i-lex. Scienze Giuridiche, Scienze Cognitive e Intelligenza Artificiale
Rivista semestrale on-line: www.i-lex.it

AI ACT: IMPATTI E PROPOSTE

Opportunità e rischi dell'over- e under-regulation

G.R. MARSEGLIA

i-lex

i-lex. Scienze Giuridiche, Scienze Cognitive e Intelligenza Artificiale
Rivista semestrale on-line: www.i-lex.it
Dicembre 2021
Fascicolo 2
ISSN 1825-1927

i-lex. Scienze Giuridiche, Scienze Cognitive e Intelligenza Artificiale
Rivista semestrale on-line: www.i-lex.it

AI ACT: IMPATTI E PROPOSTE
Opportunità e rischi dell'over- e under- regulation

G.R. MARSEGLIA[1]

Abstract. Il 21 aprile 2021 la Commissione Europea ha pubblicato una proposta di regolamento intitolata "*Regolamento del parlamento europeo e del consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (Artificial Intelligence Act) e modifica di alcuni atti legislativi dell'Unione*". L'obbiettivo di questa proposta è quello di garantire che i sistemi di Intelligenza Artificiale (AI) messi in produzione all'interno dell'Unione Europea tutelino il cittadino e dunque siano sicuri, affidabili e non ledano la dignità dell'uomo. Il policy making su una tecnologia che evolve così velocemente, però, rischia di essere insufficiente o troppo stringente rispetto alle esigenze tipiche della ricerca tecnica. In questo articolo ci proponiamo di evidenziarne questo trade-off e di proporre delle linee guida per superarlo.

Parole chiave: *intelligenza artificiale; algoretica; digital ethics; artificial intelligence act; Unione Europea.*

1. Introduzione

Tornando indietro nella storia e a partire dalla fine del diciottesimo secolo, l'economia mondiale è stata caratterizzata da dei momenti che rappresentano delle discontinuità dal punto di vista della capacità produttiva e della efficienza nella gestione dei processi: le rivoluzioni industriali[2]. Ogni qualvolta la disponibilità di nuova tecnologia ha messo le imprese nelle condizioni di promuovere innovazione e

[1] Università degli Studi di Pavia, roberto.marseglia@unipv.it

[2] Popkova, Elena G., Yulia V. Ragulina, and Aleksei V. Bogoviz. "*Fundamental differences of transition to industry 4.0 from previous industrial revolutions*." Industry 4.0: Industrial Revolution of the 21st Century. Springer, Cham, 2019. 21-29

progresso, però, ci sono stati grandi impatti e modificazioni nell'assetto sociale e, in risposta, movimenti organizzati per resistere in maniera strutturata a questi cambiamenti. A titolo esemplificativo, durante la diffusione della meccanizzazione in Inghilterra il Generale Ludd ha dato vita al movimento del *luddismo* con l'obbiettivo di proporre un sistematico sabotaggio del progresso tecnologico manomettendo fisicamente i macchinari. Più recentemente ha fatto particolare scalpore l'azione dello scienziato americano Theodore Kaczynski, altrimenti noto come *The Unabomber*, che ha perpetrato un'azione costante di terrorismo nei confronti di persone che lui considerava come simboli del progresso tecnologico; azione che nei suoi scritti condivisi anonimamente con la stampa considerava giustificata dalla lenta affermazione della *technological slavery* che voleva in questo modo scongiurare[3].

La quarta rivoluzione industriale o *Industry 4.0*, che vede protagonisti i dati e la connettività digitale, è ancora in corso e, rispetto al passato, propone sfide originali cui dobbiamo rispondere prontamente e con i giusti strumenti. In particolare, differentemente dalle precedenti rivoluzioni in cui la tecnologia evolveva lentamente durante le fasi di cambiamento, oggi si parla di una fase comunemente nota come *age of accelerating change*[4] o fase di *exponential growth of technology*[5] in cui in un tempo che varia dai tre ai sette anni circa i paradigmi tecnologici cambiano quasi completamente essendo quindi delle piccole rivoluzioni tecnologiche all'interno di una più grande rivoluzione industriale.

Questa frequenza nei cambiamenti tipica della *digital era* si riflette ovviamente anche in maniera più evidente negli impatti sociali e stimola dunque movimenti contrari al progresso e all'innovazione molto energici oggi noti come *digital luddism*[6].

---

[3] Kaczynski, Theodore John. "*Industrial society and its future.*" (1995).

[4] Taal, Amie, ed. *The GDPR Challenge: Privacy, Technology, and Compliance in an Age of Accelerating Change*. CRC Press, 2021.

[5] Cassard, Anita, and Joseph Hamel. "*Exponential Growth of Technology and the Impact on Economic Jobs and Teachings: Change by Assimilation*." Journal of Applied Business & Economics 20.2 (2018).

