

MARIAVITTORIA CATANZARITI*

Etica “artificiale”: un nuovo modello regolatorio?

ENGLISH TITLE

“Artificial” Ethics: A New Regulatory Model?

ABSTRACT

The aim of the paper is to explore the broader scope of the Ethics Guidelines for Trustworthy AI. In particular, the paper focuses on the reasons that led EU to develop an ethical approach to AI, seeking to investigate to which extent it is arguable that the ethical principles for a trustworthy AI should be based on the compliance with fundamental rights. It points out that the symbolic value of fundamental rights as embedded within this non-binding tool, shows the normative vision of EU, mitigating the possible conflict between institutional and private actors involved as well as different interests at stake. Finally, it argues that neither the ethical approach nor the legal design of AI can effectively address the issue of algorithmic inferences and their impact on individuals and society.

KEYWORDS

Artificial Intelligence – Ethics – Legal Design – Inferences – Fundamental Rights.

1. INTRODUZIONE

Come noto, nell’aprile 2019 la Commissione Europea ha affidato a un Gruppo Indipendente di Esperti di Alto Livello (High-Level Expert Group on AI, AI HLEG) la redazione degli Orientamenti etici per la realizzazione di una intelligenza artificiale affidabile¹ e delle raccomandazioni politiche e di investimento. Tale documento è stato oggetto di consultazione con più di 500 partecipanti. Esso è il risultato di un processo regolatorio composito, iniziato con la Risoluzione del Parlamento Europeo del 2017 con la quale si auspicava la creazione della European Agency for Robotics and AI, che avrebbe avuto il compi-

* Research Associate presso lo European University Institute, Firenze.

1. Orientamenti etici per un’intelligenza artificiale affidabile, <https://op.europa.eu/it/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>.

to di offrire supporto tecnico, regolatorio ed etico alle istituzioni, agli Stati Membri e agli *stakeholders* privati e pubblici coinvolti. Tale processo è tuttora *in itinere* e finora ha condotto all'adozione di vari strumenti, tra i quali uno dei principali è stato il White Paper sull'Intelligenza Artificiale del febbraio 2020 per il sostegno di un approccio normativo e orientato agli investimenti², fino alle recentissime Linee Guida per l'uso militare e non militare dell'intelligenza artificiale³. L'obiettivo finale del legislatore europeo è quello di far confluire le differenti iniziative regolatorie esistenti nell'ambito dell'AI in un regolamento che raccolga i risultati della consultazione pubblica sulle opzioni di policy coerenti con il modello dei diritti fondamentali e dei valori europei⁴.

La chiara scelta del legislatore europeo di affiancare agli strumenti regolatori classici di fonte europea primaria e secondaria nonché ai trattati internazionali sui diritti umani una varietà di strumenti di *soft law*⁵, aventi la finalità di riportare l'etica al centro dell'analisi dell'impatto dell'intelligenza artificiale sulla società, sulla psicologia individuale, sul sistema giuridico, sulla finanza, sull'ambiente e sulla fiducia nella tecnologia⁶, evidenzia una scelta segnatamente volta alla realizzazione di un'ampia condivisione istituzionale del quadro di riferimento dell'Unione Europea nel contesto dello sviluppo tecnologico (§ 35)⁷.

2. White Paper on Artificial Intelligence. A European Approach to Excellence and Trust, https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

3. Guidelines for military and non-military use of Artificial Intelligence, <https://www.europarl.europa.eu/news/en/press-room/20210114IPR95627/guidelines-for-military-and-non-military-use-of-artificial-intelligence>.

4. Durante la fase di pubblicazione dello scritto è intervenuta la proposta di regolamento del Parlamento e del Consiglio sull'approccio europeo all'AI cui fa seguito la proposta di regolamento della Commissione, dei quali il presente contributo non ha potuto tener conto.

5. Nel contesto dei modelli regolatori globali si vedano Yeung, Lodge, 2019; Bignami, 2018.

6. Si annoverano almeno tre studi del Parlamento Europeo su etica e intelligenza artificiale: European framework on ethical aspects of artificial intelligence, robotics and related technologies, [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/654179/EPRS_STU\(2020\)654179_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/654179/EPRS_STU(2020)654179_EN.pdf); The ethics of artificial intelligence: Issues and initiatives, [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU\(2020\)634452_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf); Artificial Intelligence: From Ethics to Policy, [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641507/EPRS_STU\(2020\)641507_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641507/EPRS_STU(2020)641507_EN.pdf).

7. "L'Europa deve definire quale visione normativa intende realizzare per un futuro immerso nell'IA e di conseguenza deve capire quale nozione di IA dovrebbe essere studiata, sviluppata, distribuita e utilizzata in Europa per concretizzare tale visione. Con il presente documento, intendiamo contribuire a tale sforzo introducendo la nozione di IA affidabile, che riteniamo sia il modo giusto per costruire un futuro con l'IA. Un futuro in cui la democrazia, lo Stato di diritto e i diritti fondamentali sono alla base dei sistemi di IA e in cui tali sistemi migliorano e difendono costantemente la cultura democratica consentirà anche di creare un ambiente in cui l'innovazione e la competitività responsabile possono prosperare", <https://op.europa.eu/it/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>.

Il contributo tenta di esplorare le ragioni e le implicazioni di un approccio etico all'intelligenza artificiale che nutre l'aspirazione di essere “affidabile” (*trustworthy*) per poter garantire una “competitività responsabile” (§ 17). In particolare, ci si propone di riflettere sull'adeguatezza dell'approccio etico riguardo alla questione delle inferenze algoritmiche sulle persone e sulla società.

