 145

SAGGI 153

1. Intelligenza artificiale generativa, costruzione di senso e reti sociali: una prospettiva sociologica
Francesco Mattioli 155

2. Intelligenza artificiale e Literacy. Promuovere l'approccio sociologico umano-centrico per superare i pregiudizi e favorire l'inclusione sociale
Danilo Boriati, Mariangela D'Ambrosio 183

3. Riattivare la riflessività: per un modello etico-critico di educazione digitale
Giuseppe De Ruvo 213

4. Trasformazioni digitali nel welfare: intelligenza artificiale e servizio sociale
Roberto Veraldi, Chiara Fasciani 243

5. A Model for Responsible Governance of human-centric AI in the Public Sector
Francesco Niglia 277

6. AI and Democracy: the Role of the European Parliament in Shaping the EU "AI Act"
Raffaella Cinquanta 311

7. L'IA nella gestione delle frontiere dell'Unione europea: un approccio antropocentrico per tutti?
Giulia Maria Gallotta 341

8. Sovranità e indipendenza tecnologica. La sfida e i rischi delle "nuove" intelligenze. Una valutazione di sistema
Giuseppe Romeo 373

3. BALANCING INNOVATION AND EQUITY: AN ANALYSIS OF THE EUROPEAN AI ACT

by Sergio Pappagallo*

Abstract: Questa rubrica illustra la logica regolatoria dell'AI Act europeo (Regulation (EU) 2024/1689), fornendone un'introduzione sistematica attraverso il framework del principio di precauzione. L'analisi muove dalla definizione di intelligenza artificiale contenuta nell'Articolo 3, evidenziando le criticità tecniche e filosofiche che caratterizzano tali sistemi: dai problemi di overfitting e data bias al problema epistemologico dell'induzione nelle inferenze predittive. La rubrica dimostra come il legislatore europeo abbia strutturato una risposta normativa all'incertezza tecnologica mediante un sistema di classificazione del rischio articolato su quattro livelli progressivi: dalle pratiche espressamente proibite ai sistemi ad alto rischio, fino a quelli a rischio limitato e minimo, ciascuno soggetto a obblighi e controlli differenziati.

Parole chiave: AI Act, Innovazione, Classificazione dei rischi, Principio di Precauzione, Problema dell'Induzione.

Abstract: This section illustrates the regulatory logic of the European AI Act, Regulation (EU) 2024/1689, providing a systematic introduction through the precautionary principle framework. The analysis begins with the artificial intelligence definition contained in Article 3, highlighting the technical and philosophical criticalities characterizing such systems: from overfitting and data bias problems to the epistemological problem of induction in predictive inferences. The section demonstrates how the

* PhD Student, Dipartimento di Scienze Umane, Università degli Studi "Link", https://orcid.org/0009-0005-5960-636X, ser.pappagallo@gmail.com.

Accettato Dicembre 2024 - Pubblicato Aprile 2025.

European legislator structured a normative response to technological uncertainty through a risk classification system articulated on four progressive levels: from expressly prohibited practices to high-risk systems, down to limited-risk and minimal-risk systems, each subject to differentiated obligations and controls.

Keywords: AI Act, Innovation, Risk Classification, Precautionary Principle, Problem of Induction.

Introduction

As with most legislative processes, one of the conditions necessary to regulate a phenomenon is that the phenomenon poses risks to the values that societies need to protect, and that the phenomenon is scalable. Think of an example of the first traffic codes developed precisely because of the massive use of automobiles. Similarly, AI has been implemented for a long time already. Still, recent applications and even more recent massive uses of such technologies have made it imperative to have an initial regulation that would bring multiple actors into an international agreement. Besides, the most serious errors in AI systems require specialized skills to be identified and corrected (Binns, 2018). Therefore, it is better to anticipate risk factors. The Artificial Intelligence Act (European Commission, 2024) represents a significant attempt to balance the promotion of innovation with the need to structure the safe development and use of AI, overcoming the practices of ethics washing and protecting society from the most significant risks. This approach reflects the "responsible innovation" concept discussed by Stilgoe *et al.* (2013), which emphasizes the importance of considering new technologies' ethical and social implications during the development process.