[6] Jones, Steven E. *"Against technology: From the Luddites to neo-Luddism.*" Routledge (2013).

In una situazione di contesto così dinamica, le forze che spingono verso un *mondo digitale* sono due: una che tende ad avvicinare e democraticizzare l'accesso a informazioni e servizi di qualità; l'altra che invece tende a creare divisioni tra le aziende (si pensi ad esempio al concetto di *the winner takes it all*[7] per cui nella *digital economy* il leader di un settore tende ad avere la grandissima maggioranza delle quote di mercato) e tra le persone (il *digital divide*[8], vale a dire il divario nelle possibilità e opportunità tra chi ha accesso e conosce l'utilizzo delle tecnologie e chi no). L'azione del regolatore, dunque, deve essere attenta a mitigare per quanto possibile le divisioni stimolando al contempo l'accesso alle nuove opportunità *digital-driven*.

L'Unione Europea ha certamente raccolto la sfida affrontandola in maniera puntuale. La proposta di regolamento europeo del 29 aprile 2021[9], infatti, è solamente il punto di arrivo di un percorso a step e di un'azione strutturata di studio e di *policy making* da parte dell'Unione. D'altra parte, quest'azione era stata ampliamente annunciata. In "*Un'Unione più ambiziosa – il mio programma per l'Europa* [10] " l'allora candidata alla carica di presidente della Commissione Europea (poi eletta) Ursula Von der Leyen ha annunciato, che, in caso fosse risultata vincitrice, la Commissione avrebbe presentato una normativa per un approccio europeo coordinato alle implicazioni umane ed etiche dell'Intelligenza Artificiale (IA). La promessa è stata mantenuta.

Durante la presidenza Von der Leyen, la prima azione nella direzione di un regolamento condiviso tra gli stati membri è stata, infatti, l'adozione delle linee guida etiche definite nel testo

[7] Bughin, Jacques, Tanguy Catlin, Martin Hirt, and Paul Willmott. "*Why digital strategies fail*". McKinsey Quarterly (2018)

[8] Van Dijk, Jan. *The digital divide*. John Wiley & Sons, 2020.

[9] European Commission, "*Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts*". 29.4.2021. European Union (2021). Disponibile in: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021PC0206&from=EN

[10] Ursula Von der Leyen, "*Un'Unione più ambiziosa – il mio programma per l'Europa*". European Union (2019).

"*Orientamenti etici per un'IA affidabile*[11]" (8 aprile 2019) frutto del lavoro del gruppo di esperti e la pubblicazione de "*Il Libro bianco sull'intelligenza artificiale - Un approccio europeo all'eccellenza e alla fiducia*[12]" (19 febbraio 2020).

Questi testi si aggiungono alla precedente regolamentazione nell'ambito della tutela del cittadino verso la tecnologia tra cui è opportuno annoverare i meno specifici, ma altrettanto importanti, *Direttiva sui machinery products*[13] (che copre i nuovi rischi derivanti dalle nuove tecnologie digitali e si concentra sull'integrazione sicura dei sistemi di IA nei macchinari), gli emendamenti alla *General Product Safety Directive*[14] e alla *Radio Equipment Directive*[15] (per renderle utili alle nuove problematiche poste dalla *digital era*) e la

[11] European Commission, Directorate-General for Communications Networks, Content and Technology, "*Ethics guidelines for trustworthy AI*", Publications Office, 2019, https://data.europa.eu/doi/10.2759/177365

[12] European Commission, "*White paper on Artificial Intelligence - A European approach to excellence and trust*", 19.2.2020, COM (2020). Disponibile in: https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

[13] European Parliament, "*Directive 2006/42/EC of the European Parliament and of the council on machinery and amending Directive 95/16/EC (recast)"*. 17.5.2016, Official Journal of the European Union (2006). Disponibile in: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32006L0042&from=EN

[14] European Parliament, "*Directive 2001/95/EC of the European Parliament and of the council on general product safety*". 3.12.2001. Official Journal of the European Union (2001). Disponibile in: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32001L0095&from=EN

[15] European Parliament, "*Directive 2014/53/EU of the European Parliament and of the council on the harmonisation of the laws of the Member States relating to the making available on the market of radio equipment and repealing Directive 1999/5/EC*". 16.4.2014. Official Journal of the European Union (2014). Disponibile in: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014L0053&from=IT

General Data Protection Regulation[16] (regolamento per la tutela dei dati e della privacy)[17].

In questo articolo verranno suggerite delle possibili conseguenze legate all'interpretazione dei regolamenti proposti e delle possibili soluzioni al problema dell'*over-regulation*.