Per dovizia di precisione, dalla presente analisi rimane esclusa la riflessione relativa alla descrizione dei sistemi di intelligenza artificiale rilevanti nel campo giuridico e dell'impatto dell'intelligenza artificiale sul diritto, rinviandosi sul punto ad altri scritti⁸. Tuttavia, l'approccio etico nei sistemi di intelligenza artificiale, così come delineato nel contesto europeo, ne ripercorre alcune tappe che saranno oggetto di analisi.

2. ETICA E DIRITTI FONDAMENTALI

L'approccio etico all'intelligenza artificiale (AI) si basa programmaticamente sul rispetto dei diritti fondamentali, il cui nucleo comune è rappresentato dal concetto della dignità umana, sulla quale poggia l'architettura antropocentrica europea (§ 38)⁹. Più in generale, se questo è un paradigma sempre più ricorrente in diversi segmenti regolatori e in svariati ambiti applicativi, dalla genetica nutrizionale alla finanza sostenibile e al cambiamento climatico, nel contesto degli orientamenti etici la compenetrazione delle due componenti dell'AI, legalità ed eticità¹⁰, viene giustificata alla luce della pervasività dei sistemi di intelligenza artificiale nella vita quotidiana.

La motivazione di tale scelta istituzionale è data dalla trasposizione nel campo dell'AI di una dicotomia tradizionale¹¹, quella cioè tra diritti fonda-

8. Cfr. Catanzariti, 2020a, 239-255; Catanzariti, 2020b, 149-165.

9. § 38: “Il fondamento che accomuna questi diritti può essere inteso come radicato nel rispetto della dignità umana, riflettendo così quello che definiamo un “approccio antropocentrico” in cui l'essere umano gode di uno status morale unico e inalienabile di primato in campo civile, politico, economico e sociale”, <https://op.europa.eu/it/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>. Cfr. sul punto Cath *et al.*, 2018, 525: “This approach to human dignity provides the much needed grounding in a well-established, ethical, legal, political, and social concept, which can help to ensure that tolerant care and fostering respect for people (both as individuals and as groups), their cultures and their environments, play a steering role in the assessments and planning for the future of an AI-driven world. By relying on human dignity as the pivotal concept, it should become less difficult to develop a comprehensive vision of how responsibility, cooperation, and sharable values can guide the design a ‘good AI society’”. Per una trattazione puntuale sul concetto giuridico di dignità, cfr. McCrudden, 2013, 1-58; Mazzoni, 2019, 55-60; Viola, 2013, 283.

10. In tal senso cfr. Surden, 2020, 719.

11. Cfr. Pariotti, 2013, 31: “Si intende per “positivizzazione” il processo tramite il quale i diritti si trasformano, da ideali morali, in situazioni soggettive come tali dotate di copertura normativa entro gli ordinamenti giuridici”.

mentali positivizzati, sui quali si fonda la legalità dell'intelligenza artificiale, e diritti umani inerenti all'individuo, sui quali si fonda l'eticità dell'intelligenza artificiale (§ 39). Inoltre, il documento fa riferimento nello specifico al rispetto dei diritti fondamentali nel quadro della democrazia e dello Stato di diritto, affermando che *“Tali diritti sono descritti nella Carta dei diritti fondamentali dell'Unione europea con riferimento alla dignità, alle libertà, all'uguaglianza, alla solidarietà, ai diritti dei cittadini e alla giustizia. Il fondamento che accomuna questi diritti può essere inteso come radicato nel rispetto della dignità umana, riflettendo così quello che definiamo un 'approccio antropocentrico' in cui l'essere umano gode di uno status morale unico e inalienabile di primato in campo civile, politico, economico e sociale”* (§ 38).

Il paragrafo, dal generico e ambiguo titolo *“Dai diritti fondamentali ai principi etici”* (§§ 40-45), include cinque parametri di riferimento: la dignità umana, in base alla quale le persone devono essere trattate come soggetti morali e non come oggetti da classificare e manipolare (§ 41); la libertà individuale in virtù della quale gli esseri umani devono essere liberi di prendere decisioni su se stessi e non essere oggetto di manipolazione applicandosi ciò in senso ampio alla libertà di impresa, alla libertà delle arti, alla libertà di espressione, di riunione e di associazione, al diritto al rispetto della vita privata e della riservatezza (§ 42); la democrazia, giustizia e stato di diritto, intesa come triade, che legittima i poteri autorizzati da una previsione di legge e costituisce la cornice all'interno della quale i sistemi di AI devono contribuire ai processi democratici nonché a garantirne le basi, come ad esempio il giusto processo (§ 43); l'uguaglianza, il principio di non discriminazione e di solidarietà (compresi i diritti delle persone a rischio di esclusione), che si basa sul “pari rispetto per il valore morale e la dignità di tutti gli esseri umani”, tale per cui l'AI non può generare discriminazione né “risultati ingiustamente distorti” (ad esempio, i dati utilizzati per istruire i sistemi di IA dovrebbero essere il più inclusivi possibile e rappresentare gruppi di popolazione diversi); infine, quello relativo ai diritti dei cittadini, il cui accesso ai beni e ai servizi deve essere promosso dai sistemi di AI (§ 45).

Di fatto, non sembra esservi sostanziale differenza nei contenuti degli orientamenti etici e in quello dei valori fondamentali dell'UE sanciti dall'art. 2 TUE come quello della dignità umana, della libertà, della democrazia, dell'uguaglianza, dello Stato di diritto e del rispetto dei diritti umani, compresi i diritti delle persone appartenenti a minoranze. Ciò che suscita dubbi interpretativi è piuttosto il fatto che mentre l'inosservanza di una norma etica non comporta sanzione, il rischio ovvero l'esistenza di una grave violazione dei valori fondamentali dell'UE comporta l'attivazione della procedura di cui all'art. 7 TUE. Tuttavia, come sottolineato da Sheppele, i due livelli relativi

all’art. 2 e all’art. 7 TUE (replicato dall’art. 49 TUE come requisito per l’adesione all’UE) racchiudono in essi un meccanismo politico, non giuridico, poiché è necessaria la volontà politica unanime degli Stati membri per poter sanzionare una violazione e soltanto sulla base di una decisione assunta caso per caso. Il principio della mutua fiducia sul quale si basa l’Unione Europea plasma il contenuto intrinseco dello stato di diritto¹².