1. Conceptual Framework

The AI Act provides in Article 3 a definition of AI that emphasizes adaptivity post-implementation, impacting society. In addition, the definition considers the relationship between input and output. This implies that risk management must refer to this relationship. At a technical level, there are two main problems: overfitting and data bias. Overfitting occurs when a model overfits training data, compromising its ability to generalise to new data (Barocas *et al.*, 2019). Data bias emerges when the dataset does not adequately represent reality, causing potential discrimination and perpetuating existing societal biases (O'Neil, 2016). Also, note that AIs are models that use inductive inferences to make predictions. Therefore, they are subject to the classic philosophical problem of induction because even if the input data were of high quality, inductive inferences could not guarantee true conclusions by their very nature (Henderson, 2022). Therefore, uncertainty is an ineradicable component when it is necessary to obtain an amplifying prediction from the observed phenomenon to the unobserved phenomenon. Still, it is essential to govern it to ensure its manageability on a large scale. To this end, the precautionary criterion is one of the most fitting strategies in the ethical-legal context. The regulatory framework adopted by the EU on AI uses this criterion precisely, imposing preventive measures that are all the more stringent, the more significant the impact and likelihood of risks. The inherent difficulty in accurately predicting all the consequences of the massive use of AI makes the precautionary principle particularly suitable, as it promotes rigorous assessment of AI technologies' risks and long-term implications.

2. Risk Classification

Title II of the AI Act establishes four categories of prohibited practices in using AI. The first concerns subliminal manipulation and deception: using techniques that operate beyond the users' awareness, impairing their ability to make rational choices, is prohibited. The second category concerns the exploitation of vulnerabilities: AI cannot be used to take advantage of conditions such as age, disability, or disadvantaged socio-economic situations. The third prohibition relates to discriminatory social scoring: using systems that score individuals based on social behaviors or personal characteristics is not permitted, as this reduces the person to a mere numerical datum, undermining human dignity and potentially perpetuating social inequalities. Lastly, the fourth category concerns mass surveillance and biometric identification: using real-time facial recognition and biometric tracking systems in public spaces is severely restricted and requires extreme justification and strict oversight. Then, the AI Act defines high-risk AI systems according to two key criteria. The first identifies systems that serve as security components in products already regulated by the EU in Annex II, such as medical devices, vehicles, and toys, which are subject to external compliance evaluation. The second includes systems listed in Annex III of the AI Act, including biometric identification, critical infrastructure management, education, employment, and law enforcement. The legislation provides exemptions for systems with limited impact, such as those that perform circumscribed procedural operations or optimize pre-existing human decisions. These exemptions do not include systems that profile individuals based on job performance, financial status, or behavior.

Next, the legislation defines the limited-risk categories, focusing on transparency obligations. Such technologies typically

interface with users, generate content, recognize emotions, and categorize biometric data. The legislation imposes disclosure requirements for these systems on the artificial nature of the interaction or content generated.

Finally, the minimal-risk category, which represents the majority of AI applications in the market, includes systems such as spam filters, video games, and editing tools, which are subject to minimal regulation. Although the latter do not pose significant risks to fundamental rights, the regulatory framework encourages voluntary adherence to codes of conduct to ensure ethical and transparent practices in technical implementation.

References

Acemoglu, D. (2024). *Redesigning AI: Work, Democracy, and Justice in the Age of Automation*. Cambridge, MA: MIT Press.

Acemoglu, D., & Restrepo, P. (2019). Automation and New Tasks: How Technology Displaces and Reinstates Labor. *Journal of Economic Perspectives*, *33*(2), 3-30.

Albanesi, S., Marinescu, I., & Schaumann, D. (2023). The Impact of AI on Labor Markets: Evidence from Online Job Postings. National Bureau of Economic Research Working Paper Series.

Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and Machine Learning. In *Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency*.

BBC. (2023, April). We'll need universal basic income - AI's godfather'. BBC News.

Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. In *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency*, 149-158.

Briggs, J., & Kodnani, D. (2023). The Potentially Large Effects of Artificial Intelligence on Economic Growth. Goldman Sachs Global Investment Research.

European Commission. (2022, June 16). A European approach to micro-credentials. Education and Training.

European Commission. (2024). AI Act: Regulation on artificial intelligence. Official Journal of the European Union.

Henderson, L. (Winter, 2022). The Problem of Induction. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta & Uri Nodelman (eds.), https://plato.stanford.edu/archives/win2022/entries/induction-problem/.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. New York: Crown Publishing Group.

Smeds, E., & Smeds, E. (2020). The Greener State: Public services for a carbon-neutral Europe. https://core.ac.uk/download/481470986.pdf.

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. Research Policy, 42(9), 1568-158.

Verlaat, T., de Kruijk, M., Rosenkranz, S., Groot, L., & Sanders, M.

(2020). Onderzoek Weten wat werkt: samen werken aan een betere bijstand, Eindrapport (What Works: Final Report). Utrecht: Utrecht University. https://www.uu.nl/en/publication/final-report-what-works-weten-wat-werkt.