2. Possibili impatti organizzativi della proposta

2.1. Contesto di riferimento

La *technology penetration* nelle imprese e nella popolazione residente è certamente un fenomeno in grande crescita. Chiunque, sia per ragioni professionali che personali, quotidianamente dedica grande parte del tempo a sfruttare servizi legati alla connettività digitale (e.g. inviare e-mail, leggere notizie online, frequentare social networks, comprare o vendere azioni in borsa) e si confronta con applicazioni *AI-enabled* (e.g. ricerca su Google, auto-completamento del testo, applicazioni di filtri in Instagram). Chiaramente, però, la *technology penetration* è disomogenea all'interno di una stessa geografia (e.g. diversa per regione, per grandi e piccoli centri abitati, per fasce d'età, per livello d'istruzione, ecc.) ed è ancora più disomogenea se si confrontano geografie diverse.

In aggiunta a ciò, la percezione e l'accettazione della tecnologia nelle popolazioni culturalmente diverse varia di molto e, dunque, la definizione di principi condivisi da tutti diventa un tema particolarmente complesso da definire e da gestire. Si consideri, a riprova di ciò, l'esperimento effettuato da MIT Media Lab chiamato

[16] European Parliament, "*Regulation 2016/679 of the European Parliament and of the council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*". 27.4.2016. Official Journal of the European Union (2016). Disponibile in: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN

[17] Vjollca Kopsaj (ed.), "*Problemi di filosofia pratica 2021*". Pavia, Printservice editore (2021).

The Moral Machine[18]. Questo esperimento propone a chiunque voglia partecipare alcune scelte, che sottintendono dei quesiti etici, rispetto alle decisioni che una macchina a guida autonoma dovrebbe prendere in caso di incidente inevitabile tramite un semplice sito web [19] (l'esperimento è una evoluzione de facto del più tradizionale *trolley problem* [20]). Il risultato osservato è che ci sono delle preferenze globalmente accettate; per la grandissima parte delle decisioni, però, è possibile identificare dei *cultural clusters* che suggeriscono dunque che all'interno di una singola macro-regione i principi etici possano trovare condivisione, ma non è possibile definire un'etica globale e accettata da tutti.

La prima grande sfida dell'Unione Europea è stata dunque quella di definire dei principi etici che fossero condivisi e condivisibili dai paesi membri e che potessero essere la fonte cui ispirare i regolamenti successivi. L'attività dei *policy maker* nel regolare la tecnica e la tecnologia, infatti, deve necessariamente basarsi su dei principi generali stabili e dei successivi regolamenti legati all'applicazione che variano al variare dell'evoluzione delle possibilità e dei rischi connessi al maturare e l'affermarsi delle diverse tecnologie.

La scelta dei principi etici non è una scelta semplice: le conseguenze strategiche di proporre un atteggiamento a tutela dell'impresa o del consumatore, di puntare su un mercato iper-regolato o di scegliere per un atteggiamento *laissez faire*, possono impattare in maniera fondamentale sulla capacità di innovare, sul consumo e possono dunque spostare gli equilibri su scala geo-politica.

L'*AI Act*, dunque, viene proposto con l'obbiettivo di non imbrigliare la ricerca e di non limitare l'imprenditoria, ma vuole dare una direzione identitaria europea rispetto all'uso della tecnologia e rispetto alle direzioni scientifiche e dell'innovazione in cui il valore della *human dignity* viene tutelato e preferito al valore della strategia di crescita a controllo statale o della pura libertà di impresa. Nello specifico, gli obbiettivi dell*'Artificial Intelligence* Act sono:

---

[18] Awad, Edmond, et al. "*The moral machine experiment*." Nature 563.7729 (2018): 59-64.

[19] https://www.moralmachine.net/

[20] Thomson, Judith Jarvis. "*Killing, letting die, and the trolley problem*." The Monist 59.2 (1976): 204-217.

(1) assicurare che i sistemi di IA immessi sul mercato dell'Unione e utilizzati siano sicuri e rispettino tutte le normative vigenti in materia di diritti fondamentali e i valori proposti dall'Unione;

(2) assicurare la certezza del diritto per stimolare gli investimenti e l'innovazione nell'IA;

(3) migliorare la governance e l'applicazione effettiva della normativa esistente in materia di diritti fondamentali e requisiti di sicurezza applicabili ai sistemi di IA;

(4) facilitare lo sviluppo di un mercato unico per applicazioni di IA lecite, sicure e affidabili nonché prevenire la frammentazione del mercato[21].

2.2. Principi etici e requisiti fondamentali

Per dare un'interpretazione attenta dei possibili effetti ed impatti socio-economici dell'*AI Act* è opportuno riassumere brevemente i principi etici ed i requisiti fondamentali raccolti nel documento *Orientamenti etici per un'IA affidabile*[22]. Questo documento, infatti, è, *de facto*, un tentativo da parte dell'Unione di formalizzare le *fonti-fatto* perimetrando i principi etici condivisi nell'approccio giusto con la tecnologia, di seguito brevemente riportati:

(1) ***rispetto dell'autonomia umana***: Gli esseri umani che interagiscono con i sistemi di IA devono poter mantenere la propria piena ed effettiva autodeterminazione e devono poter essere partecipi del processo democratico;

(2) ***prevenzione dei danni***: I sistemi di IA non devono causare danni né aggravarli e neppure influenzare negativamente gli esseri

[21] European Commission, "*Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts*". 29.4.2021. European Union (2021). Pag. 3.