È evidente, dunque, che l’introduzione di un meccanismo valoriale sfuma sensibilmente l’idea stessa di *enforcement*, in quanto dipende dal contesto culturale di riferimento e può ben avere premesse anche diverse¹³.

Nel caso dell’intelligenza artificiale, le norme etiche “*non sono necessariamente vincolanti dal punto di vista giuridico ma cruciali per garantirne l’affidabilità*” (§ 39). Ebbene, qual è la ragione che ha spinto un gruppo di esperti a definire con il termine di ‘orientamento etico’ e non di ‘valore’ ciò che nella sostanza può essere considerata la quintessenza dei diritti fondamentali?

La scelta compromissoria dell’AI HLEG desta anzitutto qualche perplessità epistemologica in quanto essa sembrerebbe lasciar intendere che le basi dell’AI poggino sui diritti fondamentali in un doppio senso: da un lato, come strumenti vincolanti, i diritti fondamentali ineriscono alla legalità, dall’altro, essendo valori intrinseci dell’umanità, sono sussumibili nel paradigma etico (§ 39).

La confusione deriva dal fatto che si tenterebbe così di innestare la teoria integrata dei diritti fondamentali, che non ne riduce la portata alla mera dimensione morale o giuridica, su una premessa contraddittoria, e cioè che l’etica si basi sui diritti.

Affermare che l’etica poggia sul diritto vuol dire invertire l’ordine del processo di legittimità di un imperativo, in quanto ciò implica in altri termini che è il diritto a legittimare ed autorizzare l’etica.

La visione integrata, che riconosce la presenza della dimensione morale e giuridica nei diritti umani considerandoli quali “*pretese morali particolarmente pregnanti, sostenute ed espresse da norme giuridiche*”, si basa invero su una premessa del tutto diversa, che esclude la possibilità di fondare l’etica sul diritto¹⁴.

12. Sheppele, 2018, <https://verfassungsblog.de/rule-of-law-retail-and-rule-of-law-wholesale-the-ecjs-alarming-celmer-decision/>.

13. Non si può in questa sede dilungarsi in un’analisi approfondita, ma basta forse ricordare che Rawls nella edizione del 1999 di *A Theory of Justice* aveva rivisitato il concetto di *overlapping consensus* in maniera problematica, ponendo il problema della possibile ricaduta nella sfera etica (Rawls, 1999, 340 ss.).

14. Molto convincente sul punto è la tesi di Pariotti, 2013, 204: “riconoscere l’importanza dei diritti umani per un recupero dell’universalità nell’etica contemporanea non significa sostenere che l’etica debba fondarsi sui diritti. Tutt’al contrario, è un’ulteriore conferma che i diritti stessi necessitano di poggiare su valori morali o, meglio, etico-politici”. E ancora: “i diritti sono una

A meno di non dover ritenere che l'utilizzo dell'etica in siffatto contesto non sia altro che una strategia comunicativa tra gli Stati Membri che funge da facilitatore di un processo istituzionale a venire¹⁵, per il giurista risulta problematico accettare che oltre al *corpus* normativo sia necessaria l'etica per creare fiducia nell'intelligenza artificiale¹⁶, e che essa si basi sui diritti fondamentali.

Scendendo nel merito, gli orientamenti etici contengono quattro principi etici e sette requisiti di affidabilità dell'AI, ma non si evince chiaramente quale sia il rapporto tra di essi.

Tra i primi sono inclusi: 1) il principio del rispetto dell'autonomia umana, che si fonda sull'autodeterminazione tale per cui i sistemi AI non dovrebbero manipolare gli esseri umani (§ 50); 2) il principio della prevenzione dei danni, che si fonda sulla robustezza dei sistemi che dovrebbe scongiurare la causazione di danni nei confronti degli esseri umani, incluso il rispetto dell'ambiente (§ 51); il principio di equità, in base al quale lo sviluppo, la distribuzione e l'utilizzo dei sistemi AI non soltanto non debba produrre distorsioni inique ma debba altresì accrescere le possibilità di accesso ai beni e ai servizi (§ 52); il principio di esplicabilità, sul quale si fonda l'impalcatura della fiducia nei sistemi di AI, e che significa trasparenza e accesso ai processi decisionali algoritmici in ogni fase (§ 53).

Accanto ai principi, vi sono poi i sette requisiti necessari per la realizzazione di una intelligenza artificiale affidabile: 1. intervento e sorveglianza umani, volti ad evitare un impatto negativo sui diritti fondamentali, preservare l'autonomia decisionale umana in merito ai sistemi di AI potendone modulare l'intensità a seconda che sia previsto l'intervento umano in ogni ciclo, la supervisione umana durante la progettazione e il monitoraggio del funzionamento o

nozione *poggiante* su (e non coincidente con) valori. Ragione, questa, per cui sarebbe da escludere la possibilità di un'etica *basata* sui diritti: se si ammette che i diritti, per la determinazione del loro contenuto e per la loro giustificazione, rinvino ad altri elementi, allora è esclusa la possibilità di vedere nei diritti dei concetti originari”.

15. In tal senso cfr. Artificial Intelligence. From Ethics to Policy, [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641507/EPRS_STU\(2020\)641507_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641507/EPRS_STU(2020)641507_EN.pdf), p. 24, come pure Yeung, Howes, Pogrebna, 2020, 81-82.

16. Orientamenti etici, par. 36 e 46. Nel senso del carattere suppletivo dell'etica, cfr. Mantelero, 2018, 29: “the effect of the social and ethical values on the interpretation of these human rights. These values represent the societal factors that influence the way the balance is achieved between the different human rights and freedoms, in different contexts and in different periods. Moreover, social and ethical values concur in defining the extension of rights and freedoms, making possible broader forms of protection when the regulatory framework does not provide adequate answers to emerging issues”. In linea cfr. anche White Paper, p. 36, specificamente sulla mancata copertura giuridica di concetti come trasparenza, tracciabilità e intervento umano, https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en.

il controllo umano nel complesso (§ 62-65); 2. robustezza tecnica e sicurezza, realizzati mediante un approccio di prevenzione dei rischi volto ad evitare che si possano determinare danni non intenzionali e a far sì che i sistemi operino in base alle previsioni e siano resilienti contro gli attacchi esterni, precisi, affidabili e riproducibili (§ 66-70); 3. riservatezza e *governance* dei dati, che richiedono la riservatezza e protezione dei dati durante l'intero processo di raccolta e utilizzo, l'alta qualità e integrità dei dati al fine di ridurre il margine di imprecisione ed errore dei sistemi AI, e anche la regolamentazione dell'accesso ai dati personali (§ 71-74); 4. la trasparenza che si declina in tracciabilità dei data set e dei processi, spiegabilità dei processi tecnici e decisionali e identificazione degli stessi quali sistemi non umani con possibilità di preferire sempre ad essi l'interazione umana (§ 75-78); 5. diversità, non discriminazione ed equità, che pongono quali premesse del ciclo iniziale dei sistemi di AI l'inclusione e il pluralismo al fine di evitare distorsioni inique nonché garantire l'accesso al più ampio spettro di soggetti (§ 79-82); 6. benessere sociale e ambientale, che mirano all'utilizzazione dell'AI a vantaggio delle generazioni future in un contesto di sostenibilità e responsabilità ecologiche e a far sì che non sia alterata la percezione umana di intervento sociale, relazioni sociali ed affettive, valutando anche l'impatto sociale sui processi istituzionali e politici e sul diritto di voto (§§ 83-86); 7. *accountability*¹⁷, esemplificata dalla verificabilità degli algoritmi, dei dati utilizzati e dei processi di progettazione, dalla riduzione al minimo degli effetti negativi potenziali e dalla relativa segnalazione, dall'individuazione di compromessi eticamente accettabili tra soluzioni in conflitto e infine dalla possibilità di ricorso avverso violazioni che determinano effetti negativi ingiustificati (§§ 87-91).

Tra i rischi potenziali segnalati dall'AI HLEG merita particolare attenzione il rischio della valutazione dell'integrità etica dei cittadini per punteggio in violazione dei diritti fondamentali (§§ 132-133) in netta antitesi con l'opzione culturale della iniziativa governativa cinese relativa al Social Credit System, che prevede lo *scoring* etico dei cittadini basato sull'affidabilità delle abitudini private e pubbliche e che investe la totalità delle dimensioni della vita quotidiana e le azioni non soltanto legalmente illecite ma anche moralmente e professionalmente considerate non meritevoli: dalla storia finanziaria alla capacità contrattuale e alle abitudini quotidiane, dalle relazioni sociali alle qualità personali¹⁸.

Tale rischio suscita una preoccupazione allarmante, al di là dell'analisi nel merito del progetto cinese, se si considera che il credito sociale rappresenta un

17. Si preferisce non tradurre il termine *accountability* in quanto in italiano non esiste un termine equivalente, che sarebbe inesatto rendere con responsabilità.

18. Orgad, Reijers, 2020, 2-3, https://cadmus.eui.eu/bitstream/handle/1814/66910/RSCAS%202020_28.pdf.

esempio di etica fondato sul diritto. Esso poggia su una nozione di virtù civica nella quale il collante della comunità politica è rappresentato dalla conformità alle norme sociali a scapito della possibilità di manifestare il dissenso¹⁹.

3. ARTIFICIAL ETHICS O ARTIFICIAL REASON?

A ben guardare, se da un lato può essere condivisibile che l'approccio etico all'intelligenza artificiale riponga le proprie basi nella cultura dei diritti fondamentali, rinvenire la costruzione delle basi fondanti dell'etica nei diritti non ne fa comprendere da un lato la necessità, poiché se ne duplica il contenuto, dall'altro la finalità, poiché potenzialmente in contrasto con il principio del laicismo della società democratica²⁰. Del resto, addirittura dal pensiero giusfilosofico di matrice cattolica proviene la critica alla commistione tra etica e laicità, nella nota formula del paradosso di Ernst-Wolfgang Böckenförde risalente al celebre saggio del 1967 *La formazione dello stato come processo di secolarizzazione*: "lo Stato liberale secolarizzato vive di presupposti che non può garantire: da una parte esso può esistere solo se la libertà si regola dall'interno, cioè a partire dalla sostanza morale dell'individuo; dall'altra però se lo Stato cerca di garantire da sé queste forze regolatrici interne esso rinuncia alla propria liberalità ("carattere liberale"²¹, *N.d.A.*) e ricade – su un piano secolarizzato – in quella stessa istanza di totalità dalle quali si era tolto con le guerre civili confessionali"²². È dunque innegabile che proprio nel contesto dell'innovazione un approccio etico evochi l'idea di uno stato confessionale, ponendo un problema di accettazione generalizzata di alcuni principi e valori che sono fondati su prescrizioni normative. Né si comprende la ragione del dover ribadire dal punto di vista etico la conformità ai diritti fondamentali già positivizzati. La giustificazione ontologica fornita dagli esperti, relativa cioè alla