[22] European Commission, Directorate-General for Communications Networks, Content and Technology, "*Ethics guidelines for trustworthy AI*", Publications Office, 2019. Pagg. 11-13.

umani, per cui occorre tutelare la dignità umana nonché l'integrità fisica e psichica;

(3) ***equità***: lo sviluppo, la distribuzione e l'utilizzo dei sistemi di IA devono essere equi, devono dunque garantire una distribuzione giusta ed equa di costi e di benefici e garantire che gli individui e i gruppi siano liberi da distorsioni inique, discriminazioni e stigmatizzazioni;

(4) ***esplicabilità***: tale principio implica che i processi devono essere trasparenti, le capacità e lo scopo dei sistemi di IA devono essere comunicati apertamente e le decisioni, per quanto possibile, devono poter essere spiegate a coloro che ne sono direttamente o indirettamente interessati[23].

Nello stesso documento, i principi etici proposti vengono tradotti in requisiti concreti che le applicazioni *IA-driven* devono rispettare. Questi requisiti sono di seguito brevemente riportati:

(1) ***intervento e sorveglianza umani***: i sistemi di IA dovrebbero sostenere l'autonomia e il processo decisionale degli operatori umani, come prescritto dal principio del *rispetto dell'autonomia umana*. A tal fine, questi sistemi devono agire come catalizzatori di una società democratica, prospera ed equa, sostenendo l'intervento degli utenti e promuovendo i diritti fondamentali, e devono consentire la sorveglianza umana.

(2) ***robustezza tecnica e sicurezza***: a tutela e promozione del principio di *prevenzione dei danni*, per garantire la robustezza tecnica è necessario che i sistemi di IA siano sviluppati con un approccio di prevenzione dei rischi. In questo senso viene richiesto che essi si comportino secondo le previsioni, riducendo al minimo i *"danni non intenzionali e imprevisti"* e prevenendo del tutto *"danni inaccettabili"*.

(3) ***riservatezza e governance dei dati***: sempre a tutela e promozione del principio di *prevenzione dei danni*, la riservatezza viene promossa dall'*AI Act* essendo un diritto fondamentale particolarmente interessato dall'IA stessa. Per prevenire danni alla riservatezza occorre tra l'altro un'adeguata governance dei dati che

[23] Ibid.

riguardi la qualità e l'integrità dei dati utilizzati, la loro pertinenza rispetto al settore in cui i sistemi di IA saranno distribuiti, i protocolli di accesso e la capacità di trattare i dati in modo da tutelare la riservatezza.

(4) ***trasparenza***: a tutela e promozione del *principio di esplicabilità*, *l'AI Act* richiede la trasparenza degli elementi tipici che costituiscono un sistema di IA: i dati, il sistema e i modelli di business.

(5) ***diversità, non discriminazione ed equità***: a tutela e promozione del *principio di equità*, uno dei requisiti per ottenere un'IA affidabile è quello di promuovere e consentire l'inclusione e la diversità durante l'intero ciclo di vita del sistema.

(6) ***benessere sociale e ambientale***: a tutela e promozione dei *principi di equità e di prevenzione dei danni*, è richiesto che anche la società in generale, gli esseri senzienti e l'ambiente siano considerati come portatori di interessi durante l'intero ciclo di vita del sistema di IA.

(7) ***accountability***: quest'ultimo requisito integra i precedenti sei requisiti ed è promosso a tutela del *principio di equità*, richiedendo che vengano messi in atto meccanismi a garanzia dell'accountability dei sistemi di IA e dei loro risultati, sia prima che dopo la loro attuazione[24].

2.3. Conseguenze organizzative

Ogni anno la Commissione Europea pubblica l'indice DESI (*The Digital Economy and Society Index*), un rapporto utile a riassumere gli indicatori sulle prestazioni digitali dell'Europa e a tracciare i progressi dei paesi dell'Unione [25]. Il rapporto del 2021 mostra un'Europa che sui temi digitali viaggia a due velocità: da una parte gli stati del nord Europa, che mostrano perfomance ottime rispetto a tutte le dimensioni di misura, dall'altra gli stati dell'Europa dell'est, a cui si sommano Portogallo ed Italia, che invece mostrano performance pessime in tutte (o quasi) le dimensioni di misura.

[24] Ibid. Pagg. 14-20.