19. Síthigh, Siems, 2019, 17, https://cadmus.eui.eu/bitstream/handle/1814/60424/LAW_2019_01.pdf.

20. Al par. 40 le Ethics Guidelines affermano quanto segue: "Nell'esauriente corpus di diritti indivisibili previsti dal diritto internazionale in materia di diritti umani, dai trattati UE e dalla Carta dei diritti fondamentali dell'Unione europea, le seguenti famiglie di diritti fondamentali sono particolarmente pertinenti per quanto riguarda i sistemi di IA. Molti di questi diritti, in determinate circostanze, sono giuridicamente applicabili nell'UE, pertanto il rispetto dei loro termini è giuridicamente vincolante. Ma anche una volta rispettati i diritti fondamentali giuridicamente applicabili, la riflessione etica può aiutarci a comprendere come lo sviluppo, la distribuzione e l'utilizzo dell'IA possano implicare i diritti fondamentali e i valori soggiacenti, e può contribuire a fornire un orientamento più dettagliato nel percorso di ricerca di ciò che *dobbiamo* fare piuttosto che ciò che (attualmente) *possiamo* fare con la tecnologia".

21. Si è preferito ipotizzare una traduzione libera, non concordando con la traduzione del termine originario tedesco di *Freiheitlichkeit* con la parola "liberalità" che si trova nella versione italiana dell'opera.

22. Böckenförde, 1967, 68-69.

natura etica dei diritti fondamentali appare francamente debole, rivelando invece un problema nella applicabilità e nell’efficacia degli stessi. Ora, se tale debolezza derivi dalla impossibilità dell’*enforcement* ovvero da una necessità di rafforzamento del *framework* giuridico esistente, che possa rappresentare un vademecum di affidabilità per gli operatori del settore, non è del tutto chiaro. Tuttavia, la non vincolatività degli orientamenti etici è percepita come lacuna da colmare mediante un’operazione di duplicazione dei contenuti giuridici e di ancoraggio della normativa secondaria europea ai diritti fondamentali. Sotto il primo profilo, l’operazione appare anche superflua nel merito; sotto il secondo profilo, è anche concettualmente sbagliata, in quanto manca in tal senso una norma di legge autorizzativa²³.

Appare tuttavia probabilmente più convincente andare oltre questa lettura e tentare un tipo di interpretazione della policy europea che consideri la scelta dell’approccio etico in altra ottica, e cioè come un fenomeno di importazione culturale di un modello di matrice anglosassone che in un contesto culturale diverso da quello europeo significa altra cosa. Ciò che appare interessante e plausibile perché in linea con la tendenza extraeuropea nei confronti della innovazione da un lato e della sostenibilità digitale dall’altro.

Se per un verso l’etica può sopperire alla *self-regulation* di stampo americano che spesso comporta la derogabilità dei diritti, per altro verso la propagazione di una portata globale del modello europeo, una sorta di *Brussels Effect* di stampo etico²⁴, a livello formale e giuridico può forse essere più efficace?

Nel contesto regolatorio nel quale l’Europa è stata spesso definita come “regulatory standard-setter”²⁵, la conformità ai principi etici mediante meccanismi di autocertificazione serve a scampare quello che oltre Oceano è solitamente demandato alla *self-regulation* e che in Europa susciterebbe resistenze.

Come evidenziato da Chiti e Marchetti, la tematizzazione dell’avvento dell’intelligenza artificiale come una “crisi” spiega anche il tipo di interventi messi a punto dall’Unione Europea: da un lato i finanziamenti; dall’altro gli strumenti normativi; infine, gli orientamenti etici. Ebbene, la finalità ravvisata dagli Autori nella combinazione di questi interventi è quella di ammorbidire la durezza dell’intervento regolatorio “per stabilizzare i comportamenti dei privati e [...] promuovere l’innovazione”²⁶. Per quanto riguarda gli orientamenti etici, in particolare, non sarebbe chiaro se debba trattarsi di un’etica

23. Principi come *accountability* e *transparency* sono ad esempio contenuti nel Regolamento Generale della Protezione dei Dati Personali.

24. Bradford, 2019, 7, 142, 147.

25. Cremona, Scott, 2019, 11; Bradford, 2015, 158; Ebers, Cantero Gamito, 2021, 8. Emblematico sotto questo profilo lo scopo territoriale estremamente ampio della proposta di regolamento sull’AI (art. 2), <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence>.

26. Chiti, Marchetti, 2020, 39.

sociale, cui possa conformarsi la comunità nel suo complesso ovvero di un'etica individuale²⁷, e quale sia il rapporto con il diritto. Questi Autori definiscono gli orientamenti etici come “norme giuridiche non vincolanti che hanno un contenuto apertamente morale”²⁸ in quanto non sempre è facile distinguere se essi rispondono a questioni relative ai diritti e ai doveri ovvero ad imperativi morali²⁹.

Ebbene, prescindendo un momento dalla disinvoltura con cui viene utilizzato il termine “etica” si può però comprendere meglio a che cosa si riferisce il gruppo di esperti: a un *corpus* di principi che ha un'ambizione regolatoria senza averne uno *standing* e tuttavia pretende legittimità³⁰. Etica è dunque tutto ciò che va *beyond the law*, e non esattamente ciò che si basa sui diritti. Per questo necessita di *moral suasion*, attraverso tecniche composite ed eterogenee: a volte replica il contenuto dei diritti, per cui non si riscontrano differenze strutturali né di contenuto tra obblighi giuridici ed etici; altre volte amplia il raggio della portata applicativa dei diritti (si pensi ad esempio all'applicabilità della Carta dei Diritti nelle sole materie del diritto dell'Unione), altre volte pone sullo stesso piano obblighi etici e obblighi giuridici.

In sostanza, i principi etici servirebbero a scongiurare l'uso della *self-regulation* da parte delle imprese che in ambiente europeo verrebbe visto con sospetto, in quanto produrrebbe una delega in bianco da parte della regolazione nei confronti dell'autonomia negoziale privata³¹. Concetti come impatto sociale e sostenibilità sono difficilmente vincolanti nel breve periodo e quando non hanno copertura giuridica non è neanche necessaria una deroga espressa. Per questo l'etica può essere meglio intesa nella pratica come una sorta di sistema di autocertificazione *soft* che funge da biglietto da visita di affidabilità: ciò anche in considerazione del fatto che la presenza di un *ethical committee* all'interno di ogni singola azienda sarebbe discrezionale e dispendioso, mentre la *checklist* di conformità ai sette requisiti dell'intelligenza artificiale affidabile rappresenta certamente un meccanismo più accessibile.

4. LA MORALITÀ DELLE INFERENZE

Sotto altro profilo, appare singolare la pretesa etica della regolazione del funzionamento delle macchine. Come a dire che le macchine possano essere morali³². Rispetto alla domanda con la quale si apriva il famoso saggio di Alan

27. Ivi, pp. 40 e 41.

28. Ivi, p. 48.

29. La critica di fondo consiste nel rifiuto della fonte etica di produzione del diritto.

30. Waldman, 2020, 107.

31. Per un'analisi teorica del rapporto tra *self-regulation* e *soft law*, cfr. Pariotti, 2017, 9-28.

32. Renda, 2021, 667. Per un'attenta disamina del pensiero filosofico settecentesco sulle macchine cfr. Punzi, 2003, 1-428.

Turing del 1950³³, se cioè le macchine possano pensare, l’etica si pone di molto al di fuori dell’*imitation game*, perché la verifica non è necessariamente contestuale. Dal punto di vista dell’analisi filosofico-giuridica i principi etici possono valere come giustificazione *ex post* di un intervento già posto in atto, in maniera duplice: tentando di coniugare due logiche diverse, quella del diritto e quella dell’etica, e di renderle il più possibile omogenee nel loro carattere cogente, oppure demandando a un campo *non enforceable* come quello dell’etica uno spazio di libertà che lasci gli attori e in particolare le imprese libere di operare.

Nel primo caso si pone il problema di valutare in quale misura etica e diritto siano funzionalmente complementari e cosa significhi nella pratica questa complementarità: o forse, in sostanza, quale delle due logiche prevalga a conti fatti³⁴. Nel secondo caso la delega nei confronti dell’etica è già di per sé foriera di una scelta valoriale di fondo che imposta in qualche modo l’ordine delle preferenze partendo dal livello valoriale ma non vincolante, e giocando sull’ambiguità.

Il nodo da sciogliere, per il quale non appare adeguato né l’intervento del diritto come *legal design* dell’algoritmo né tantomeno l’approccio etico appiattito sui diritti fondamentali, è quello relativo alle inferenze algoritmiche che, creando connessioni logiche tra i dati immessi nei sistemi, specie in quelli di *machine learning*, agiscono sul piano cognitivo della rappresentazione della realtà³⁵.

Da un lato, le modifiche del processo cognitivo che deriva dalla interazione di causalità e correlazione sono autonome rispetto al piano degli effetti. In altri termini, quando gli algoritmi incidono sulla vita degli individui, diventano processi egemonici e non è pensabile che essi siano neutrali per il solo fatto di essere il frutto di aggregazione di dati grezzi. Non bisogna però illudersi che spostandosi sul piano etico dell’AI il determinismo tecnologico cessi di produrre forme di potere e di soggezione. L’etica non può neutralizzare l’idea di *affordance*, intesa come limite o opportunità per l’azione sociale all’interno della tecnologia³⁶. Sia l’AI sia la produzione di *training data* sono metodi estrattivi e non relazionali (Zuboff, 2019, 128), che assorbono dalla materialità delle risorse ambientali consumabili e degli individui in carne e ossa, il che determina certamente un impatto etico in termini di costi e benefici³⁷.

33. Turing, 1950, 433.

34. Sul punto cfr. Ebers, 2020, 92.

35. Per un’analisi giuridica delle inferenze, cfr. Watcher, Mittelstadt, 2019, 494-620.

36. Graber, 2020, 197.

37. Sui presupposti della sostenibilità digitale, cfr. Stuermer, Abu-Tayeh, Myrach, 2017, 247-262.

Dall'altro, la formazione dell'autonomia individuale contribuisce a un personale convincimento di giustizia³⁸ e l'etica non può avere un significato univoco e valevole *erga omnes* al pari di un obbligo giuridico. Come giustamente osservato da Floridi, le Information and Communication Technologies (ICT's) "hanno re-ontologizzato i sistemi"³⁹: blindare programmaticamente l'AI con il *passe partout* dei diritti fondamentali sia come base della legalità dell'AI sia come base dell'eticità dell'AI non risolve la complessità delle questioni sottese.

Proprio in quest'ottica, appare interessante l'analisi dei test etici condotta da alcuni studiosi dell'Oxford Internet Institute⁴⁰. Le regole del test etico applicabile all'AI secondo questi Autori sono definite dalle stesse norme e consuetudini che governano l'ambito cognitivo nel quale agisce l'artefatto dell'AI. Lo studio muove da tre casi nei quali l'uso degli algoritmi può determinare la possibilità di errore nell'interpretazione della realtà. Tali sono: la trasformazione di dati in prove; la giustificazione di un'azione che non è eticamente neutrale; la divisione di responsabilità. Questi tre casi possono a loro volta dar luogo a questioni etiche: una prova non fondata può dar luogo ad azioni ingiustificate; una prova non accessibile può determinare opacità; una prova fuorviante (non è vero che gli algoritmi sono privi di *bias*, perché riflettono i valori degli sviluppatori, valori sociali preesistenti, limitazioni tecniche e dati di contesto) può dar luogo a pregiudizi; un risultato non trasparente può causare effetti discriminatori; alcuni effetti trasformativi sugli individui possono rappresentare sfide all'autonomia individuale; infine la tracciabilità rappresenta il fattore determinante della responsabilità morale⁴¹.