[25] European Commission, "*The Digital Economy and Society Index"*, European Union (2021)

Nell'ultimo anno l'Italia ha mostrato un miglioramento importante rispetto all'anno precedente, ma rispetto alla dimensione del *human capital,* si classifica ancora terz'ultima in Europa e addirittura ultima per quanto riguarda le "*advanced skills*"[26]. Il risultato dell'analisi proposta suggerisce dunque che le imprese italiane abbiano urgente bisogno di promuovere azioni di trasformazione digitale e programmi di innovazione strutturati e che la diffusione di competenze digitali sia di base (*digital literacy*) che avanzate debba essere una priorità strategica.

Nel novembre del 2021, il Governo Italiano ha promosso il *Programma Strategico Intelligenza Artificiale 2022-2024*[27], in cui, concordemente alle esigenze evidenziate dall'indice DESI, viene programmato uno sforzo di investimento verso la ricerca e la formazione nell'ambito dell'Intelligenza Artificiale e vengono promosse attività di partenariati pubblico-privati al fine di stimolare il trasferimento tecnologico e l'adozione delle tecnologie avanzate in azienda.

Da un punto di vista del quadro normativo e della programmazione degli investimenti ci sono dunque due forze che spingono nelle direzioni opposte. Da un lato l'AI Act, che insieme alla GDPR propone dei limiti nell'utilizzo e nelle applicazioni in azienda dell'intelligenza artificiale. Queste limitazioni sono state pensate secondo un intervento legislativo orizzontale (i.e. *non-industry specific*) dall'Unione Europea che propone di limitare l'utilizzo seguendo un approccio proporzionato basato sul rischio, impedendo dunque l'utilizzo per situazioni *non tollerabili*, regolando le situazioni *ad alto rischio* e definendo dei codici di condotta per i sistemi di IA a rischio *medio e rischio basso*[28]. Dall'altro lato, il programma di investimenti nazionale[29] ed il programma di

[26] Ibid.

[27] Caputo, B. et al., "*Programma Strategico Intelligenza Artificiale 2022-2024",* Governo Italino (2021)

[28] European Commission, "*Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts*". 29.4.2021. European Union (2021). Pag. 12. Tit. 2.

[29] Piano Nazionale di Ripresa e Resilienza.
https://www.governo.it/sites/governo.it/files/PNRR.pdf

investimenti nella ricerca europeo (*Horizon Eurpe*[30]) puntano a recuperare posizioni sulla *digital literacy,* sulle *advanced skills* e sull'adozione di tecnologia nelle aziende.

Al di fuori del regolamento di tutela della privacy e trattamento dei dati personali (GDPR)[31], infatti, questo regolamento per la prima volta pone delle questioni di merito su quando sia o non sia lecito sviluppare sistemi informatici intelligenti e sia o non sia lecito sperimentare algoritmi intelligenti che affronti un certo problema. In aggiunta alle considerazioni teoriche sui sistemi di controllo e sulle eventuali limitazioni applicative valutate sulla base del rischio, infatti, il documento elenca negli allegati esempi e ambiti applicativi che devono essere sempre considerati ad alto rischio e dunque tratti come tali[32].

In un momento in cui l'azione del governo italiano e di diversi governi europei è rivolta verso un'accelerazione nell'adozione della tecnologia, questo tipo di approccio regolatorio, se non opportunamente governato, può diventare un freno alla diffusione della tecnologia stessa e alle opportunità ad essa associate.

Non solo. Dal punto di vista della gestione dell'attività di ricerca o dell'attività di impresa, il regolamento pone due questioni fondamentali. In primo luogo, è necessario che le caratteristiche tecniche garantiscano il soddisfacimento dei principi di *equità* e *prevenzione dei danni*. In secondo luogo, è necessaria una valutazione *ex-ante* del rischio in funzione dei requisiti fondamentali precedentemente elencati.

---

[30] https://horizoneurope.apre.it/

[31] European Parliament, "*Regulation 2016/679 of the European Parliament and of the council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*". 27.4.2016. Official Journal of the European Union (2016). Disponibile in: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN

[32] European Commission, "*Annexes to the proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts*". 29.4.2021. European Union (2021). Annex III https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_2&format=PDF

Per quanto concerne la valutazione dei rischi *ex-ante*, la proposta di regolamento recita quanto segue:

> *"Title III contains specific rules for AI systems that create a high risk to the health and safety or fundamental rights of natural persons. In line with a risk-based approach, those high-risk AI systems are permitted on the European market subject to compliance with certain mandatory requirements and an ex-ante conformity assessment."*[33]

Nella misura in cui, dunque, gli stakeholder vogliano industrializzare un servizio o un prodotto abilitato dall'intelligenza artificiale, la prima azione necessaria, prima ancora di cominciare con l'implementazione, sarà stimarne il rischio per assicurarsi che l'applicazione non ricada nella categoria dei sistemi ad alto rischio. Questa analisi non è un'analisi semplice e molte imprese possono avere necessità di chiedere una consulenza esterna. Qualora poi venga stabilito che il sistema di intelligenza artificiale possa generare degli alti rischi sarà necessario un ulteriore parere di conformità, sempre *ex-ante*.