Tale filone di ricerca, che si colloca all'interno della cosiddetta *Information Ethics*, analizza in maniera profonda le implicazioni etiche della tecnologia dell'informazione, dalla creazione dell'informazione al suo utilizzo. Nell'accezione proposta da Floridi, si nota una profonda distanza della speculazione filosofica sull'etica rispetto agli orientamenti etici, in quanto essa muove dal presupposto che le informazioni abbiano un valore etico intrinseco che può essere universalmente accettabile nella misura in cui non sia sottoposto a coercizione⁴² e che scaturisce dalla condivisione contestuale dello spazio informativo (definito "infosfera") tra agenti umani e agenti artificiali dotati entrambi un *relational self*⁴³ tale per cui non esisterebbe una separazione tra

38. Stevens, 2020, 156, 168.

39. Floridi, 2010, 4.

40. Mittelstadt, Allo, Taddeo, Wachter, Floridi, 2016, 4.

41. Ivi, pp. 4-10.

42. Floridi, 2013, 322.

43. Floridi, 2015, 11.

realtà e spazio virtuale. Nel contesto della presente analisi, tale prospettiva appare assai più interessante, quantomeno dal punto di vista speculativo, rispetto a un modello di duplicazione dei diritti in base al quale gli orientamenti etici sono stati concepiti, perché essa non contempla una duplicazione tra spazio reale e artificiale e rivendica l'autonomia dell'etica dell'informazione come branca autonoma della filosofia dell'informazione.

5. CONCLUSIONI

Alcuni ritengono che sia posta eccessiva enfasi sull'impatto dell'intelligenza artificiale sul diritto, mentre poco si rifletterebbe sul fatto che vi siano diversi altri fattori – tra i quali la politica, le possibilità finanziarie, l'assetto istituzionale – che contribuiscono a determinare come i valori giuridici fondanti di un ordinamento siano integrati al suo interno⁴⁴. Tuttavia, è inevitabile riflettere sul ruolo del diritto in presenza di norme etiche che ricalcano quelle giuridiche. Ebbene, dinanzi a un goffo tentativo di un'etica che replica i precetti giuridici, sembra sempre più cogente il bisogno di un diritto che sia all'altezza dei saperi specializzati e sia in grado di accogliere le sfide dei tempi elevandosi culturalmente come forma di sapere inclusivo.

Probabilmente, il carattere frammentario e settoriale delle normative in materia di intelligenza artificiale ha creato un equivoco di fondo: che il ruolo del diritto, cioè, potesse essere relegato al *legal design* dell'algoritmo modellato a seconda della disciplina applicabile, e per questa ragione il legislatore europeo, probabilmente deluso dallo scarto cognitivo del “poter fare tutto quello che si può fare” con la tecnologia, ha riposto nell'approccio etico la speranza di creare il migliore dei mondi possibili, tanto più se conforme ai diritti fondamentali. Nella pratica, come esposto, ciò ha dato luogo a due fenomeni che non devono essere sottovalutati: da un lato, l'espressa domanda di conformità dell'etica al diritto; dall'altro, il livellamento tra principi etici e diritti, che da un lato indebolisce il senso di conformità giuridica, tradendone la esatta delimitazione dei presupposti, e dall'altro offusca la portata e le implicazioni dei precetti etici.

Dunque, etica o diritto, in fin dei conti, per l'intelligenza artificiale? Se l'esperienza giuridica veicola meccanismi di elaborazione riflessiva del sapere, l'etica ne sembra una declinazione un po' inusuale e poco produttiva. Se si ritiene invece che il diritto della quarta rivoluzione industriale abbia un suo nuovo lessico, allora è necessario fare i conti con il fatto che le metafore linguistiche possano mutare, e con esse anche la consapevolezza che induce a definire i problemi.

44. Cfr. Surden, 2020, 721.

RIFERIMENTI BIBLIOGRAFICI

- Bignami, F. (2018). *Comparative Law and Regulation. Understanding the Global Regulatory Process*. Elgar.
- Böckenforde, W. (1967). *La formazione dello Stato come processo di secolarizzazione* (a cura di M. Nicoletti). Morcelliana.
- Bradford, A. (2015). Exporting Standards: The externalization of the EU's Regulatory Power Via Markets. *International Law Review of Law and Economics*, 42, 158-173.
- Bradford, A. (2019). *The Brussels Effect. How the European Union Rules the World*. Oxford University Press.
- Catanzariti, M. (2020a). Enhancing Policing Through Algorithmic Mass-Surveillance. In L. Marin, S. Montaldo (eds.), *The Fight Against Impunity* (pp. 239-255). Hart.
- Catanzariti, M. (2020b). La razionalità algoritmica dei processi decisionali. In S. Gozzo, C. Pennisi, V. Asero, R. Sampugnaro (eds.), *Big Data e processi decisionali* (pp. 149-165). Egea.
- Cath, C., Floridi, L., Mittelstadt, B., Taddeo, M., Watcher, S. (2018). Artificial Intelligence and the 'Good Society': the US, EU, and UK approach. *Science and Engineering Ethics*, 24, 505-528.
- Chiti, E., Marchetti, B. (2020). Divergenti? Le strategie di Unione Europea e Stati Uniti in materia di intelligenza artificiale. *Rivista della Regolazione dei Mercati*, 1, 28-50.
- Cremona, M., Scott, J. (2019). Introduction. In M. Cremona, J. Scott (eds.), *EU Law Beyond EU Borders: The Extraterritorial Reach of EU Law* (pp. 1-20). Oxford University Press.
- Ebers, M. (2020). Regulating AI and Robotics: Ethical and Legal Challenges. In M. Ebers, S. Navas Navarro (eds.), *Algorithms and Law* (pp. 37-99). Cambridge University Press.
- Ebers, M., Cantero Gamito, M. (2021). Algorithmic Governance and Governance of Algorithms: An Introduction. In M. Ebers, M. Cantero Gamito (eds.), *Algorithmic Governance and Governance of Algorithms: Legal and Ethical Challenges* (pp. 1-22). Springer.
- Floridi, L. (2010). Ethics after the Information Revolution. In *The Cambridge Handbook of Information and Computer Ethics* (pp. 3-19). Cambridge University Press.
- Floridi, L. (2013). *The Ethics of Information*. Oxford University Press.
- Floridi, L. (2015). *The Online Manifesto. Being Human in Hyperconnected Era*. Springer.
- Graber, C.B. (2020). Artificial Intelligence, Affordances and Fundamental Rights. In M. Hildebrandt, K. O'Hara (eds.), *Life and the Law in the Era of data-driven agency* (pp. 194-213). Elgar.
- Hildebrandt, M., O'Hara, K., (2020). Introduction: Life and the Law in the Era of Data-Driven Agency. In M. Hildebrandt, K. O'Hara (eds.), *Life and the Law in the Era of Data-Driven Agency* (pp. 1-15). Elgar.
- Mantelero, A. (2018). AI and Big Data. A Blueprint for a Human Rights, Social and Ethical Impact Assessment. *Computer Law & Security Review*, 34(4), 754-772.
- Mazzoni, C.M. (2019). *Quale dignità. Il lungo viaggio di un'idea*. Leo S. Olschki.
- McCrudden, C. (2013). In Pursuit of Human Dignity: An Introduction to Current Debates. In C. McCrudden (ed.), *Understanding Human Dignity* (pp. 1-58). Oxford University Press.

- Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S., Floridi, L. (2016). The Ethics of Algorithms: Mapping the debate. *Big Data & Society*, 3(1), 1-21.
- Orgad, L., Reijers, W. (2020), How to Make the Perfect Citizen? Lessons from China's Model of Social Credit System. *EUI Working Paper RSCAS 28*, https://cadmus.eui.eu/bitstream/handle/1814/66910/RSCAS%202020_28.pdf.
- Pariotti, E. (2013). *Diritti umani: concetto, teoria, evoluzione*. Cedam.
- Pariotti, E. (2017), Self-regulation, concetto di diritto, normatività giuridica. *Ars Interpretandi*, 2, 9-28.
- Punzi, A. (2003). *L'ordine giuridico delle macchine. La Mettrie Helvetius d'Holbach. L'uomo macchina verso l'intelligenza collettiva*. Giappichelli.
- Rawls, J. (1999). *A Theory of Justice*. Harvard University Press.
- Renda, A. (2021). Moral Machines. The Emerging EU Policy on “Trustworthy AI”. In W. Barfield (ed.), *The Cambridge Handbook of Algorithms* (pp. 667-690), Cambridge University Press.
- Sithigh, D.M., Siems, M. (2019), The Chinese Social Credit System: A Model for Other Countries? *EUI Working Paper LAW 1*, https://cadmus.eui.eu/bitstream/handle/1814/60424/LAW_2019_01.pdf.
- Sheppele, K.L. (2018), Rule of Law Retail and Rule of Law Wholesail. *verfassunsblog.de*, July.
- Stevens, D. (2020), In Defense of ‘Toma’: Algorithmic Enhancement of a Sense of Justice. In M. Hildebrandt, K. O'Hara (eds.), *Life and the Law in the Era of Data-Driven Agency* (pp. 156-174). Elgar.
- Stuermer, M., Abu-Tayeh, G., Myrach, T. (2017). Digital Sustainability: Basic Conditions for Sustainable Digital Artifacts and Their Ecosystems. *Sustainability Science*, 2, 247-262.
- Surden, H. (2020). The Ethics of AI in Law: Basic Questions. In M.D. Dubber, F. Pasquale, S. Das (eds.), *Oxford Handbook of Ethics of AI* (pp. 720-736). Oxford University Press.
- Turing, A.M. (1950). Computing Machinery and Intelligence. *Mind*, 49, 433-460.
- Viola, F. (2013). Lo statuto normativo della dignità umana. In A. Abignente, F. Scamardella (a cura di), *Dignità della persona. Riconoscimento dei diritti nelle società multiculturali* (pp. 283-295). Editoriale Scientifica.
- Wachter, S., Mittelstadt, B. (2019). A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review*, 2, 494-620.
- Waldman, A.E. (2020), Algorithmic Legitimacy. In W. Barfield (ed.), *The Cambridge Handbook of Algorithms* (pp. 107-120). Cambridge University Press.
- Yeung, K., Howes, A., Pogrebna, G. (2020). AI Governance by Human Rights – Centered Design, Deliberation, and Oversight: An End to Ethics Washing. In M.D. Dubber, F. Pasquale, S. Das (eds.), *The Oxford Handbook of Ethics of AI* (pp. 78-106). Oxford University Press.
- Yeung, K., Lodge, M. (2019). *Algorithmic Regulation*. Oxford University Press.
- Zuboff, S. (2019). *The Age of Surveillance Capitalism. The Fight for a Human Future and the New Frontier of Power*. Public Affairs.