Nonostante lo sforzo da parte del regolatore di chiarire quanto più possibile cosa rientri e cosa non rientri nelle applicazioni ad alto rischio, anche a causa della rapida evoluzione della tecnologia e della applicazione su vasta scala e nei contesti più diversi dell'intelligenza artificiale possibili situazioni dubbie e meritevoli di approfondimento possono verificarsi tanto che nell'introduzione alla proposta di regolamento viene specificato come gran parte degli stakeholder abbiano commentato le bozze del regolamento stesso chiedendo chiarificazioni rispetto alla definizioni (tra le altre) di rischio e di alto rischio:

[33] European Commission, "*Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts*". 29.4.2021. European Union (2021). Pag. 13

> *"Stakeholders mostly requested a narrow, clear and precise definition for AI. Stakeholders also highlighted that besides the clarification of the term of AI, it is important to define 'risk', 'high-risk', 'low-risk', 'remote biometric identification' and 'harm'.*[34]*"*

Da un punto di vista tecnico, il soddisfacimento dei principi di *equità* e di *prevenzione dei danni* è tutt'altro che banale. A titolo esemplificativo, la professoressa Fei-Fei Li, direttore dell'Institute for Human Centered Artificial Intelligence dell'università di Stanford, scrive che identificare i *bias* nei dati che alimentano gli algoritmi è alle volte impossibile[35] e, con l'aumentare della complessità dei modelli di descrizione dei dati e dei modelli predittivi, eliminare i *bias* diventa consistentemente più difficile (con conseguente violazione del principio di *equità* e del requisito di *non-discriminazione*). C'è di più. La presenza o meno dei *bias* nei dati può certamente portare a dei *bias* negli *outcomes* degli algoritmi, ma questi *bias* non necessariamente sono dannosi per la dignità dell'uomo o violano i principi etici. La valutazione del rischio di violazione di questi principi dunque è un'operazione estremamente complessa e fare sì che il rischio di utilizzo sia valutato con precisione implicherebbe che le aziende debbano dotarsi internamente (o della consulenza) di organismi notificati di valutazione, che, similmente ai comitati etici degli istituti socio-sanitari, dovranno essere chiamati a esprimersi sulla conformità o meno di ogni singolo caso d'uso e dovranno essere composti da eticisti e da ingegneri esperti della materia.

Questa ricaduta da un punto di vista organizzativo porterebbe a costi aggiuntivi e tempi di studio delle applicazioni dilatati e non compatibili con i tempi tipici di implementazione della tecnologia e dell'innovazione e che potrebbe quindi costituire un fattore di resistenza alla urgente *technology adoption*.

---

[34] Ibid. pag. 9

[35] Adeli, Ehsan, et al. "*Representation learning with statistical independence to mitigate bias.*" Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2021.

È infatti noto che tra i fattori che determinano in maniera più rilevante la resistenza al cambiamento (in questo caso la *digital transformation*) vi sono, oltre al costo di gestione del cambiamento, *misunderstaning* (confusione e complicazione nel capire cosa si sta cambiando, come e perché), *fear* (paura delle conseguenze del cambiamento e paura del fallimento), *poor training* (scarsa preparazione degli stakeholders e scarsa conoscenza della materia)[36].

In ultimo, il rischio legato all'*outcome* deve essere confrontato con il rischio ambientale che dipende dalle condizioni al contorno. A titolo esemplificativo, l'AI Act non consente ai sistemi intelligenti di decidere se consentire o meno l'accesso a un servizio sanitario (e.g. triage)[37], ma, nella misura in cui dovesse esserci un'ulteriore ondata pandemica e un triage automatico consentisse di alleviare lo stress sul SSN, sarebbe utile avere a disposizione un sistema funzionante già testato anche se mai messo in produzione.

Una soluzione possibile è un approccio analogo ai *Regulatory Sandboxes*[38] per i servizi *fintech*, che consistono in un ambiente controllato in cui è possibile testare, per un periodo di tempo limitato, prodotti e servizi tecnologicamente innovativi sotto la supervisione dell'autorità di vigilanza. Il titolo V dell'AI Act prevede ed incoraggia le autorità nazionali alla possibilità di costruire dei *regulatory sandboxes* in cui imprese e istituti di ricerca dei singoli stati membri possano avere uno spazio protetto e che sia una utile misura di sostegno all'innovazione.

Lo sviluppo di questo tipo di iniziative metterebbe le aziende nelle condizioni di cominciare subito la sperimentazione stimolando dunque l'attività di ricerca e di trasferimento tecnologico, identificando problemi ed ostacoli in poco tempo, consentendo uno studio di fattibilità tecnica più rapido e basato sul *fast-prototyping* e

[36] Dent, Eric B., and Susan Galloway Goldberg. "*Challenging "re-sistance to change".*" The Journal of applied behavioral science 35.1 (1999): 25-41.

[37] European Commission, "*Annexes to the proposal for a regulation of the Euro-pean Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts*". 29.4.2021. European Union (2021). Annex III. P.to 5.a

[38] Zetzsche, Dirk A., et al. "*Regulating a revolution: From regulatory sandboxes to smart regulation.*" Fordham J. Corp. & Fin. L. 23 (2017): 31.

consentendo all'autorità di vigilanza di interrompere o delimitare il perimetro e l'applicazione del caso d'uso in esame in tempi ridotti.

Stabilire se per il settore *fintech* lo strumento del *regulatory sandbox* proposto da Banca d'Italia è stato o meno un successo è ancora prematuro essendo la prima finestra per l'invio delle candidature d'accesso ancora aperta[39] (per una durata di apertura totale di 60 giorni) e dovendo aspettare dalla chiusura della stessa che si esaurisca la fase di istruttoria (45 giorni), la pubblicazione degli esiti (21 giorni) e la fine della sperimentazione (fino a 18 mesi)[40].

È opinione concorde di molti portatori di interesse[41] che anche per la sperimentazione sull'intelligenza artificiale i *regulatory sandboxes* possano essere una buona soluzione per sostenere la ricerca e l'innovazione. Certamente la varietà di casi d'uso, di ambiti applicativi e di settori industriali coinvolti oltre che il numero di aziende e di stakeholders propongono delle sfide da un punto di vista della gestione e del controllo, poiché dare un giudizio rispetto al funzionamento di un sistema di intelligenza artificiale necessita certamente conoscenza della tecnologia, ma anche, almeno in parte, conoscenza del settore industriale di riferimento.

3. Conclusioni

Di fronte alla diffusione di tutte le moderne tecnologie, i *policy makers* sono sempre posti di fronte al *trade-off* legato *all'over-regulation* e *all'under-regulation*.

Per tutelare i propri cittadini, la loro privacy e la loro dignità e a tutela delle infrastrutture informative e degli asset intangibili che sarebbero facile preda delle grandi potenze digitali internazionali, l'Unione Europea non può assumere un atteggiamento di *laissez faire*. L'indipendenza e la libertà dei cittadini dell'Unione, infatti, passa anche dal buon governo della traccia digitale che quotidianamente i

[39] https://www.bancaditalia.it/focus/sandbox/index.html

[40] D. M. 30 aprile 2021, n. 100, Ministero dello Sviluppo Economico

[41] European Commission, "*Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts*". 29.4.2021. European Union (2021). Pag. 9

cittadini lasciano utilizzando i comuni dispositivi elettronici. Questa traccia digitale non può essere strumento per limitare la libertà di pensiero e d'espressione di nessuno, né la base per costruire sistemi di *scoring* che limitino le libertà personali dei cittadini non allineati a chi controlla i loro dati[42].

Il rischio dell*'under regulation*, dunque, oltre che riferibile all'uso improprio dei dati dei cittadini, è anche legato alla perdita del governo e della proprietà del dato (banalmente per ragioni di tipo economico) e la perdita del controllo delle infrastrutture di scambi, di elaborazione e di storage dei dati che, oggi, costituiscono un asset strategico per ogni paese.

D'altra parte, *l'over-regulation* non consentirebbe agli Stati Membri che sul tema della *digital adoption* fanno più fatica – tra cui l'Italia – di evolvere e di recuperare il terreno perduto, né alle imprese di innovare ed essere competitive nel mercato globale.

L'AI Act, pur essendo in linea con la linea dell'Unione Europea di tutela dei diritti e della dignità dei suoi cittadini, dunque, se non viene attuato in maniera attenta, può costituire un freno al progresso e alla ricerca negli Stati Membri.

Similmente a quanto già successo nell'industria *fintech*, la possibilità di *regulatory sandboxes* come stimolo per la ricerca e come strumento per sperimentare applicazioni nuove con rischi d'impresa mitigati deve essere una soluzione da tenersi in considerazione. Questa soluzione, infatti, permetterebbe alle imprese di innovare e sperimentare pur muovendosi all'interno di quadri normativi complessi e, a tratti, restrittivi. D'altra parte, quando si tratta con tecnologie avanzate, conoscere gli effetti (positivi o negativi che siano), l'efficacia, l'utilità di un'applicazione senza sperimentarla è spesso molto complesso e anteporre dunque la sperimentazione a considerazioni di rischio ed efficacia può, nella grande maggior parte dei casi, essere tecnicamente la soluzione preferibile.

I *regulatory sandboxes* sono inseriti all'interno dell'AI Act come strumento suggerito agli Stati Membri, ma non imposto. La formulazione attuativa efficace, efficiente e flessibile di questo

[42] Dai, Xin. "*Toward a reputation state: The social credit system pro-ject of China.*" Available at SSRN 3193577 (2018).

strumento deve essere considerata una priorità strategica per abbassare le resistenze del tessuto industriale del paese alla *technology adoption* e per rendere più vivace l'ecosistema dell'innovazione che vi ruota attorno.

4. Ringraziamenti

Questo lavoro è stato parzialmente supportato dal Ministero dell'Università e della Ricerca italiano (MUR) attraverso il progetto PRIN ``XFAST-SIMS: Simulazione extra veloce e accurata di sistemi strutturali complessi" (No. 20173C478N).

Bibliografia

Adeli, Ehsan, et al. "Representation learning with statistical independence to mitigate bias." Proceedings of the IEEE/CVF Winter Conference on Applica-tions of Computer Vision. 2021.

Awad, Edmond, et al. "*The moral machine experiment*." Nature 563.7729 (2018): 59-64.

Bughin, Jacques, Tanguy Catlin, Martin Hirt, and Paul Willmott. "*Why digital strategies fail*". McKinsey Quarterly (2018)

Caputo, B. et al., "*Programma Strategico Intelligenza Artificiale 2022-2024*", Governo Italino (2021)

Cassard, Anita, and Joseph Hamel. "*Exponential Growth of Technology and the Impact on Economic Jobs and Teachings: Change by Assimilation*." Journal of Applied Business & Economics 20.2 (2018).

Dai, Xin. "*Toward a reputation state: The social credit system project of China.*" Available at SSRN 3193577 (2018).

Dent, Eric B., and Susan Galloway Goldberg. "*Challenging "resistance to change".*" The Journal of applied behavioral science 35.1 (1999): 25-41.

European Commission, "*Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts*". 29.4.2021. European Union (2021). Disponibile in:

https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021PC0206&from=EN

European Commission, Directorate-General for Communications Networks, Content and Technology, "*Ethics guidelines for trustworthy AI*", Publications Office, 2019,

https://data.europa.eu/doi/10.2759/177365

European Commission, "*White paper on Artificial Intelligence - A European approach to excellence and trust*", 19.2.2020, COM (2020). Disponibile in:

https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

European Parliament, "*Directive 2006/42/EC of the European Parliament and of the council on machinery and amending Directive 95/16/EC (re-cast)*". 17.5.2016, Official Journal of the European Union (2006). Disponibile in:

https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32006L0042&from=EN

European Parliament, "*Directive 2001/95/EC of the European Parliament and of the council on general product safety*". 3.12.2001. Official Journal of the European Union (2001). Disponibile in:

https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32001L0095&from=EN

European Parliament, "*Directive 2014/53/EU of the European Parliament and of the council on the harmonisation of the laws of the Member States relating to the making available on the market of radio equipment and repealing Directive 1999/5/EC*". 16.4.2014. Official Journal of the European Union (2014). Disponibile in:

https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014L0053&from=IT

European Parliament, "*Regulation 2016/679 of the European Parliament and of the council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*". 27.4.2016. Official Journal of the European Union (2016). Disponibile in:

https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN

European Commission, "*The Digital Economy and Society Index*", European Union (2021)

Jones, Steven E. "*Against technology: From the Luddites to neo-Luddism*". Routledge, 2013.

Kaczynski, Theodore John. "*Industrial society and its future.*" (1995).

Marseglia, G. Roberto, Dal Mas, Francesca, Massaro, Maurizio, Bagnoli, Carlo; "*L'Artificial Intelligence Act: risvolti pratici dell'etica dell'Intelligenza Arificiale*"; in Vjollca Kopsaj (ed.),

"Problemi di filosofia pratica 2021". Pavia, Printservice editore (2021); ISBN: 9788898765980.

Popkova, Elena G., Yulia V. Ragulina, and Aleksei V. Bogoviz. "*Fundamental differences of transition to industry 4.0 from previous industrial revolutions*." Industry 4.0: Industrial Revolution of the 21st Century. Springer, Cham, 2019. 21-29.

Senate of the United States, "*Clarifying Lawful Overseas Use of Data (CLOUD) Act*", 2018.

Taal, Amie, ed. "*The GDPR Challenge: Privacy, Technology, and Compliance in an Age of Accelerating Change.*" CRC Press, 2021.

Thomson, Judith Jarvis. "*Killing, letting die, and the trolley problem*." The Monist 59.2 (1976): 204-217.

Ursula Von der Leyen, "*Un'Unione più ambiziosa – il mio programma per l'Europa*". European Union, 2019.

Van Dijk, Jan. "*The digital divide*". John Wiley & Sons, 2020.

Zetzsche, Dirk A., et al. "*Regulating a revolution: From regulatory sandboxes to smart regulation*." Fordham J. Corp. & Fin. L. 23 (2017): 31.